

DEVELOPMENT OF LOW TEMPERATURE BIOLOGICAL SMALL ANGLE X-RAY
SCATTERING, CRYOSAXS, AND STUDY OF ICE FORMATION IN AQUEOUS
SOLUTIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jesse Bennett Hopkins

February 2016

© 2016 Jesse Bennett Hopkins

DEVELOPMENT OF LOW TEMPERATURE BIOLOGICAL SMALL ANGLE X-RAY
SCATTERING, CRYOSAXS, AND STUDY OF ICE FORMATION IN AQUEOUS
SOLUTIONS

Jesse Bennett Hopkins, Ph. D.

Cornell University 2016

This dissertation is based on the study of biological systems and aqueous solutions at low temperatures. The focus of the dissertation is on the continued development of low temperature biological small angle X-ray scattering (SAXS), cryoSAXS. CryoSAXS is a new method for biological SAXS that involves cryocooling the sample to 100 K. At 100 K the sample is much more resistant to damage from X-rays, enabling use of volumes hundreds or thousands of times smaller than in conventional room temperature SAXS. This work was divided into three main pieces. First, we designed, microfabricated, and tested new small volume, fixed path length sample holder for use in cryoSAXS. Next, as a first step towards answering the question of how much the radiation tolerance of a sample increases upon cryocooling, we developed methods to quantify radiation damage to SAXS profiles and tested these using standard room temperature SAXS. Finally, we measured the radiation tolerance of cryocooled SAXS samples, and found that the cryocooling decreases the radiation sensitivity by three orders of magnitude. The last chapter of the dissertation describes ongoing work to develop time resolved cryoSAXS, and other outstanding questions in cryoSAXS.

In addition to cryoSAXS, in a complementary research project we studied ice formation in aqueous solutions. In particular, we quantified the critical warming rate, the warming rate required to prevent ice formation upon warming a vitrified solution. This has particular relevance to biological cryopreservation, as ice formation can irreparably damage cryopreserved samples.

BIOGRAPHICAL SKETCH

Jesse was born in 1987 in Portland Oregon. After a number of intervening years full of interesting yet mostly irrelevant details, in 2009 he received his B.A. in Physics and Mathematics from Oberlin College. He was awarded Highest Honors in physics, the Carl E. Howe prize in Physics, which is given to the top graduating physics student as judged by academic and research achievements, and was inducted into the Phi Beta Kappa and Sigma-Xi Honors Societies. His undergraduate research was the study of materials for hydrogen storage using infrared spectroscopy in the lab of Stephen Fitzgerald. He has done his graduate work in Physics at Cornell University in lab of Robert Thorne, receiving several honors, most recently the Margaret C. Etter Student Lecture Award (May 2014), awarded for the best student lectures presented at the American Crystallographic Association annual meeting. When he is not doing physics, he hikes, plays with his cats, reads science fiction and fantasy books, plays board games, and writes autobiographical sketches in the third person.

ACKNOWLEDGMENTS

There are more people to thank than I can or will remember. This work could not have been completed without my advisor, Robert Thorne, giving me advice, mentoring, feedback, and the freedom to pursue science I thought might be interesting. I also owe a great debt to Matthew Warkentin, a postdoc who took a young graduate student under his wing and offered training, mentoring, and friendship. Andrea Katz has been my steadfast collaborator for the past three years, and much of this work would not have been completed without her insight, fortitude, and friendship. Steve Meisburger provided my initial SAXS training, and has become a collaborator and friend. Lois Pollack, co-PI on the cryoSAXS project, has provided key insights and encouragement. Numerous other members of the Thorne and Pollack groups, as well as scientists and support staff at CHESS, MacCHESS, CNF, and BioCAT helped me along the way.

There are two other key people in my physics life. One is my undergraduate advisor and mentor, Stephen Fitzgerald, who taught me so much about physics and research. He remains an example to me of how to do and think about physics. The other is my high school physics teacher, Robert Schlichting. His boundless enthusiasm and love for physics, in fact for any science, and his insightful, engaging teaching helped me take my first few steps down this path.

I benefitted greatly from many wonderful friends who helped keep me sane and grounded throughout the ups and downs of graduate school. Whether it was indulging in our love of board games, brunches, hiking, Game of Thrones, or the ritualistic grilling of meat and vegetables, you were always there for me. I won't list you all here, but you know who you are, and how important you've been.

My family has always been supportive of my academic choices, something I greatly appreciate. Thank you for that, for your love, and, for feeding and housing me for 18+ years, helping me pay

for college, and other such small details.

Finally, I want to thank my wonderful partner Aryn Taylor, who has had to put up with me, including the odd hours, the late nights, and the stress, throughout the defeats and the victories of research life. Your support means the world to me. Also, thanks for packing all of those boxes.

I also need to acknowledge the various institutions where work took place, and funding agencies.

This work was funded by the National Science Foundation (DBI-1152348) and the National Institutes of Health (NIH) (GM065981-05A1). This work was performed in part at CNF, a member of the National Nanotechnology Infrastructure Network, which is supported by the National Science Foundation (grant No. ECCS-0335765). This work is based in part upon research conducted at 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 MacCHESS facility, which is supported by award GM-103485 from the National Institute of General Medical Sciences, National Institutes of Health. This work made use of the Nanobiotechnology Center shared research facilities at Cornell. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. This project was supported by grant 9 P41 GM103622 from the National Institute of General Medical Sciences of the National Institutes of Health. Use of the Pilatus 3 1M detector was provided by grant 1S10OD018090-01 from NIGMS. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

TABLE OF CONTENTS

Biographical sketch	iii
Acknowledgements	v
Table of contents	vii
1. Introduction	1
1.1 Overview and background	1
1.2 An introduction to small angle X-ray scattering	4
1.2.1 The basics	4
1.2.2 Data processing	7
1.3 Summary of initial cryoSAXS work	11
1.3.1 Achieving reproducible cryocooling for SAXS	11
1.3.2 Results from macromolecules	14
1.3.3 Summary and next steps	16
1.4 Structure of this dissertation	17
References	21
2. A microfabricated fixed path length silicon sample holder improves background subtraction for cryoSAXS	24
2.1 Preface	24
2.2 Introduction	24
2.3 Sample holder design, fabrication, and assessment	27
2.3.1 Design	27
2.3.2 Wafer processing	30
2.3.3 Evaluation of etched features	31
2.4 Methods for SAXS	32
2.4.1 Sample preparation	32
2.4.2 Beamline setup	37

2.4.3	Data collection at 100 K	39
2.4.4	Data collection at room temperature	40
2.4.5	Data processing	41
2.5	Results and analysis	42
2.5.1	Characterization of sample holders	42
2.5.2	CryoSAXS of glucose isomerase	45
2.5.3	Comparison to theoretical scattering curves	50
2.5.4	Repeatability	54
2.5.5	Origin of cooling-dependent variations in data below $q \sim 0.02 \text{ \AA}^{-1}$	54
2.5.6	Attempts to mitigate fracturing	60
2.6	Conclusion	61
2.7	Supporting information	62
2.7.1	Efforts to mitigate fractures	62
	References	65
3.	Quantifying radiation damage in biological small angle x-ray scattering	67
3.1	Preface	67
3.2	Introduction	67
3.3	Overview of a radiation damage experiment	70
3.4	Definition of dose	70
3.5	Calibration of experiment and sample parameters	71
3.6	Measurement of radiation damage data	75
3.7	Calculation of dose	77
3.7.1	Determining the flux density	77
3.7.2	Determining the absorbed energy fraction	78
3.7.3	Correcting for diffusion turnover	79
3.8	Quantifying radiation damage data	82
3.8.1	Calculate parameters from scattering profiles, normalize	82

3.8.2 Calculate sensitivities from fits to dose curves	85
3.8.3 Which sensitivities are important?	85
3.8.4 Quantification in a nutshell	92
3.9 Additional considerations	92
3.9.1 Buffer composition, temperature, macromolecule concentration, degassing	92
3.9.2 X-ray energy and dose rate	94
3.9.3 Beam heating	94
3.9.4 Critical Dose	95
3.10 Summary	98
3.11 Supporting information	99
3.11.1 Methods	99
3.11.1.1 Sample preparation	99
3.11.1.2 Beamline setup	100
3.11.1.3 Data processing and metrics for radiation damage	101
3.11.2 The attenuation coefficient	104
3.11.3 The diffusion correction	107
3.11.4 Effect of radiation on lysozyme, xylanase, and glucose isomerase	110
3.11.4.1 Measured Sensitivities	110
3.11.4.2 Differences between proteins	113
3.11.4.3 Effect of concentration	118
3.11.4.4 Other parameters affecting damage rates in SAXS	121
3.11.5 Correlation between metrics	123
3.11.6 Beam heating calculations	125
3.11.7 Rate of oxygen absorption in degassed samples	129
References	131
4. CryoSAXS increases radiation tolerance by orders of magnitude	136

4.1 Preface	136
4.2 Introduction	136
4.3 Methods	139
4.3.1 Sample preparation	139
4.3.2 Beamline setup and data collection	140
4.3.3 Calculating dose	143
4.3.4 Data processing	143
4.4 Results	147
4.4.1 Damage to the sample holders	147
4.4.2 Damage at 100 K	147
4.4.3 Damage at room temperature	162
4.5 Discussion	170
4.5.1 Comparison between room temperature and 100 K sensitivities	170
4.5.2 Origin of damage at 100 K	173
4.5.3 Comparison to previous results	180
4.5.4 Comparison to expected results	182
4.5.5 Room temperature damage	183
4.5.6 Implications for minimum sample volume	186
4.6 Conclusion	191
4.7 Supporting information	192
4.7.1 Density measurement	192
4.7.2 Illuminated volume	195
4.7.3 Lysozyme data	203
4.7.3.1 Additional methods	203
4.7.3.2 Additional results	204
4.7.3.2.1 100 K	204
4.7.3.2.2 Room temperature	211

4.7.4 Additional discussion	214
4.7.4.1 Comparison between room temperature and 100 K rates	214
4.7.4.2 Comparison between samples and buffers and to previous results	215
4.7.5 Additional Conclusions	216
References	218
5. Effect of common cryoprotectants on critical warming rates and ice formation in aqueous solutions	223
5.1 Preface	223
5.2 Introduction	224
5.3 Materials and methods	226
5.3.1 Materials	226
5.3.2 Experimental apparatus for warming	226
5.3.3 Warming rate data collection	231
5.3.4 Critical cooling rates	233
5.3.5 Theoretical modeling	233
5.4 Results	239
5.4.1 Critical warming and cooling rate measurements	239
5.4.2 Model calculations	245
5.5 Discussion	250
5.5.1 Concentration dependence	250
5.5.2 Variation of critical warming rate with cooling rate	252
5.5.3 Excess of CWR over CCR	253
5.5.4 Predictions of critical warming rates based upon previous data	254
5.5.5 Predictions of critical warming and cooling rates based upon models of ice nucleation and growth	255
5.5.6 Application to biological cryopreservation	257
5.6 Conclusions	258

References	260
6. Current and future work on CryoSAXS	264
6.1 Current problems and possible solutions	264
6.1.1 Problems	264
6.1.2 In situ deposition and cooling as a possible solution	267
6.1.3 Improvements in the sample holders	273
6.1.4 In-vacuum cryoSAXS	274
6.2 Other questions	275
6.2.1 Time resolved cryoSAXS	275
6.2.2 Hydration shells and modeling scattering at 100 K	277
6.2.3 High profile target	278
6.2.4 Screening for proper buffer conditions	279
6.3 Summary of future work	281
References	282
Appendices	
A. Protocol for sample holder fabrication	284
B. SAXS data analysis program	294

CHAPTER 1

INTRODUCTION

1.1 Overview and background

Structural biology is the study of the molecular structure and dynamics of biological macromolecules, particularly proteins and nucleic acids, and how alterations in their structures affect their function. Structural biology incorporates the principles of molecular biology, biochemistry and biophysics.

- Nature.com Subject Area Definition (nature.com)

Structural biology has been enormously successful in helping to explain how and why the machinery of life works as it does, providing the framework on which most ideas in molecular biology are built and understood. It should be no surprise that major advances in structural biology are often linked with advances in methods for determining structure (Campbell, 2002). The workhorse technique of structural biology has been and continues to be X-ray crystallography, which provides atomic resolution structures of macromolecules. There are almost 110,000 macromolecular structures in the Protein Data Bank (PDB), and 89% of them have been determined by X-ray crystallography (rcsb.org). The other major high resolution structural technique is solution NMR spectroscopy, accounting for 10% of deposited PDB structures. Both of these techniques have their limitations: of the approximately 61,000 proteins that the Protein Structure Initiative TargetTrack database lists as purified, only 16% have structures deposited in the PDB (sbkb.org). For crystallography the major challenge lies in producing diffraction quality crystals: only 27% of the purified targets have been crystallized, 76% of those crystals were 'diffraction quality,' and 53% of the diffraction quality crystals yielded crystal structures (10% of the initial purified targets). Crystallography is also limited in the information it can reveal, as important structural motions, alternative conformations or intrinsically disordered regions may not be present in the crystal structure. For NMR, peak

overlap and broadening often limits the ability to solve structures for larger targets (Putnam et al., 2007; Marion, 2013; Skou et al., 2014). Cryo-Electron Microscopy is another structural technique that has just recently been able to achieve near atomic resolution structures for smaller molecules, and may become more widely used in the future (Bai et al., 2015).

Small angle X-ray scattering (SAXS) is a technique that can provide a wealth of structural information about macromolecules in solution with few restrictions on experimental conditions (Vestergaard & Sayers, 2014). SAXS can only provide relatively low (~nm) resolution, but the fact that it can do so for molecules in solution of almost any size makes it very powerful. Beyond basic structural information such as size, weight, and shape, it also provides insight, both static and time resolved, into complex formation, interconversion between allosteric states, folding and unfolding, mixtures, flexible systems, and other processes (Petoukhov & Svergun, 2013; Dyer et al., 2014; Kirby & Cowieson, 2014; Skou et al., 2014; Vestergaard & Sayers, 2014). Additionally, it can provide complementary information to aid in the application of high resolution techniques such as crystallography or NMR (Putnam et al., 2007; Grant et al., 2011; Chaudhuri, 2015), structural modeling including MD and protein and complex docking (Schneidman-Duhovny et al., 2012; Chen & Hub, 2014, 2015). Increasingly, it is used in conjunction with other biophysical assays such as size exclusion chromatography, multi angle light scattering, dynamic light scattering, and differential ultracentrifugation (Jacques & Trewhella, 2010; Bras et al., 2014; Acerbo et al., 2015; Hynson et al., 2015) to provide additional structural or biochemical information.

A successful SAXS experiment generally requires homogeneous, monodisperse samples in solution. This places requirements not just on sample preparation, but also upon data collection. The X-rays used to probe the molecules may also damage them, causing aggregation, unfolding,

conformational changes, and fragmentation, any of which can render the data unusable. Minimizing radiation damage requires spreading the incident X-rays out over a large volume, either by defocusing the X-ray beam at the sample (instead focusing on the detector) (Hura et al., 2009), flowing or oscillating the sample through the X-ray beam (Fischetti et al., 2003; Jeffries et al., 2015), or translating the sample cell in the beam (Hong & Hao, 2009). For this reason, SAXS experiments generally require 10-30 μL of sample, usually with a macromolecule concentration near 1 mg/mL (Dyer et al., 2014; Skou et al., 2014). Sensitivity can also vary by at least a factor of 20 between different samples, and the same sample with different buffer additives (Jeffries et al., 2015), making it hard to determine proper exposure times to maximize sample use without damage.

Radiation damage was significantly reduced for both macromolecular X-ray crystallography and electron microscopy by cryocooling samples to near 100 K (Dubochet et al., 1981; Hope, 1988; Holton, 2009; Bai et al., 2015). For crystallography, cooling to 100 K reduces damage rates by a factor of ~ 30 -50 for typical proteins (Warkentin et al., 2013a). We suspected that cryocooling could yield similar or larger gains for SAXS. This would allow measurement of scattering profiles that are difficult or impossible to obtain at room temperature due to sample limitations. Additionally, cryocooling would prevent time dependent sample changes, allowing simple measurements on unstable or aggregation prone samples and opening up the possibility for low volume time resolved SAXS. It may also enable more routine shipping and mounting for mail-in high-throughput data collection, similar to techniques used in cryocrystallography (Blundell et al., 2002). All of these reasons make cryoSAXS a very exciting area of development, and most of this dissertation is devoted to detailing our work to enable this technique.

1.2 An introduction to small angle X-ray scattering

This section will explain some of the basic principles of SAXS and the methods by which SAXS data is collected and processed. The goal is not to provide a comprehensive review of the topic, but to provide enough information that the data collection and processing described in this dissertation can be followed. For those seeking a more comprehensive overview of small angle scattering principles and methods, there are a number of excellent review articles and books (Glatter et al., 1982; Svergun & Koch, 2003; Jacques & Trehwella, 2010; Schneidman-Duhovny et al., 2012; Blanchet & Svergun, 2013; Svergun et al., 2013; Dyer et al., 2014; Skou et al., 2014).

1.2.1 The basics

At the most basic level, a SAXS experiment proceeds as follows. A sample is prepared, and put in an X-ray beam. The incident X-rays elastically scatter off of the sample, and that scattering is recorded on an X-ray detector. For biological SAXS, samples are typically biological macromolecules dissolved in a buffer solution. Standard experiments are carried out with monodisperse, homogeneous solutions of macromolecules. While polydisperse and heterogeneous solutions can be used, these typically require advanced processing or unusual experimental methods, and are not addressed here. Typical protein solutions are measured at concentrations near 1 mg/mL, in pH buffers with small amounts of salt added to stabilize inter-molecular interactions.

Figure 1.1 shows a schematic of a SAXS experiment, with an incident beam described by

$k_0 = |\vec{k}_0| = 2\pi / \lambda$. The scattered intensity has $k_1 = |\vec{k}_1| = 2\pi / \lambda$ (elastic scattering). The scattering

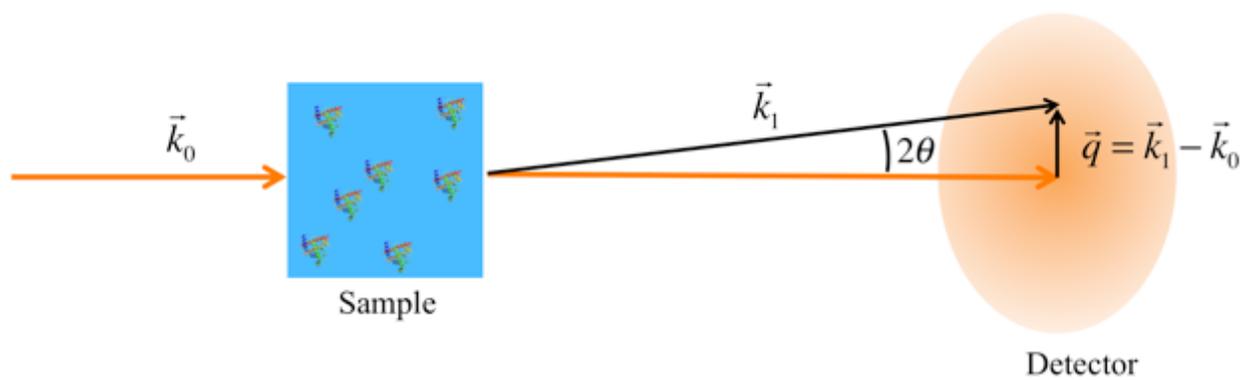


Figure 1.1 – A schematic of a small angle X-ray scattering experiment. X-rays with wavevector \vec{k}_0 illuminate the sample volume. These X-rays are scattered with wavevector \vec{k}_1 at an angle 2θ . The scattering vector is $\vec{q} = \vec{k}_1 - \vec{k}_0$, and in real space describes the position on the detector. As shown, the scattering pattern depends only on the distance from the incident beam, not the azimuthal angle.

vector can be defined as $\vec{q} = \vec{k}_1 - \vec{k}_0$. If the angle between the incident and scattered intensity is 2θ , as shown, the magnitude of the scattering vector is then

$$|\vec{q}| = q = \frac{4\pi \sin\theta}{\lambda} . \quad (1.1)$$

As shown in Figure 1.1, q is also the radial distance on the detector from the beam center.

Scattering in SAXS is treated in the first Born approximation, which means it is described by Fourier transformation from real space (laboratory) coordinates to reciprocal space coordinates (scattering vector) (Svergun & Koch, 2003). In a solution SAXS experiment, the scattering from the macromolecules and from the pure buffer are collected separately, and the latter is subtracted from the former. This yields a scattering profile where the signal comes from the contrast, the difference in electron density between the particle and the buffer. Of note is that macromolecules perturb a small layer of buffer near their surface, the hydration shell, increasing the density of this buffer (Svergun et al., 1998). As a result, the scattering pattern after buffer subtraction is from both the macromolecule and the hydration shell. The measured scattering intensity $I(\vec{q})$ is assumed to be from a homogenous solution of randomly oriented particles with uncorrelated positions and orientations, and thus should be proportional to the scattering from a single particle averaged over all orientations. This is given, in spherical coordinates, by

$$I(q) = 4\pi \int_0^{D_{max}} r^2 \gamma(r) \frac{\sin(qr)}{qr} dr , \quad (1.2)$$

where

$$\gamma(r) = \left\langle \int \Delta\rho(\vec{u}) \Delta\rho(\vec{u} + \vec{r}) d\vec{u} \right\rangle_\omega \quad (1.3)$$

is the spherically averaged autocorrelation of the excess electron density $\Delta\rho(\vec{r}) = \rho(\vec{r}) - \rho_s$, where $\rho(\vec{r})$ is both the particle and hydration layer density and ρ_s is the bulk solvent density. In Equation (1.2), D_{max} is the maximum dimension of the macromolecule plus hydration shell.

There are a few other basic experimental details that are worth mentioning. Occasionally the terms ‘upstream’ and ‘downstream’ will be used. These are directions relative to the sample: upstream means closer to the X-ray source (the synchrotron/injection device) along the X-ray beam, downstream means farther from the source (towards the detector). The measured scattering angles (the angle between the scattered intensity and the incident beam) are ~ 0.5 - 15° . In order to get good resolution across this range of angles, the X-ray detector is typically placed ~ 2 m downstream of the sample. Macromolecule solutions usually are weakly scattering, so a vacuum flight tube is typically placed between the sample and the detector, minimizing background scattering occurring after the sample. In order to further reduce background scatter, highly polished metal blades (in sets of 4, defining a rectangular aperture called a slit) are placed close to the incident beam upstream of the sample. These remove scattering from upstream windows, optics, and other components.

1.2.2 Data processing

All of the data in this dissertation was collected on area detectors, which give a two dimensional scattering pattern. However, the scattered intensity, Equation (1.2), is one dimensional. Given this, the first step in data processing is to convert the scattering patterns into one dimensional scattering profiles. The scattering vector, q , is calibrated by measuring the X-ray energy, the position of the incident X-ray beam relative to the detector, and the distance from the detector to

the sample. Once the calibration is taken, the two dimensional scattering patterns can be azimuthally integrated into one dimensional scattering profiles, yielding the measured $I(q)$.

With minimal further processing the scattering profile can yield information about the shape, flexibility, monodispersity, size, and molecular weight of the sample. The rate of intensity decay at high q gives information on the shape of the macromolecule. For example for spheres, such as globular proteins, $I(q) \propto q^{-4}$ at high q , whereas for rods, like short duplex DNA strands, $I(q) \propto q^{-3}$ (Svergun & Koch, 2003). More advanced interpretation using similar principles can inform on flexibility and foldedness of the macromolecule (Rambo & Tainer, 2011). At the lowest q , a homogeneous monodisperse sample (of globular protein) should have $I(q) \propto \exp(-c * q^2)$, where $c = 1/3R_g^2$ is a constant that depends on the radius of gyration, R_g , of the macromolecule.

Deviation from this behavior indicates a lack of homogeneity or monodispersity, typically either the presence of larger aggregates or interaction between the macromolecules, such as electrostatic repulsion. Fitting the low q region to this exponential form yields size information, and allows the interpolation of that fit to zero scattering angle, $I(0)$. This intensity at $q = 0$ is proportional to the molecular weight, and with proper calibration can be used to get an absolute molecular weight of the macromolecule. Particle volumes can also be obtained for the scattering profile (Rambo & Tainer, 2011).

The typical next step in data processing is to invert the Fourier transform and examine the real space pair distance distribution $p(r) = r^2\gamma$, with γ defined as in Equation (1.3). This is given by

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty q^2 I(q) \frac{\sin(qr)}{qr} dr . \quad (1.4)$$

As $I(q)$ is not known for all space, and is only sampled at finite intervals, directly carrying out the transformation in Equation (1.4) is not usually attempted. The standard method to find $p(r)$ is to use an indirect Fourier transform method, the most common of which is implemented by the GNOM program (Svergun, 1992; Petoukhov et al., 2012).

Once $p(r)$ is calculated it provides information about particle size, molecular weight, and particle shape. The particle size information comes in two ways. The $p(r)$ function represents the distribution of interparticle dimensions, so where it goes to zero gives the maximum dimension of the particle, D_{max} . The second piece of size information comes from the fact that the R_g is the normalized second moment of the distance distribution of the particle around the center of its scattering length density distribution

$$R_g = \frac{\int_0^{D_{max}} r^2 p(r) dr}{2 \int_0^{D_{max}} p(r) dr} . \quad (1.5)$$

The forward scattering, $I(0)$ can also be calculated from $p(r)$ as

$$I(0) = 4\pi \int_0^{D_{max}} p(r) dr , \quad (1.6)$$

which can then be used to obtain molecular weights. The main advantage to calculating R_g and $I(0)$ from the $p(r)$ function is that it uses all of the measured $I(q)$ data, rather than just the lowest q values. For systems that are not perfectly homogeneous or monodisperse or otherwise have bad low q data, use of Equations (1.5) and (1.6) can yield more reliable results. Finally, the shape of the $p(r)$ function can be used to gain significant insight into the shape of the particle. We will not go into detail here, but Figure 5 from (Svergun & Koch, 2003) illustrates this well.

Next, one usually attempts to model the three dimensional shape of the macromolecule. Because SAXS data is one dimensional, the resolution and reliability of reconstructions is limited. The

typical approach, implemented in the DAMMIN and DAMMIF programs (Svergun, 1999; Franke & Svergun, 2009; Petoukhov et al., 2012), starts with a volume of ‘beads’ (dummy atoms of some size), which are randomly assigned to one of two phases, either solvent or macromolecule. An energy for the system is defined from the chi-squared fit of the calculated scattering to the measured scattering and additional imposed constraints such as connectivity of the model. A simulated annealing approach is used, where moves are accepted if they lower the energy and accepted with Boltzmann probability if they increase the energy. The temperature is periodically lowered until moves stop improving the energy. For these reconstructions, a move is changing the phase of a random bead. Because this approach does not produce a unique solution, typically 10 or more structures are produced, and then a most likely structure is calculated based on the overlap (Volkov & Svergun, 2003).

If high resolution structural information (typically from crystallography or NMR) is available, one final data processing step is taken: the SAXS data is compared to the high resolution data. This is typically done in two ways. First, the shape reconstruction described in the previous paragraph can be overlaid on the high resolution structural model, allowing direct comparison of differences in shape. As the SAXS reconstruction is not high resolution, nor completely reliable, it is also usual to compare the predicted scattering from the high resolution structure to the measured $I(q)$. There are many programs that calculate predicted scattering, but the most common is CRY SOL (Svergun et al., 1995; Petoukhov et al., 2012). This comparison allows validation of a crystallographic structure with the solution structure, and allows further interpretation of the data. It is worth noting that it has been found that the effects of the hydration layer scattering must be included in the calculated scattering from the high resolution model to

obtain good agreement with the data, and that there are many methods for calculating the hydration layer scattering (Schneidman-Duhovny et al., 2012).

1.3 Summary of initial cryoSAXS work

The work in this dissertation builds upon the initial proof of principle cryoSAXS experiments from (Meisburger et al., 2013). Here I give a brief overview of the work and summarize the principle points. The interested reader is encouraged to read the entire paper for more details. I must also mention that while I helped with that initial proof of principle, it was really pushed forward by Steve Meisburger and Matt Warkentin.

The setup for this experiment started with a typical room temperature SAXS setup. We then removed the typical sample holder, and placed a 100 K nitrogen cold gas stream at the sample position. Samples were loaded in small, open ended, plastic capillaries, or nylon or polyimide loops, and cryocooled by placing them in the cold gas stream. Data was collected through the open ends of the capillaries and the centers of the loops. This is shown in Figure 1.2 (Figure 2a in the paper). In the following sections I highlight the main results of the paper, obtained using this setup.

1.3.1 Achieving reproducible cryocooling for SAXS

As discussed in Section 1.2, measuring a macromolecular scattering profile with SAXS involves two measurements: the macromolecular solution and the buffer solution without the macromolecule. The buffer scattering profile is then subtracted from the macromolecular solution profile, leaving just the scattering from the macromolecule (and hydration shell). The accuracy of the subtraction depends on measuring reproducible scattering profiles for both

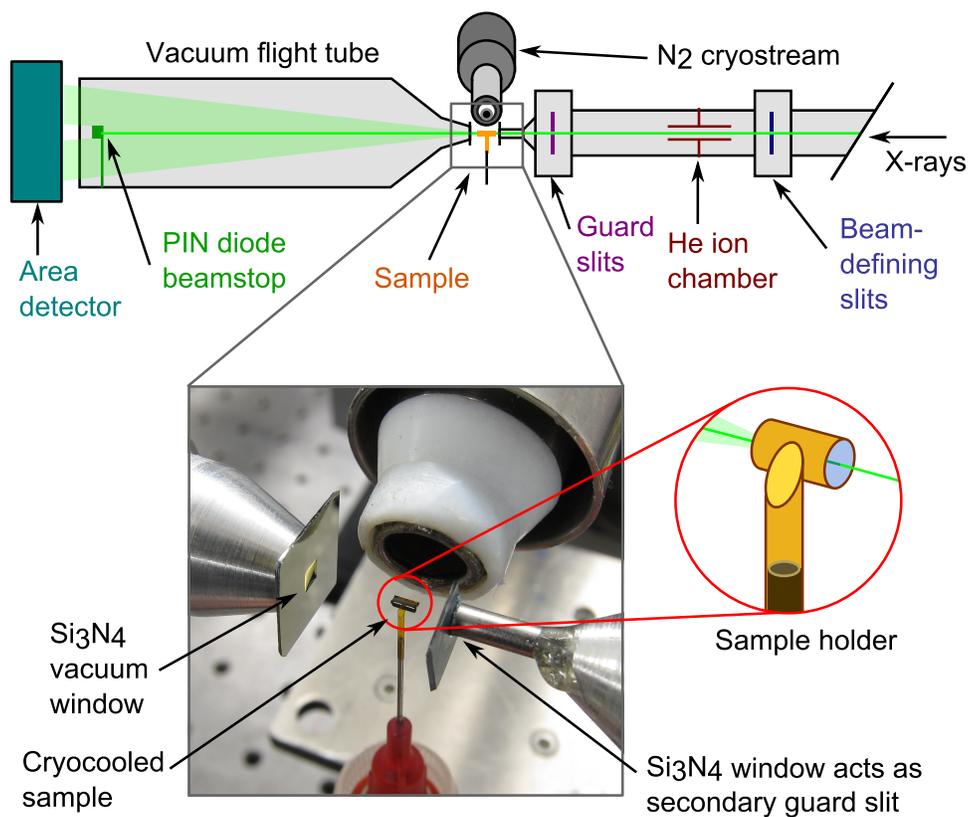


Figure 1.2 – The top is a schematic of the experimental layout. The inset of the sample position shows an image of a windowless sample, constructed for a capillary, in the nitrogen cryostream. The inset from that shows a cartoon of the sample holder and the direction of the X-ray beam. Figure 2 from (Meisburger et al., 2013).

sample and buffer. In order to get reliable SAXS measurements at 100 K we first needed to eliminate any variability in scattering from nominally identical samples that were cryocooled to 100 K. We determined that, likely due to the stochastic nature of homogeneous ice nucleation (both in time and space), ice formation in the sample gave irreproducible signals upon cooling. Additionally, the magnitude of the scattering was many times what would be expected from protein. In order to eliminate this variable we vitrified the samples, preventing ice formation through a combination of fast cooling and the addition of chemical cryoprotectants to the buffer solutions. We tested a number of cryoprotectants and determined that Polyethylene Glycol 200 (PEG 200) was the best, based on biocompatibility, effectiveness by weight at preventing ice formation, and contrast (the electron density difference between protein and buffer, which affects the total scattered intensity). It was determined that 45% w/w PEG 200 was necessary to completely suppress ice formation, or alternatively completely vitrify the sample, and achieve reproducible scattering profiles at 100 K. This is shown in Figure 1 of (Meisburger et al., 2013). It is noteworthy that the concentration of cryoprotectant to eliminate ice scattering in SAXS was ~5% higher than the concentration that produced a visually clear sample.

The use of windowless sample holders (open ended capillaries or loops) meant that the X-ray illuminated sample path length varied from cooling to cooling depending on the precise volume loaded, sample position, and shape of the meniscus. This variation was corrected for by measuring the transmissions of the sample and buffer, and the background scattering of the beamline without any sample or buffer present. These, along with the usual buffer and sample measurements, were combined into the three curve subtraction protocol described in the paper, compensating for the lack of fixed path length.

1.3.2 Results from macromolecules

After establishing the method for achieving reliable background subtraction at 100 K, we measured scattering from macromolecules, both at 100 K and room temperature. There were three major questions that had to be answered. First: how did the addition of 45% w/w PEG 200 affect the macromolecule? Second: did the cryocooling affect the macromolecule? Third: did cooling to 100 K significantly increase the radiation tolerance of the macromolecules? The answers we found will be described in turn.

Figure 1.3 shows scattering profiles for the protein glucose isomerase at 100 K, at room temperature in the 45% w/w PEG 200 buffer, and at room temperature in a standard buffer, Figure 4 from the paper. It also shows the predicted scattering profile from a high resolution structure of the protein. All of the profiles were in excellent agreement, indicating that neither the addition of cryoprotectant nor the cryocooling significantly altered the structure of the protein. Comparison of the R_g measured for the different profiles also showed good agreement, with the R_g at 100 K being ~3% larger than that at room temperature. This increase was attributed to changes in the hydration layer. Shape reconstructions fit the measured crystal structure well (Figure 3c in the paper). All of this showed that cryoSAXS can provide information from proteins at 100 K that is not significantly different from that provided by room temperature SAXS.

We also tested the radiation tolerance of the samples. Radiation damage in SAXS is typically quantified by looking at the change in the scattering profiles or R_g as a function of dose, where dose is energy absorbed per unit mass and has units of Gray (1 Gy = 1 J/kg). Glucose isomerase at 100 K was exposed for a total of 3.7 MGy of dose with no radiation damage (no observed

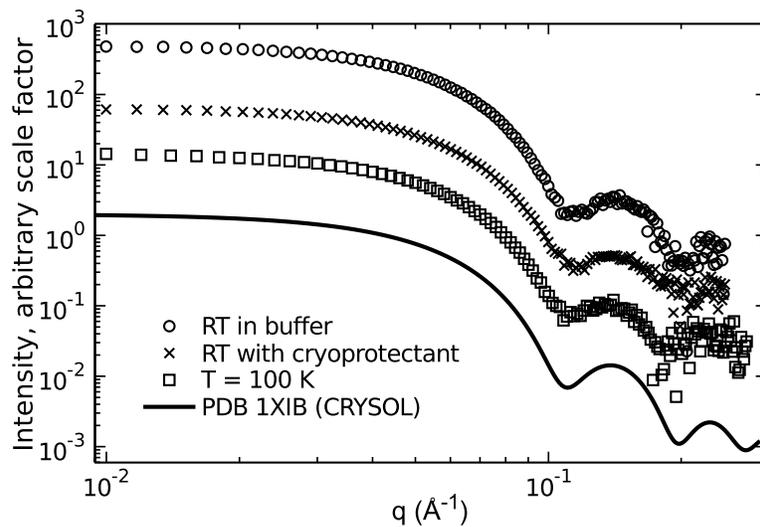


Figure 1.3 – Scattering profiles for glucose isomerase. From top to bottom: at room temperature in standard buffer, at room temperature in 45% w/w PEG 200 buffer, at 100 K, and predicted from the crystallographic structure. As can be seen, all of the profiles are very similar, and show the same characteristic bounces at high q . This is Figure 4 from (Meisburger et al., 2013).

change in the scattering profile or R_g), Figure 5 in the paper. At room temperature, radiation damage to glucose isomerase has been observed to start at 7 kGy of dose (Jeffries et al., 2015), so cryoSAXS increased the dose tolerance of the protein by at least a factor of 500.

This increased dose tolerance should enable corresponding reductions in minimum sample volumes. To test this, we collected data from 13-25 nL volume samples, 1000 times less than the 10-30 μ L typically used for room temperature SAXS. We collected scattering profiles for glucose isomerase, lysozyme, and 24 base pair duplex DNA. All of the profiles agreed with predicted scattering from high resolution structures, and the shape reconstructions also agreed well with the high resolution structures.

1.3.3 Summary and next steps

The experiments described above were sufficient to demonstrate the feasibility of cryoSAXS. Development of a robust system for cryoSAXS could have many exciting applications and open up new scientific avenues. In particular, the increased dose tolerance should enable measurements on systems previously unavailable to SAXS, increases in automation and high throughput, and perhaps enable time resolved SAXS with modest sample volumes. However, there remained several significant obstacles to routine cryoSAXS, and our attempts to resolve these comprise the majority of this dissertation. The first of these obstacles was the variability of the path length upon freezing. This required extra measurements, extra data processing, and engendered some skepticism as to the validity of our results. Accompanying that, though not mentioned in the paper, we struggled to achieve proper background subtraction even with the extra measurements and processing. The question remained how much of that difficulty was

fundamental to the technique, or at least our approach to it, and how much could be eliminated by refining our experimental procedure.

The second obstacle was determining exactly how much the dose tolerance of the sample increased when cooled to 100 K. This is required to assess the feasibility of different cryoSAXS experiments, and to better plan necessary sample volumes, exposure times, and more during experiments. Further, it should provide insight into the mechanism of radiation damage for samples at 100 K.

There was also a more general question: how widely applicable is cryoSAXS? Because of the required high concentration of cryoprotectants, and the mostly unknown effect of the cryocooling, cryoSAXS may only work for a small portion of macromolecules. This question is still very relevant, and we have only begun to explore it. One way to address this question is to work to minimize the effect of cryoprotectant on scattering and structure, by proper choice of cryoprotectant and increasing cooling rate to minimize necessary concentration. Further exploration of cryoprotectants, cooling methods, buffer conditions, and different proteins, in particular flexible proteins and large, weakly assembled complexes, is still ongoing and necessary for the widespread adoption of cryoSAXS.

1.4 Structure of this dissertation

This dissertation is constructed using the papers method, so each of Chapters 2-5 is a published paper (2 and 5) or paper draft (3 and 4). Chapters 2-4 represent work directly related to the cryoSAXS project described previously, while Chapter 5 is from a different project. In order to provide proper context and continuity, a preface has been written as the first section of each

Chapter, before the main body of the paper is presented. Brief summaries of these Chapters are given here.

One of the main problems with the proof of concept for cryoSAXS was the need for additional measurements and processing of the data. These additional requirements create a barrier to entry for standard SAXS users, and also raised some skepticism among experts as to the viability of the technique. In Chapter 2, the design, fabrication, and testing of a novel cryoSAXS sample holder that removes the need for these additional measurements and processing is described (Hopkins et al., 2015). This removes some of the difficulty of cryoSAXS, and provides validation of the initial results. As will be discussed, the effort was not completely successful, and efforts to design a further improved sample holder are ongoing.

The allure of cryoSAXS is the capacity to greatly increase radiation tolerance of biological samples. This increased radiation tolerance would allow the use of smaller sample volumes, which opens up a number of new scientific avenues (Meisburger et al., 2013). However, in order to quantify the increase, methods and baseline measurements were first needed for radiation tolerance at room temperature. Chapter 3 details work to accurately quantify radiation damage in room temperature SAXS, and discusses how it extends previous methods and results (Kuwamoto et al., 2004; Jeffries et al., 2015).

Once a proper baseline was established at room temperature, the next step was to quantify radiation damage at low temperatures. Chapter 4 reports the results of cryoSAXS experiments that reveal the rate of radiation damage at 100 K. Though the rate of damage has some dependence on buffer conditions and the particular protein, an increase in the radiation tolerance of three orders of magnitude is observed, relative to room temperature. This is compatible with

the initial estimates for cryoSAXS (Meisburger et al., 2013), and also provides insight into ways to further increase the radiation tolerance at both 100 K and room temperature.

In order to develop a successful method for cryoSAXS, an understanding of the ice formation behavior in aqueous solutions was required. This has been an ongoing research focus in the Thorne group, as it has practical applications to cryopreservation and protein crystallography (Kriminski et al., 2003; Berejnov et al., 2006; Warkentin et al., 2006, 2008; Warkentin & Thorne, 2009) and provides insight into fundamental physics (Warkentin et al., 2013b). Chapter 5 presents research on the ice formation behavior of aqueous solutions upon warming (Hopkins et al., 2012). The work is directly relevant to the cryopreservation of biological samples, as ice formation on cooling or warming can destroy the samples. The work reports measurements of the warming rates necessary to prevent ice formation in aqueous solutions. These are then compared to the cooling rates needed to prevent ice formation in the same solutions, and the differences are understood in terms of the fundamental physics of ice formation and growth as a function of temperature in these solutions.

The development and improvement of cryoSAXS is an area that is ongoing and will continue after my graduation. Chapter 6 describes the current efforts to develop a method for time resolved cryoSAXS, the motivation for this technique, and current challenges facing successful implementation. A brief overview is also given of what I see as the other outstanding questions or challenges for cryoSAXS along with a set of possible experiments that could address some of these questions and challenges.

Over the course of the work composing this dissertation, a detailed fabrication procedure was developed for manufacturing the sample holders described in Chapter 2. This is laid out in detail in Appendix A. Appendix B gives a brief overview of key parts and features of the custom code

base established for processing SAXS data, particular radiation damage data. This is done in the hopes that future students might be able to use this, instead of having to recreate the entire structure from scratch.

REFERENCES

- Acerbo, A., Cook, M., & Gillilan, R. (2015). *J. Synchrotron Radiat.* **22**, 1–7.
- Bai, X., McMullan, G., & Scheres, S. H. . (2015). *Trends Biochem. Sci.* **40**, 49–57.
- Berejnov, V., Hussein, N. S., Alsaied, O. A., & Thorne, R. E. (2006). *J. Appl. Crystallogr.* **39**, 244–251.
- Blanchet, C. E. & Svergun, D. I. (2013). *Annu. Rev. Phys. Chem.* **64**, 37–54.
- Blundell, T. L., Jhoti, H., & Abell, C. (2002). *Nat. Rev. Drug Discov.* **1**, 45–54.
- Bras, W., Koizumi, S., & Terrill, N. (2014). *IUCrJ.* **1**, 478–491.
- Campbell, I. D. (2002). *Nat. Rev. Mol. Cell Biol.* **3**, 377–381.
- Chaudhuri, B. N. (2015). *Protein Sci.* **24**, 267–276.
- Chen, P. & Hub, J. S. (2014). *Biophys. J.* **107**, 435–447.
- Chen, P. & Hub, J. S. (2015). *Biophys. J.* **108**, 2573–2584.
- Dubochet, J., Booy, F. P., Freeman, R., Jones, a V, & Walter, C. a (1981). *Annu. Rev. Biophys. Bioeng.* **10**, 133–149.
- Dyer, K., Hammel, M., & Rambo, R. (2014). *Struct. Genomics Gen. Appl.* **1091**,.
- Fischetti, R. F., Rodi, D. J., Mirza, A., Irving, T. C., Kondrashkina, E., & Makowski, L. (2003). *J. Synchrotron Radiat.* **10**, 398–404.
- Franke, D. & Svergun, D. I. (2009). *J. Appl. Crystallogr.* **42**, 342–346.
- Glatter, O., Holmes, K. C., Kirste, R., Kostorz, G., Laggner, P., Leopold, H., Muller, K., Oberthur, R. C., Pilz, I., Porod, G., et al. (1982). *Small Angle X-ray Scattering* New York: Academic Press.
- Grant, T. D., Luft, J. R., Wolfley, J. R., Tsuruta, H., Martel, A., Montelione, G. T., & Snell, E. H. (2011). *Biopolymers.* **95**, 517–530.
- Holton, J. M. (2009). *J. Synchrotron Radiat.* **16**, 133–142.
- Hong, X. & Hao, Q. (2009). *Rev. Sci. Instrum.* **80**, 014303.
- Hope, H. (1988). *Acta Crystallogr. Sect. B Struct. Sci.* **44**, 22–26.

- Hopkins, J. B., Badeau, R., Warkentin, M., & Thorne, R. E. (2012). *Cryobiology*. **65**, 169–178.
- Hopkins, J. B., Katz, A. M., Meisburger, S. P., Warkentin, M. A., Thorne, R. E., & Pollack, L. (2015). *J. Appl. Crystallogr.* **48**, 227–237.
- 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.
- Hynson, R. M. G., Duff, A. P., Kirby, N., Mudie, S., & Lee, L. K. (2015). *J. Appl. Crystallogr.* **48**, 1–7.
- Jacques, D. & TrewHELLa, J. (2010). *Protein Sci.* 642–657.
- Jeffries, C. M., Graewert, M. A., Svergun, D. I., & Blanchet, C. E. (2015). *J. Synchrotron Radiat.* **22**, 1–7.
- Kirby, N. M. & Cowieson, N. P. (2014). *Curr. Opin. Struct. Biol.* **28**, 41–46.
- Kriminski, S., Kazmierczak, M., & Thorne, R. E. (2003). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **59**, 697–708.
- Kuwamoto, S., Akiyama, S., & Fujisawa, T. (2004). *J. Synchrotron Radiat.* **11**, 462–468.
- Marion, D. (2013). *Mol. Cell. Proteomics*. **12**, 3006–3025.
- Meisburger, S. P., Warkentin, M., Chen, H., Hopkins, J. B., Gillilan, R. E., Pollack, L., & Thorne, R. E. (2013). *Biophys. J.* **104**, 227–236.
- Nature.com. Nature subject area definition: structural biology. <http://www.nature.com/subjects/structural-biology>, accessed 6/15/15.
- Petoukhov, M. V & Svergun, D. I. (2013). *Int. J. Biochem. Cell Biol.* **45**, 429–437.
- 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.
- Putnam, C. D., Hammel, M., Hura, G. L., & Tainer, J. a (2007). *Q. Rev. Biophys.* **40**, 191–285.
- Rambo, R. P. & Tainer, J. a (2011). *Biopolymers*. **95**, 559–571.
- Rcsb.org. PDB Current Holdings Breakdown. <http://www.rcsb.org/pdb/statistics/holdings.do>, accessed 6/13/15.
- Sbkb.org. PSI TargetTrack. <http://sbkb.org/tt/>, accessed 6/13/15.

- Schneidman-Duhovny, D., Kim, S. J., & Sali, A. (2012). *BMC Struct. Biol.* **12**,
- Skou, S., Gillilan, R. E., & Ando, N. (2014). *Nat. Protoc.* **9**, 1727–1739.
- Svergun, D., Barberato, C., & Koch, M. H. J. (1995). *J. Appl. Crystallogr.* **28**, 768–773.
- Svergun, D. I. (1992). *J. Appl. Crystallogr.* **25**, 495–503.
- Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.
- Svergun, D. I. & Koch, M. H. J. (2003). *Reports Prog. Phys.* **66**, 1735–1782.
- Svergun, D. I., Koch, M. H. J., Timmins, P. A., & May, R. P. (2013). *Small Angle X-ray and Neutron Scattering from Solutions of Biological Macromolecules* New York: Oxford University Press.
- Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S., & Zaccai, G. (1998). *Proc. Natl. Acad. Sci. U. S. A.* **95**, 2267–2272.
- Vestergaard, B. & Sayers, Z. (2014). *IUCrJ.* **1**, 523–529.
- Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Crystallogr.* **36**, 860–864.
- Warkentin, M., Berejnov, V., Husseini, N. S., & Thorne, R. E. (2006). *J. Appl. Crystallogr.* **39**, 805–811.
- Warkentin, M., Hopkins, J. B., Badeau, R., Mulichak, A. M., Keefe, L. J., & Thorne, R. E. (2013a). *J. Synchrotron Radiat.* **20**, 7–13.
- Warkentin, M., Sethna, J. P., & Thorne, R. E. (2013b). *Phys. Rev. Lett.* **110**, 015703.
- Warkentin, M., Stanislavskaja, V., Hammes, K., & Thorne, R. E. (2008). *J. Appl. Crystallogr.* **41**, 791–797.
- Warkentin, M. & Thorne, R. E. (2009). *J. Appl. Crystallogr.* **42**, 944–952.

CHAPTER 2

A MICROFABRICATED FIXED PATH LENGTH SILICON SAMPLE HOLDER IMPROVES BACKGROUND SUBTRACTION FOR CRYOSAXS

2.1 Preface

As mentioned in Chapter 1, a significant difficulty with the first cryoSAXS paper (Meisburger et al., 2013) was the additional measurements and processing of the data required to compensate for the variable path length of the sample. We found that we were not always able to obtain precise agreement when applying this correction. This led to improper background subtraction, which is the reason that the Bayesian indirect Fourier transform method was used to recover a constant offset from the scattering profiles for the small volume drops (Meisburger et al., 2013). We also thought there was a possibility that some of the irreproducibility was coming from variability in the shape of the sample-air interface. These were all valid reasons to question the reliability of the cryoSAXS method.

In order to overcome these difficulties, we decided to make a cryoSAXS sample holder that would give fixed path length and consistent sample-interface shape. The rest of this chapter is the paper published detailing the design, fabrication, and testing of such a sample holder (Hopkins et al., 2015). Note that in order to adhere to the dissertation formatting guidelines, all of the original supporting figures have been integrated directly into the body of the text.

2.2 Introduction

Small angle X-ray scattering (SAXS) provides low resolution structural information about biological macromolecules and complexes in solution. The technique's growing popularity arises

from its ability to provide structural information without crystallizing or labeling the molecules (Svergun & Koch, 2003; Mertens & Svergun, 2010). Currently, almost all SAXS on biological samples is performed at temperatures between 4° and 20° C, and the corresponding sample preparation and handling methods are well established. However, samples may be susceptible to damage from the X-ray exposure, and to degradation by other processes on typical experimental timescales.

The cryoSAXS method involves cooling samples rapidly to 100 K in the presence of chemical cryoprotectant (Meisburger et al., 2013). The fast cooling and cryoprotectant are both necessary to prevent the formation of crystalline ice, and to obtaining a vitrified sample with repeatable scattering at 100 K. Meisburger et al. (2013) reported that the radiation tolerance of biological samples in cryoSAXS is two to five orders of magnitude larger than in room temperature SAXS. CryoSAXS scattering profiles of the SAXS standard glucose isomerase agreed well with those obtained from room temperature samples and with theoretical scattering curves predicted from the crystal structure. Data were acquired with no radiation damage from illuminated sample volumes as small as 13 nL.

CryoSAXS has the potential to measure scattering profiles from molecules that are difficult or impossible to examine with room temperature SAXS. Because sample volumes can be orders of magnitude smaller for cryoSAXS than for room temperature data collection, measurements are enabled from macromolecules or complexes that are difficult or costly to express or otherwise obtain. The smaller volumes also increase the number of experiments that can be run with a fixed amount of sample, allowing, for example, greater screening of buffer conditions and ligand interactions. Cooling to cryogenic temperatures also prevents time dependent sample changes, facilitating measurement on unstable or aggregation-prone samples. Ultrafast cooling could be

used to trap time-dependent conformations of macromolecules with millisecond resolution. Finally, due to increased stability of cryocooled samples, cryoSAXS could enable routine mail-in and high-throughput data collection, similar to techniques now standard in cryo macromolecular crystallography (Blundell et al., 2002).

Although cryocooling has been used sporadically in the context of biological SAXS for decades (Kam et al., 1981; Iwamoto, 2009; Roessle & Svergun, 2011), reliable background subtraction was not reported until recently (Meisburger et al., 2013). The windowless sample holders used in that recent work led to variations in path length between sample and buffer. To obtain accurate subtraction between sample and buffer SAXS profiles, additional measurements of the scattering profile of the instrument background and of the transmission factors of sample and buffer were required to correct for these path length variations (Meisburger et al., 2013).

The use of a fixed path length sample holder for cryoSAXS experiments could eliminate the need for these additional measurements, the added uncertainties they introduce, and the extra data collection time they require. With fixed path length sample holders, the protocol for acquiring a cryoSAXS macromolecular scattering profile would be the same as at room temperature: normalize buffer and sample scattering profiles by the transmitted intensity and subtract (Skou et al., 2014).

This work describes fabrication of fixed path length silicon sample holders for cryoSAXS and characterization of these holders at a synchrotron beamline. The sample holders allow for small sample volumes, repeatable cryocooling, and more routine data collection and analysis. They have low background scattering (comparable to room temperature window materials used in this study), and low X-ray absorption at moderate X-ray energies, ~ 10 keV. Scattering profiles from samples in these holders are reproducible down to a scattering vector q ($q = 4\pi \sin\theta / \lambda$, where

2θ is the scattering angle and λ is the incident X-ray wavelength) of $\sim 0.02 \text{ \AA}^{-1}$ and agree well with those acquired in the windowless holders reported by (Meisburger et al., 2013). Irreproducible and often anisotropic scattering below $q \sim 0.02 \text{ \AA}^{-1}$ seems to arise from fractures that occur upon cooling to 100 K. Selective masking of anisotropic scattering features in SAXS images can limit the effect of fractures on the scattering profile. But even with the constraint on the minimum q vector, a wide range of biological molecules and complexes with maximum dimension up to $\sim 150 \text{ \AA}$ can be studied. The fixed path length, low background scattering, and low absorption of these sample holders represent a step towards routine biological SAXS data collection at 100 K.

2.3 Sample holder design, fabrication, and assessment

2.3.1 Design

The optimization of sample holders for cryoSAXS is subject to several constraints. The holders should have low X-ray absorption and low background scattering. The X-ray path length should be long enough to generate adequate SAXS signal. The X-ray windows should be stiff and rigidly supported so that the path length between them remains fixed even if the sample volume changes during cooling. The unobstructed window area should be sufficiently large so that the X-ray beam is not clipped and so does not generate grazing incidence scatter. The sample volume and holder thermal mass should be small to allow rapid cooling, and thus the use of minimal cryoprotectant concentrations to prevent ice formation. Finally, the holders should be easy to load and clean.

A simple holder that meets these criteria is an appropriately sized open-topped box with parallel vertical X-ray windows. Silicon is an ideal material for such a holder due to its relatively low X-

ray absorption, rigidity, relatively low cost, and to the availability of numerous high aspect ratio fabrication techniques (Wu et al., 2010).

Anisotropic wet etching yields smooth, high aspect ratio features in silicon (Bean, 1978; Heuberger, 1990; Hölke & Henderson, 1999; Dwivedi et al., 2000) and is cost effective. Potassium hydroxide (KOH) preferentially etches the (110) over the (111) planes of silicon, with etch selectivities reported from 200:1 (Hölke & Henderson, 1999) to 600:1 (Bean, 1978; Kendall, 1979). This feature allows fabrication of thin, high aspect ratio (up to 600:1 (Wu et al., 2010)), vertical (111) features on (110) silicon wafers. The etched (111) surfaces can be extremely smooth, with RMS roughness of 1-2 nm (Ezoe et al., 2006). These high aspect ratio, smooth (111) surfaces can serve as high quality X-ray windows in a sample holder that meets all of the above criteria.

The sample holder geometry is determined by both the etch mask and by the anisotropic KOH etching. On a (110) silicon wafer there are two parallel sets of vertical $\{111\}$ planes that intersect at a 70.52° angle (Kaminsky, 1985). The etch mask, Figure 2.1a, defines an open-topped parallelepiped box formed by intersecting $\{111\}$ planes, with one set of planes to serve as X-ray windows (horizontal lines) and the other as rigid window supports (skew lines). Due to the anisotropic etch rates of the crystal planes, features appear beyond those defined in the mask, discussed in Section 2.2.3.

KOH rapidly etches exposed convex corners, requiring protection of any such features (Kim & Cho, 1998; Pal et al., 2011). To eliminate exposed corners, the etch mask shown in Figure 2.1a was tiled horizontally with no break in horizontal parallel lines. The center-to-center spacing between each sample holder was typically 5 mm. A single 3" wafer was patterned with 13 such rows, creating approximately 160 sample holders per wafer.

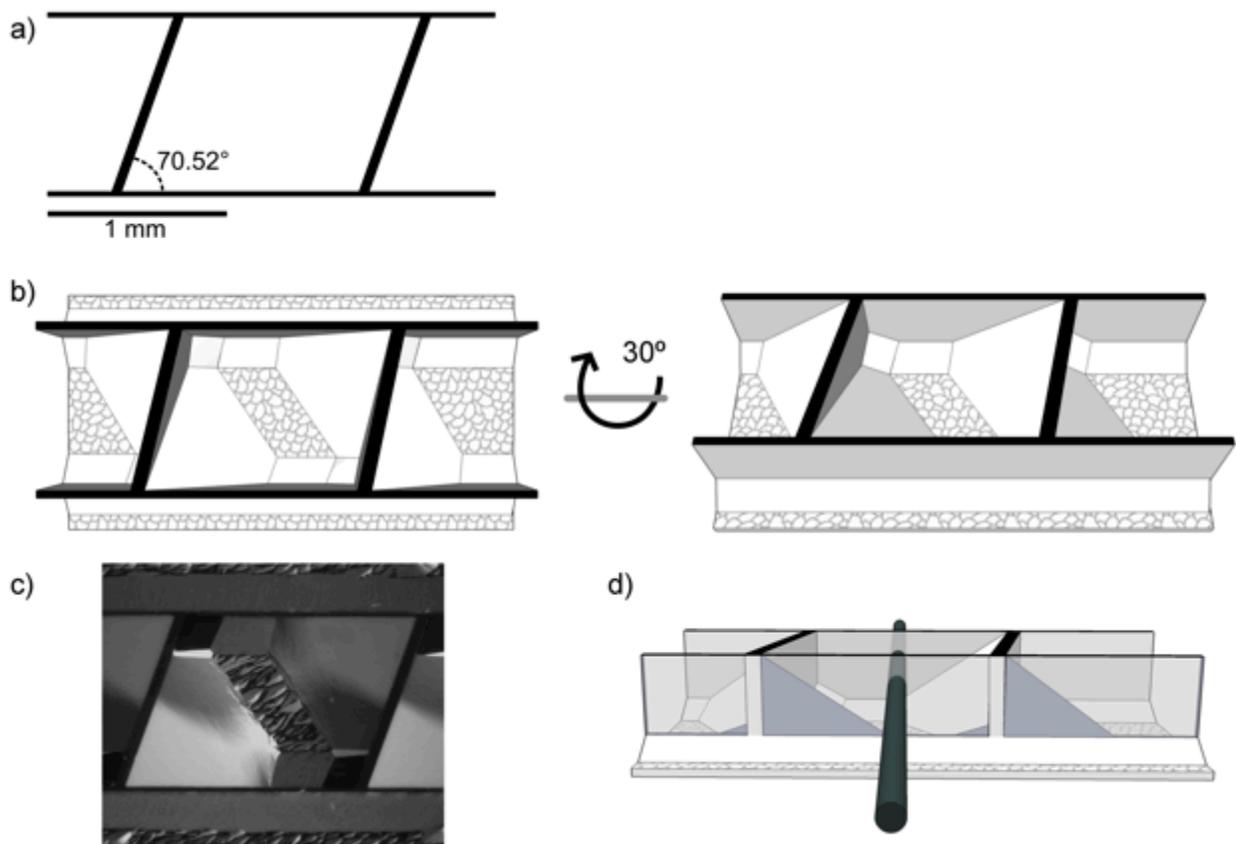


Figure 2.1 – a) Sample holder mask schematic. The horizontal lines are $30\ \mu\text{m}$ thick and the slanted crossbars are $56\ \mu\text{m}$ thick (measured perpendicular to the edge). b) Cartoon of a sample holder in two different orientations. The first orientation is a top down view. The second has been rotated 30° . The X-ray path length is 1.0 mm, the X-ray window thickness is $\sim 27\ \mu\text{m}$, and the X-ray windows are $\sim 620\ \mu\text{m}$ tall. c) Optical image of an empty sample holder, top down view. Composited from four images to achieve focus for the entire holder. The scale is the same as in a). d) Cartoon view of the holder rotated 80° from the top down view. The walls have been made semi-transparent so that the X-ray accessible volume is visible. The X-ray beam is shown in black. Parts a, b, and d of this figure were originally created by Andrea Katz.

The sample holder design has 30 μm thick X-ray windows separated by 1.0 mm, with side supports separated by 1.36 mm, and a target window height of 600 μm . These dimensions were chosen after experimentation with a variety of window thicknesses and sample volumes. The choice of window thickness is a compromise between rigidity (thick) and low X-ray absorption (thin), while the choice of sample volume is a compromise between fast cooling rates (small) and X-ray signal (large).

2.3.2 Wafer processing

To prepare silicon wafers for KOH etching, they were first cleaned and then patterned with a KOH-insoluble mask of low stress silicon nitride. Double-side polished (110) silicon wafers with a diameter of 76.2 mm and a thickness of $740 \pm 10 \mu\text{m}$ were purchased from Virginia Semiconductor (Fredericksburg, VA). Bare wafers were RCA cleaned (Kern, 1990) by immersion in a 6:1:1 deionized water: 29% by weight ammonium hydroxide: 30% by weight hydrogen peroxide solution at 70 °C for 10 minutes; rinsing in deionized water; immersion in a 6:1:1 deionized water: 37% by weight hydrochloric acid: 30% by weight hydrogen peroxide solution at 70 °C for 10 minutes; and then rinsing in deionized water. Approximately 100 nm of low stress silicon nitride was thermally grown on the cleaned wafers in an LPCVD furnace. A negative photoresist, AZ nLOF 2020 (AZ Electronic Materials, Branchburg, NJ), was spun and baked onto the cleaned wafer. The wafer's flats (which run along {111} planes) were visually aligned with a quartz photomask patterned with the design in Figure 2.1a (Heidelberg Mask Writer DWL2000) and the photoresist exposed in a contact aligner (SUSS MA6). After exposure and development the pattern was descummed using an oxygen plasma (YES CV200RFS Oxygen Plasma Asher). Unmasked nitride was removed using a tetrafluoromethane reactive ion etch

(Oxford PlasmaLab 80+ RIE System). The photoresist was then stripped, and the wafers were RCA cleaned as above.

With the silicon nitride etch mask complete, wafers were etched in 45% w/w KOH (Avantor, Center Valley, PA) solution containing 500 nL/L Triton X-100 ultra-grade surfactant (Fluka, Buchs, Switzerland) for 16 hours in a temperature regulated ultrasonic bath at 60° C. The etch vessel was covered to limit evaporation. The surfactant and the ultrasonic bath were used to minimize surface roughness of etched features. Wafers were removed from KOH and immediately placed in a beaker of 60 °C, 16.7 M Ω -cm deionized water filtered through a 0.2 μ m hollow fiber filter (Barnstead NanoPure II deionizing and filtration system). Every 2 minutes the wafers were transferred to a fresh deionized water rinse for a total of 10 rinses. The temperature of each successive rinse was decreased by 10 °C until room temperature was reached, and subsequent rinses were at room temperature. This rinse protocol was necessary to minimize the formation of KOH crystals on the wafers. After the rinse, wafers were RCA cleaned as above, then dipped in 0.8% hydrofluoric acid for 30 s and rinsed. The final fabricated wafer was scored with a diamond tipped scribe and cleaved to separate individual sample holders.

2.3.3 Evaluation of etched features

Figure 2.1b shows a sketch of the etched sample holder. The KOH etch naturally terminates on crystal planes such as the {111} and {311}, resulting in sloping walls not included in the etch mask (Hölke & Henderson, 1999). Figure 2.1c shows an optical image of a sample holder, and Figure 2.1d is a sketch showing the limited incursion of these sloping walls into the X-ray path. The nominal volume of a sample holder is 840 nL, of which the sloping walls occupy ~200 nL, giving a total sample volume of ~640 nL.

The parallel X-ray windows are a critical aspect of the sample holders. The windows were optically measured to be about 27 μm thick and 620 μm tall, with small variations across individual wafers and between wafers. The 1.5 μm {111} under-etch of the X-ray windows indicates an etch selectivity of 413:1, comparable to previously measured selectivities (Bean, 1978; Kendall, 1979; Hölke & Henderson, 1999). The window surface was examined in Zeiss Supra 55VP and Zeiss Ultra 55 SEMs. After optimization of etch parameters, these surfaces appeared featureless to the SEM at 1.5 kV. Windows were further examined with a Veeco Icon AFM by cleaving off individual windows and securing them to a substrate. The RMS roughness was measured to be 1.27 nm over 100 μm^2 . SEM and AFM images are shown in Figures 2.2 and 2.3.

2.4 Methods for SAXS

2.4.1 Sample preparation

Sample holder performance was evaluated by examining SAXS profiles of glucose isomerase from *Streptomyces rubiginosus* (Hampton Research, Aliso Viejo, CA, HR7-100). The protein was buffer exchanged in a spin column (Amicon 30 kDa MW cut-off, EMD Millipore, Billerica, MA) into 100mM Tris pH 8.0 and 1 mM magnesium chloride.

The choice of an appropriate cryoprotectant is based on several criteria. Most important, the cryoprotectant must have no adverse effects on macromolecule solubility or structure. Of those biocompatible cryoprotectants, those that have high ice prevention efficacy per unit mass and that give the best X-ray contrast (the difference in electron density between the macromolecule and the solvent (Svergun & Koch, 2003)) are preferred. The strength of the SAXS signal

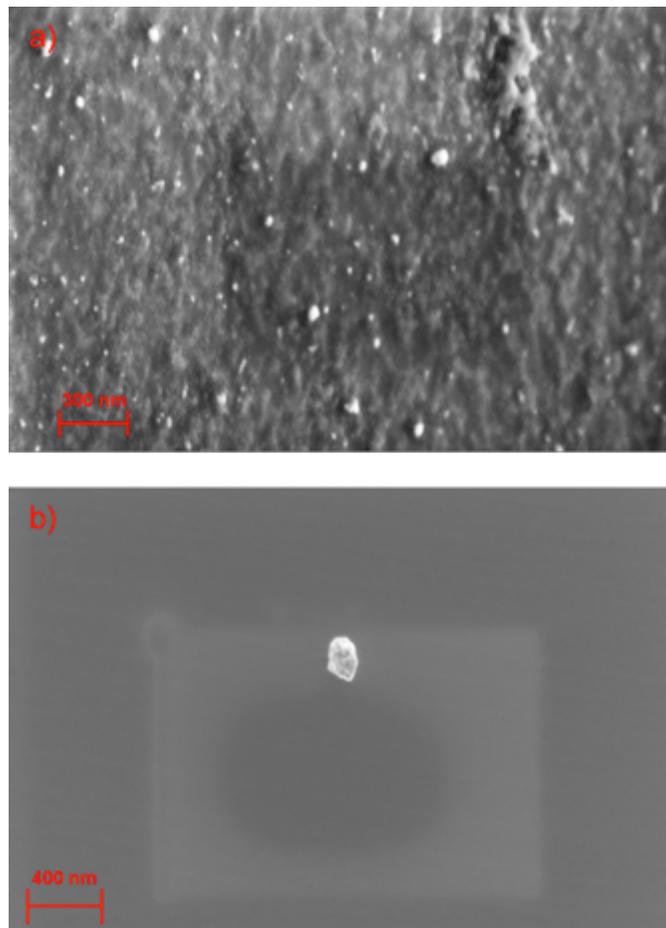


Figure 2.2 – a) SEM image (3 keV beam energy) of a rough silicon X-ray window, before etch parameters were optimized. b) SEM image (1.5 keV beam energy) of a smooth silicon X-ray window after optimization of etch parameters. A dust particle is included for focus because windows are featureless to the SEM.

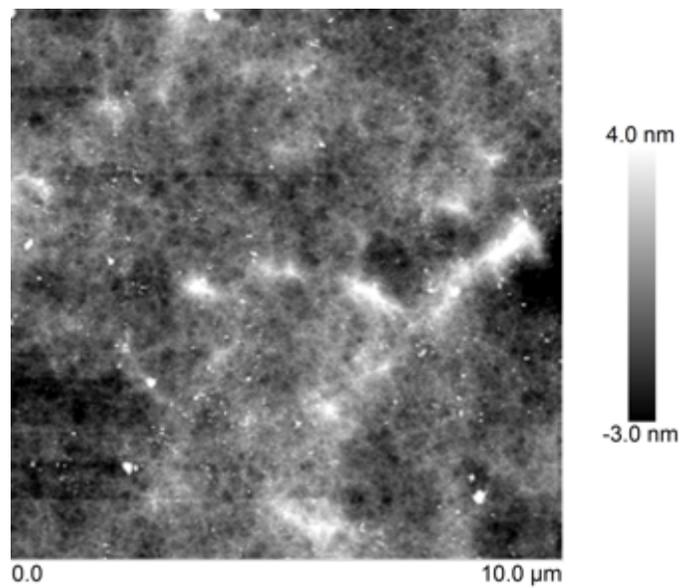


Figure 2.3 – AFM image of a silicon X-ray window with optimized etch parameters. The height variation is on the order of nanometers over $100 \mu\text{m}^2$. The window was cleaved from the sample holder and adhered horizontally to a substrate before AFM analysis.

depends in part upon the square of the X-ray contrast, so higher contrast yields better signal-to-noise.

In previous work, several cryoprotectants were screened (Meisburger et al., 2013). Polyethylene glycol (average molecular weight 200) (PEG 200) was found to be the best of those tested, and a concentration of 45% w/w PEG 200 was used to ensure solution vitrification upon cooling. Propylene glycol (PG) (Sigma, St. Louis, MO), untested by Meisburger et al. (2013), provides larger X-ray contrast and was found here to be superior for cryoSAXS experiments.

A 36% w/w PG solution was sufficient for consistent vitrification of the samples. At room temperature the electron density of this 36% w/w solution is $0.344 \text{ e}^-/\text{\AA}^3$ while the electron density of a 45% w/w PEG 200 solution is $0.355 \text{ e}^-/\text{\AA}^3$ (after accounting for the volume change upon mixing in both solutions (Muller & Rasmussen, 1991; Khattab et al., 2012)). As the electron density of proteins is $\sim 0.420 \text{ e}^-/\text{\AA}^3$ (Svergun & Koch, 2003), using PG instead of PEG 200 results in a $\sim 17\%$ increase in the contrast.

While scattering considerations are important, the effect of the cryoprotectant on the macromolecule is of primary concern. Glucose isomerase in the 36% w/w PG solution described below was tested by SAXS at room temperature to ensure that the PG had no adverse effects on the protein, Figure 2.4.

To prepare the glucose isomerase and matching buffer solutions, a solution was made with 730 mg/mL of PG and the same molality of other components as in the non PG-containing buffer. This PG solution was combined with non PG-containing protein and buffer solutions in a 1:1 volume ratio using positive displacement pipettes, creating samples with a 36% w/w PG concentration. Prior to mixing with the PG solution, protein concentration was measured using a NanoVue Plus Spectrophotometer (GE Healthcare Bio-Sciences, Pittsburgh, PA).

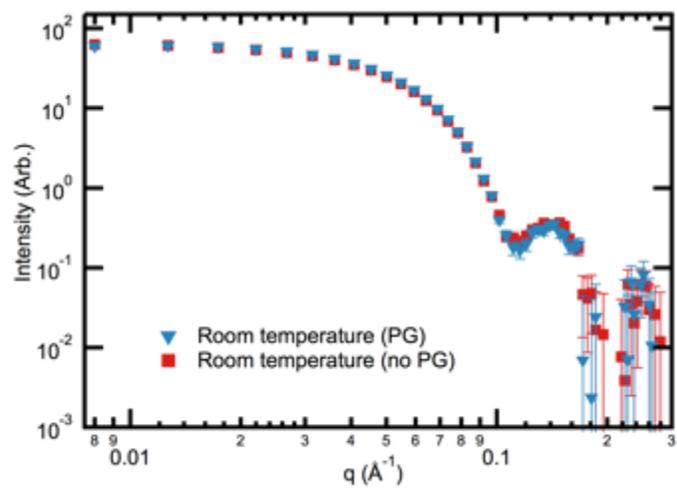


Figure 2.4 – Comparison of glucose isomerase (2.8 mg/mL) in PG at room temperature and in the same buffer without PG at room temperature.

The final concentration of glucose isomerase was 2.8 mg/mL.

2.4.2 Beamline setup

SAXS measurements were carried out at the G1 station at the Cornell High Energy Synchrotron Source (CHESS). A schematic of the beamline setup is shown in Figure 2.5. The upstream flight tube contained three sets of slits and a nitrogen ion chamber between the first two sets of slits for measuring incident beam intensity. A mica window (25 μm thick, scratch-free ruby mica discs, Attwater Group, Preston, Lancashire, England) was inserted downstream of the ion chamber, separating upstream nitrogen and downstream vacuum (6.2 mTorr) regions of the flight tube. One set of slits was used to define the beam, and the two most downstream sets served as guard slits. A 1'' air gap was introduced after the final guard slit to facilitate placement of the sample holder in the beam and cryocooling of the sample via a nitrogen cryostream (700 series, Oxford Cryosystems, Oxford, United Kingdom). The cryostream was equipped with an air-blade shutter built in-house. An aluminum flight tube downstream of the sample was held under vacuum (6.2 mTorr). Both the upstream and downstream flight tubes were capped at the air gap by mica windows, and the downstream end of the downstream flight tube was capped by a 50 μm thick Mylar window.

SAXS data were collected using a Pilatus 100 K detector (Dectris, Baden, Switzerland). To avoid potential nonlinearity associated with PIN diode beamstops, a 140 μm molybdenum foil was used as a semi-transparent beamstop, allowing normalization of the scattering profiles via direct measurement of transmitted intensity on the detector. For the foil thickness chosen, the

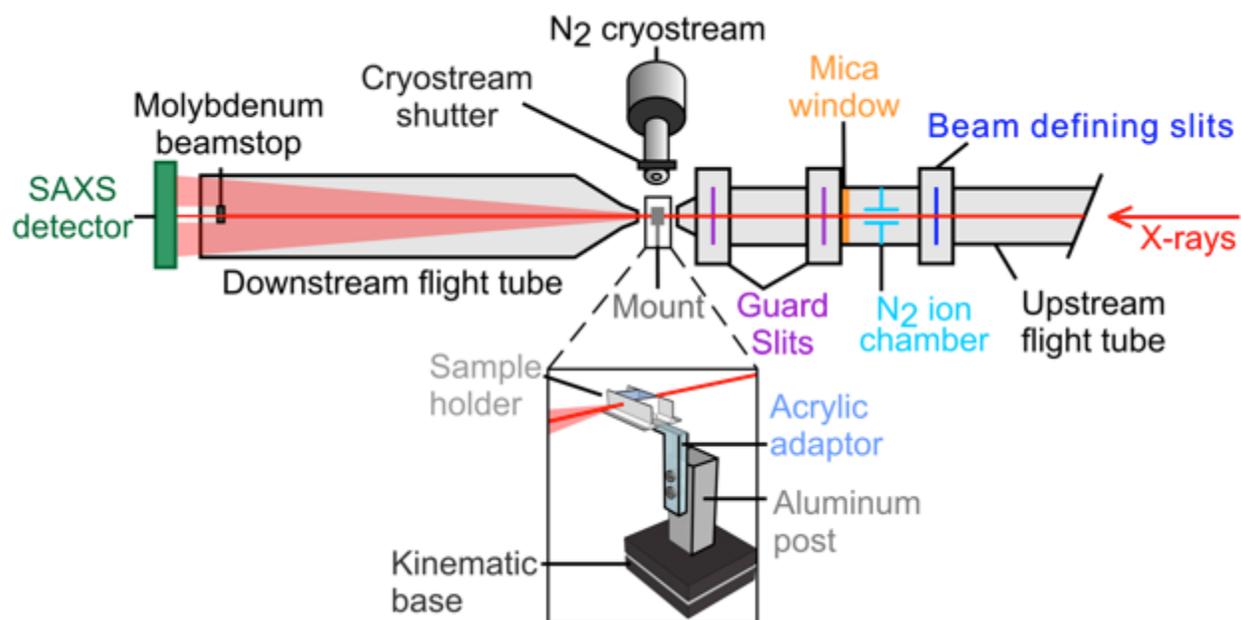


Figure 2.5 – Schematic (not to scale) of the SAXS setup with a detailed view of the sample holder and its mount (rotated and sample holder enlarged for clarity). This figure was originally created by Andrea Katz.

maximum counts in a single pixel during a ten second exposure were typically slightly less than 750,000. This number is well below the Pilatus pixel depth of 20 bits (1,048,576 counts) (and the corresponding count rate is well below the maximum count rate of 5,000,000 ph/s (Kraft et al., 2009)) but is sufficient for normalization with a precision of $1/\sqrt{750,000} = 0.1\%$. Extraneous diffraction rings from the molybdenum were absorbed by a horizontal overhanging ledge that was folded into the top of the foil.

For the SAXS measurements, the X-ray energy (wavelength) was 10.53 keV (1.18 Å), the flux after the beam defining slits measured by the nitrogen ion chamber was typically $\sim 6.5 \times 10^9$ ph/s, and the beam dimension was defined by the slits to be 50 $\mu\text{m} \times 50 \mu\text{m}$, giving a nominal illuminated sample volume of 2.5 nL. The sample-to-detector distance was calibrated using silver behenate ($d = 58.38 \text{ \AA}$ (Huang et al., 1993)) powder (The Gem Dugout, State College, PA) and was 1571.0 mm. The useful q-space range extended from 0.008 \AA^{-1} to 0.284 \AA^{-1} .

2.4.3 Data collection at 100 K

For data collection at 100 K, individual silicon sample holders were epoxied (Hardman Double Bubble epoxy, extra fast setting, Royal Adhesives and Sealants, Wilmington, CA) onto acrylic adaptors, shown in Figure 2.5. Adaptors were laser cut from a 1.6 mm thick acrylic sheet using a VersaLaser VLS3.50 (Universal Laser Systems, Scottsdale, AZ). These adaptors screwed onto a machined aluminum post attached to the top plate of a 1" kinematic base (Thorlabs, Newton, NJ). The bottom plate of the 1" kinematic base was attached to a mini rotation stage (Thorlabs), held on a sample positioning motor. The mount was placed in the air gap using the kinematic base, and the empty sample holder was positioned to center the beam in the holder.

After centering, the mount was removed and the sample holder loaded using a 1 μL 7000 series 25s gauge blunt tip modified microliter syringe (Hamilton Company, Reno, NV) controlled by a vertically mounted PHD 2000 syringe pump (Harvard Apparatus Holliston, MA) to ensure accurate and reproducible volumes. As the sample contracts upon cooling, 800 nL sample volumes were loaded to ensure complete filling of the sample holder at 100 K. To achieve consistent loading, the sample holder was placed on a second kinematic base and reproducibly positioned beneath the syringe using a 3-axis micrometer translation stage (Thorlabs). The cryostream was shuttered and the mount was replaced on the kinematic base in the beam. Finally, to cool the sample *in situ*, the cryostream was unshuttered. An average sample cooling rate of 25 K/s in the cryostream was measured with a 0.01" wire diameter bare wire E type thermocouple (Omega Engineering, Stamford, CT) inserted into the sample. Individual image exposure times were 10.0 s, and between 10 and 40 images were averaged together to obtain scattering profiles, depending upon the desired signal-to-noise.

After data collection, the mount was removed from the beam. The sample holder was cleaned by immersion and agitation in deionized filtered water for 30 s, and blown dry with compressed air. Buffer and sample were measured in the same holder. Care was taken to ensure that the X-ray beam illuminated the same position on the same sample holder for both buffer and sample.

2.4.4 Data collection at room temperature

Due to the small sample volume and the open top of the silicon sample holders, the sample evaporates too rapidly at room temperature to obtain reliable scattering profiles. Instead, acrylic sample holders similar to those described by (Hong & Hao, 2009) and with nominal sample volumes of ~ 3 μL were laser cut and windowed with 25 μm thick polystyrene (Goodfellow

Corporation, Coraopolis, PA). These sample holders screwed onto the same mounting apparatus used in the 100 K experiments. A sample holder was first centered in the X-ray beam and then removed. Then the sample holder was loaded using gel loading pipette tips and sealed using tape to prevent evaporation, and the sample holder was replaced in the X-ray beam. Room temperature SAXS data were acquired with the cryostream either shuttered or removed, using individual image exposure times of 1 s. The individual images were checked for radiation damage, and final scattering profiles were generated from 40 images averaged together. After data collection, the sample was removed, and the sample holder was filled and rinsed multiple times with deionized filtered water. As at 100 K, buffer and sample were measured in the same sample holder and with the X-ray beam passing through the same position on the sample holder.

2.4.5 Data processing

BioXTAS RAW version 0.99.9.14b was used at the beamline for data processing (Nielsen et al., 2009). SAXS curves were later reanalyzed using MATLAB (The MathWorks, Natick, MA) code written in-house. For room temperature data, processing was done according to standard protocols (Skou et al. 2014). A standard mask was applied, and the detector images were angularly integrated to generate the scattering profile. After integration, scattering profiles were normalized by the transmitted intensity to account for variation in the incident intensity or sample absorption. Buffer profiles were then subtracted from scattering profiles of macromolecule in solution to generate the scattering profile of the macromolecules alone.

For 100 K data, the same protocol was followed after integration. However, some images displayed anisotropic scattering from fractures that was masked and removed as discussed in Section 2.4.5. Additionally, to verify that the fractures were not affecting the scattering profile in

the q range of interest, profiles from separately loaded and cooled but otherwise identical instances of the same sample or buffer were taken and compared, also discussed in Section 2.4.5.

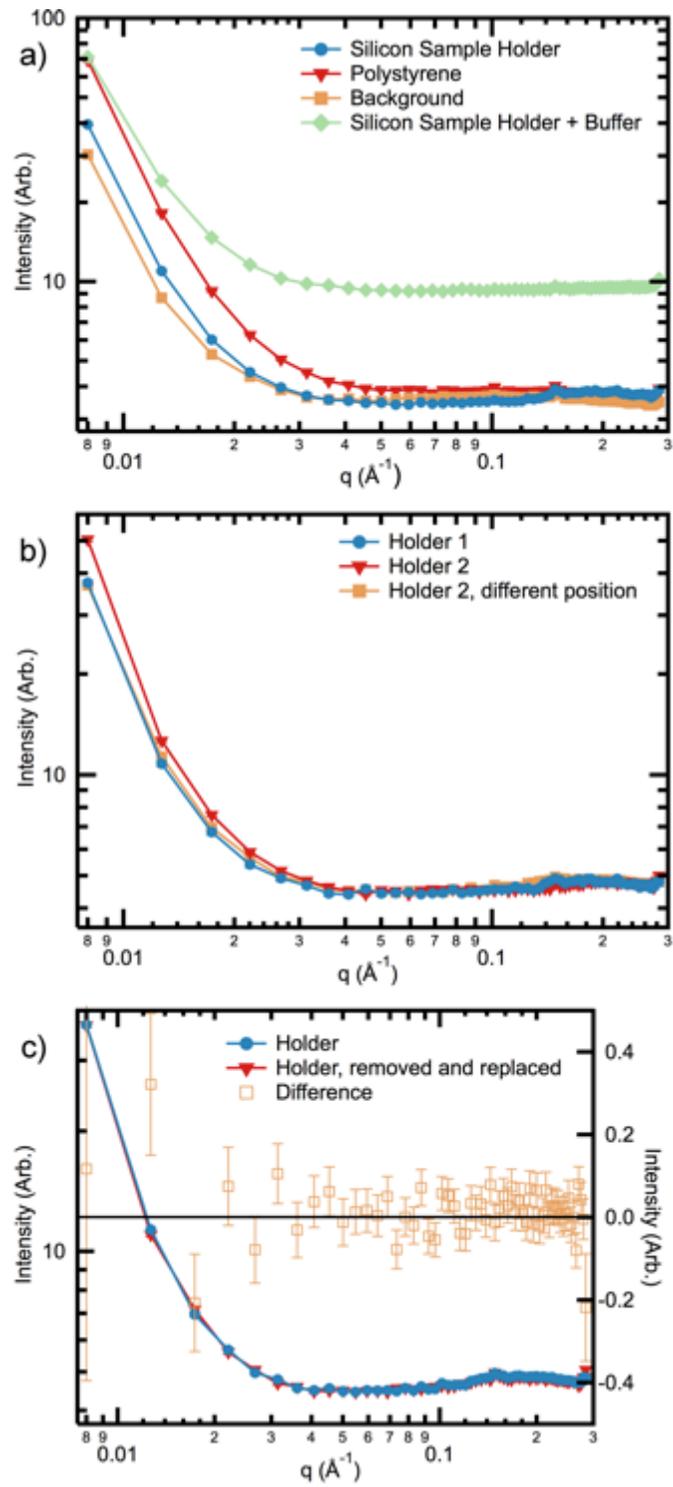
2.5 Results and analysis

2.5.1 Characterization of sample holders

The X-ray scattering properties of the sample holders were evaluated to determine their suitability for cryoSAXS experiments. Figure 2.6a compares the scattering profiles at 100 K from an empty silicon sample holder, an empty room temperature sample holder, and the 100 K instrument background. At mid and high q , the scatter of the silicon sample holders with 30 μm thick windows is comparable to that of the room temperature holders with 25 μm polystyrene windows; at low q , the silicon sample holder scatter is smaller, and comparable to that from the instrument background. The transmission of the silicon sample holders is 71%, similar to that of 10 μm wall thickness quartz capillaries commonly used as sample holders in room temperature SAXS (Nielsen et al., 2012).

Figure 2.6b compares the scattered intensity from two different sample holders from the same wafer, and from two different positions on the same sample holder. None of these profiles are in perfect agreement, notably at low q . Scattering profile differences could result from variable surface topography of the etched X-ray windows. Additionally, the sample holder absorbs upstream parasitic scatter -- scatter from any other non-sample sources including slits, other beamline components and air -- so the measured scattering profile also depends upon the holder's position relative to the beam. Consequently, any change in the holder's position relative to the beam due to holder translation or irreproducibility in holder mounting and centering will lead to variation in the background scatter. The ability to accurately reposition the sample holder

Figure 2.6 – a) Comparison of scatter from a silicon sample holder, a room temperature sample holder with two 25 μm polystyrene windows, and the instrument background, all at 100 K. For comparison the scattering profile of a 100 K buffer in a silicon holder is included. Lines added to guide eye. b) Comparison of scatter between two different silicon sample holders of the same design from the same wafer mounted nominally identically, and between positions 100 μm apart on sample holder 2. Lines added to guide eye. c) Scattering profiles from the same silicon sample holder, before and after being removed and replaced in the X-ray beam, plotted on the left axis. The difference curve is the difference between these profiles, plotted on the right axis, and demonstrates the reproducibility of our sample holder mounting system. Lines added to guide eye.



is thus critical for background subtraction. The rigid mounting scheme shown in Figure 2.5 ensures consistent holder position and provides identical SAXS profiles when a given holder is removed and replaced, as shown in Figure 2.6c. The sample holders and mounting apparatus were robust when subject to numerous freeze-thaw and cleaning cycles. These features are essential to obtaining accurate and reliable buffer subtraction.

2.5.2 CryoSAXS of glucose isomerase

Figure 2.7a shows buffer subtracted scattering profiles of glucose isomerase measured at 100 K and at room temperature. A cryoSAXS profile of glucose isomerase in a 45% w/w PEG 200 buffer collected in a windowless sample holder, reported in (Meisburger et al., 2013), is also shown. All three traces agree well down to $q = 0.013 \text{ \AA}^{-1}$. Figure 2.8 shows the residual between the present room temperature and 100 K data. Non-zero residuals are expected to result from changes in solvent structure and density (discussed below) on cooling to 100 K even if the shape of the protein is identical; they may also occur due to changes in molecular structure or to experimental variability. While direct comparison of scattering profiles provides the most straightforward evaluation of data quality, comparison of calculated quantities, such as the radius of gyration (R_g), allows for quantitative comparisons with literature results. Pair-distance distribution functions ($P(r)$) (Figures 2.9 and 2.10) were calculated from SAXS profiles using GNOM (Svergun, 1992) for the entire available q range. The GNOM R_g values are $33.34 \pm 0.08 \text{ \AA}$ and $32.70 \pm 0.04 \text{ \AA}$ at 100 K and room temperature, respectively. These compare well with values obtained for glucose isomerase in PEG 200 buffer at 100 K of $33.4 \pm 0.1 \text{ \AA}$ (Meisburger et al., 2013) and in cryoprotectant-free buffer at room temperature of $32.7 \pm 0.2 \text{ \AA}$

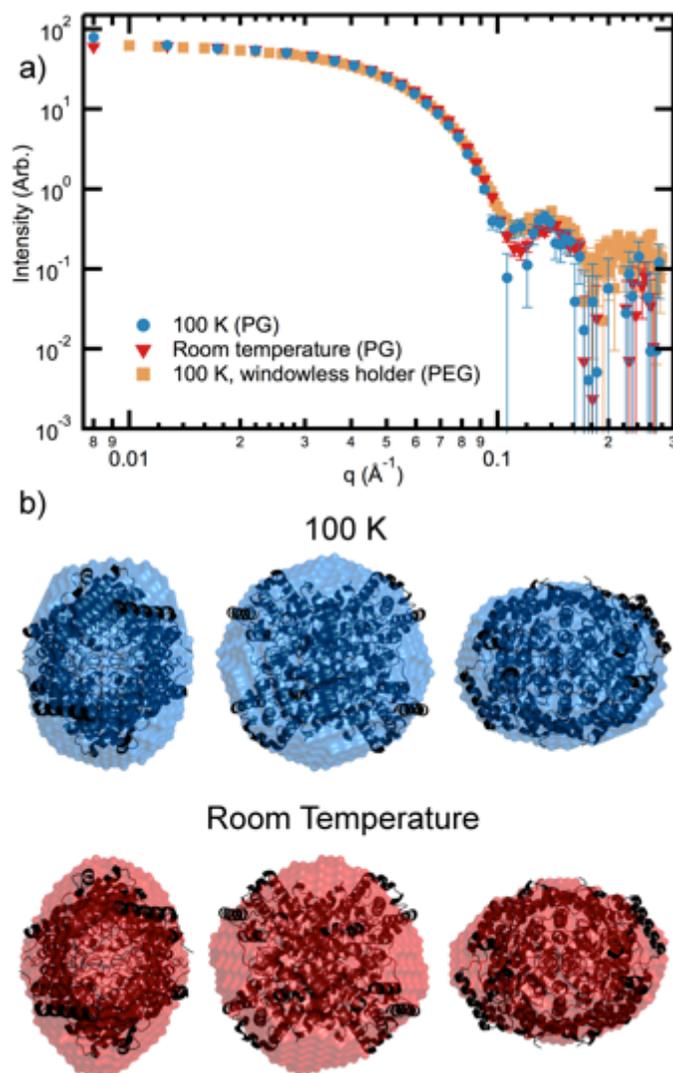


Figure 2.7 – a) Buffer subtracted glucose isomerase at 100 K and at room temperature. Data from glucose isomerase in PEG 200 in a windowless holder (Meisburger et al. 2013) at 100 K is included for comparison. Arbitrarily scaled to match at mid q . b) Envelope reconstructions (DAMFILT envelopes) of both the 100 K and room temperature data sets aligned with the crystal structure (black). Three orthogonal views are presented.

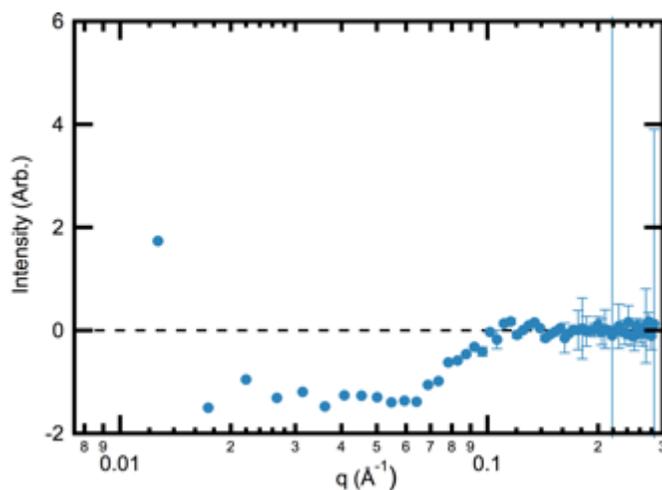


Figure 2.8 – Residual obtained by subtracting the scattering intensity at room temperature from the scattering intensity at 100 K for glucose isomerase. The original intensity curves are shown in Figure 2.7. Note that due to differences in solvent structure and density between 100 K and room temperature the curves are not expected to be identical, even if the protein has identical atomic structure, so the residual should not be zero. The lowest q point has an intensity value of ~ 18 , and is not included on the plot so residuals at high q are more visible. The q and intensity axes have the same range as Figure 2.12. The occasionally large error bars come from propagation of fractional errors for points near zero.

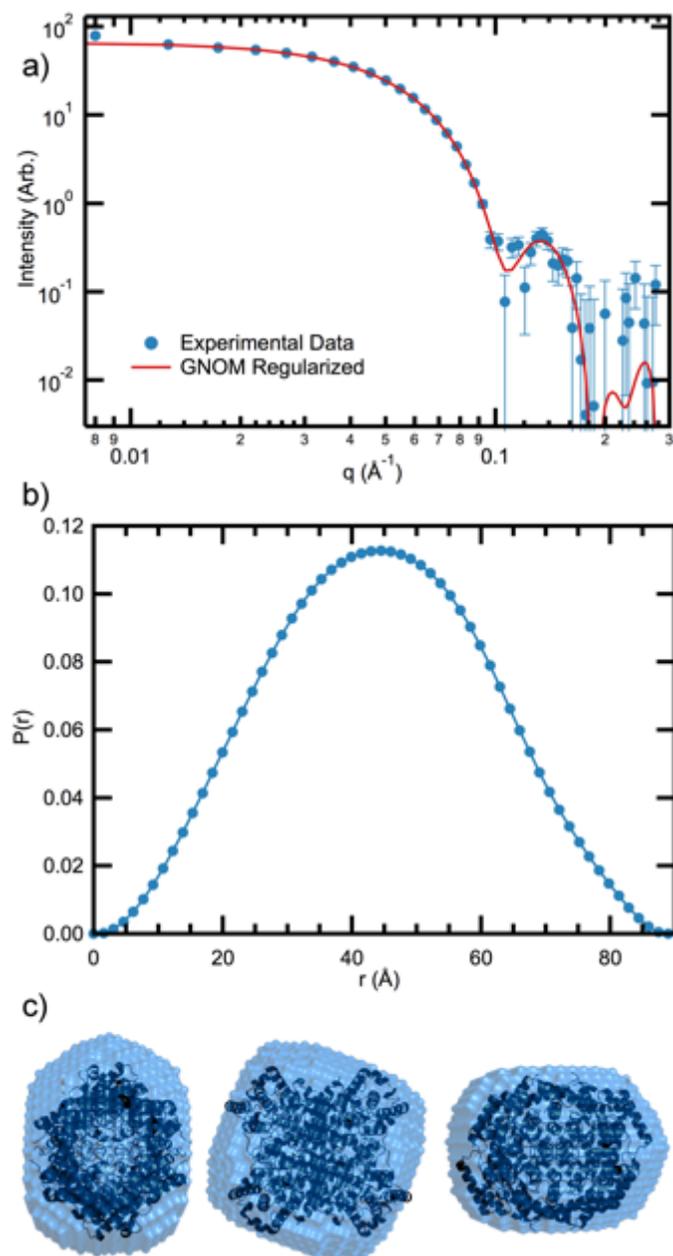


Figure 2.9 – All data in this figure are for glucose isomerase (2.8 mg/mL) at 100 K in PG. a) Experimental data, and regularized $I(q)$ from GNOM. $\chi^2 = 1.071$, total estimate = 0.934. b) $P(r)$ computed by GNOM with $D_{\max} = 89$ \AA . c) Three orthogonal views of the DAMAVER envelope aligned with the glucose isomerase crystal structure.

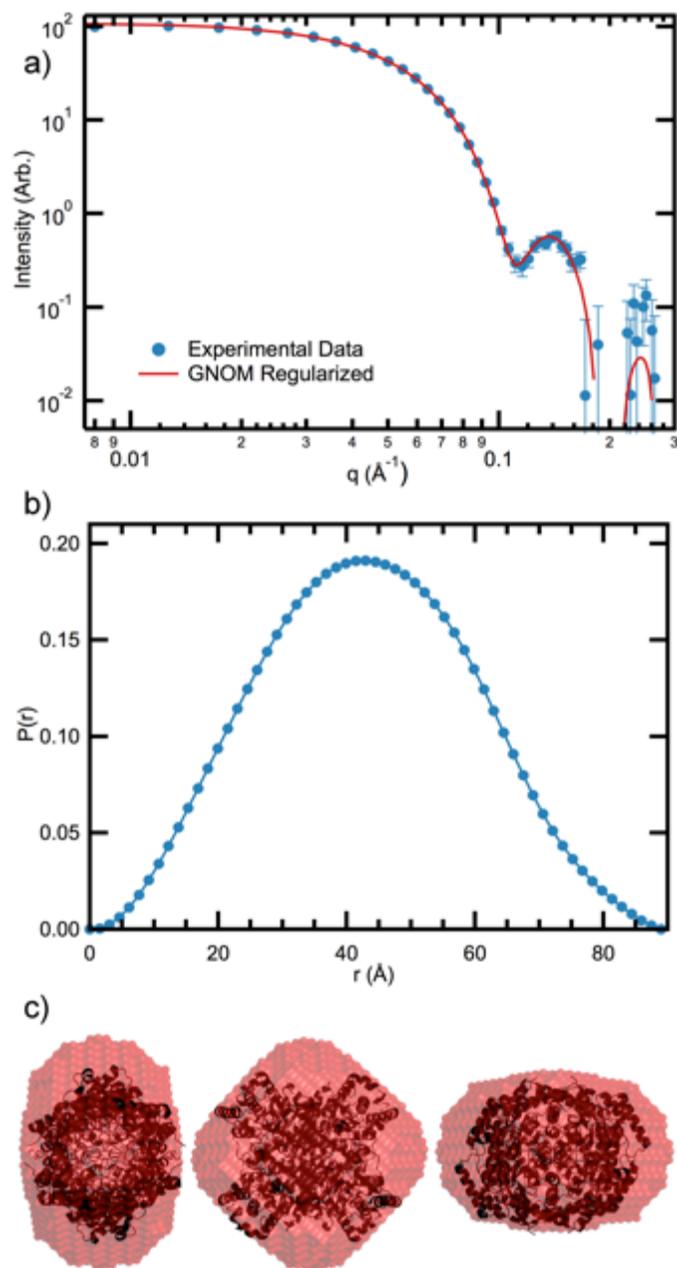


Figure 2.10 – All data in this figure are for glucose isomerase (2.8 mg/mL) at room temperature in PG. a) Experimental data, and regularized $I(q)$ from GNOM. $\chi^2 = 1.357$, total estimate = 0.915. b) $P(r)$ computed by GNOM with $D_{\text{max}} = 89$ \AA . c) Three orthogonal views of the DAMAVER envelope aligned with the glucose isomerase crystal structure.

(Kozak, 2005), and show the same $\sim 2\%$ increase on cooling from room temperature to 100 K observed by Meisburger et al. (2013). A protein's SAXS signal includes a contribution from its hydration layer of ordered water molecules. Cooling to 100 K increases hydration water ordering in protein crystals (Nakasako, 2002), and a similar effect in protein solutions could explain the observed increase in R_g at 100 K (Meisburger et al., 2013). Table 1 summarizes the above measured and literature R_g values and gives R_g values derived from Guinier fits.

Changes in the scattering profile can reflect changes in protein shape. Comparison of 3D envelope reconstructions, while less rigorous, allows for more intuitive visualization of differences. Envelope reconstructions were computed with DAMMIF in slow mode with P42 symmetry applied (Franke & Svergun, 2009; Kozak & Taube, 2009). At each temperature, twenty reconstructions were averaged with DAMAVER (Volkov & Svergun, 2003). The normalized spatial discrepancies between reconstructions at 100 K and room temperature were 0.27 ± 0.02 and 0.31 ± 0.04 respectively. The averaged structures were aligned to the crystallographic structure (PDB 1XIB with symmetry mates added in PYMOL) using SUPCOMB (Kozin & Svergun, 2001) and visualized in PYMOL version 1.3r1 (DeLano Scientific LLC). The averaged envelopes, shown in Figure 2.7b, fit the crystal structure well and show little variation between temperatures.

2.5.3 Comparison to theoretical scattering curves

When the crystal structure of a protein is known, theoretical scattering profiles can be generated and compared to experimental results. In previous cryoSAXS work, theoretical scattering profiles generated by CRY SOL (Svergun et al., 1995) agreed qualitatively with the experimental scattering profiles (Meisburger et al., 2013). However, CRY SOL's default parameters are for

cryoprotectant-free solvent at room temperature, whereas in cryoSAXS the solvent contains cryoprotectant and is at 100 K. Accurate theoretical predictions are necessary for quantitative comparison of theory and experiment, for decomposition of multi-component solutions and, in the present case, for evaluating how changes in protein and in solvent structure on cooling from room temperature to 100 K impact measured scattering profiles. Thus, the possible effects of solvent related parameter variations on CRY SOL's predictions were explored.

The relevant CRY SOL parameters are solvent electron density and protein hydration shell electron density. At room temperature, as described in Section 2.3.1, the electron density of a 36% w/w PG/water solution is $0.344 \text{ e}^-/\text{\AA}^3$ and a 45% w/w PEG/water solution is $0.355 \text{ e}^-/\text{\AA}^3$, and the electron density of pure water is $0.334 \text{ e}^-/\text{\AA}^3$. As the solution contraction on cooling has not been measured, the electron density at 100 K is not known. The effect of the cryoprotectant and cryocooling on the hydration shell density is also not known.

To simulate the effect of cryocooling, the solvent and hydration shell electron densities were both increased by a factor of 1.14. This factor was chosen as the fractional increase between the solution density at room temperature and the density of the high density amorphous (HDA) ice state of water (DeBenedetti, 2003). This density factor was used, first, because it is one of the few relevant experimentally available values and, second, because it is likely to provide an upper bound on the actual density increase on cooling. It was assumed that the hydration shell density follows the behavior of the bulk solution. Figure 2.11 compares the CRY SOL curves generated using both the default and modified parameters and experimental data collected at 100 K. The change from default to modified parameters changes R_g from 33.3 \AA to 34.5 \AA , and at higher q modulates the overall intensity and shifts the positions of scattering profile peaks. The residual between these curves is shown in Figure 2.12. While the modified parameters most likely do not

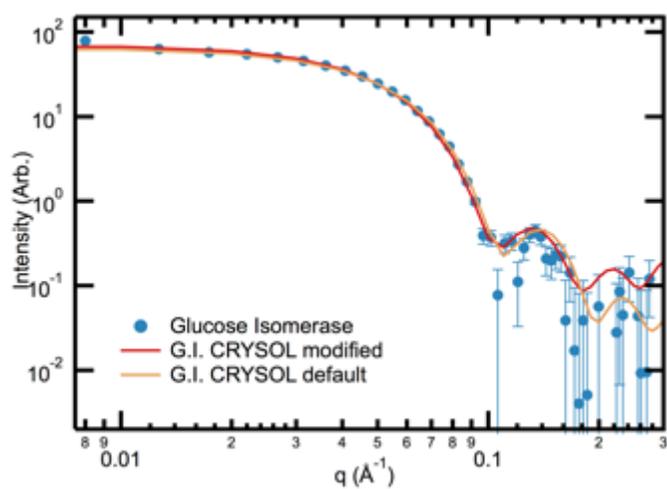


Figure 2.11 – Scattering profile of glucose isomerase at 100 K and theoretical scattering profiles generated by CRY SOL from a PDB file (tetramer generated with symmetry using 1XIB). Scaled to match at $q = 0.05 \text{ \AA}^{-1}$. The modified parameters are discussed in Section 2.4.3.

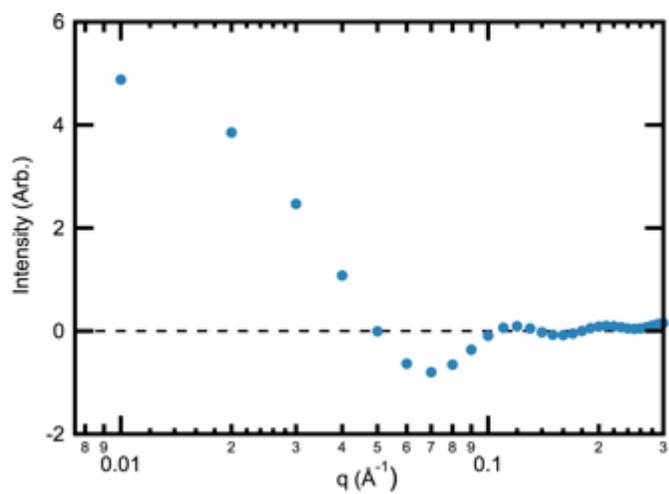


Figure 2.12 – Residual obtained by subtracting the scattering intensity calculated for glucose isomerase using CRY SOL with default solvent parameters from that calculated with modified solvent parameters. The original intensity curves are shown in Figure 2.11. The modified solvent parameters mimic the change of the solvent on cooling to 100 K, so this curve should be, and is, similar to that in Figure 2.8.

represent the actual sample conditions, differences in the CRY SOL profiles show that accurate determination of these parameters is necessary to generate quantitatively accurate predictions.

2.5.4 Repeatability

In cryoSAXS experiments using windowless sample holders (Meisburger et al., 2013), differences in sample thickness and cooling introduce variations in scattering profiles from separately loaded and cryocooled but otherwise identical solutions of the same sample, and these make analysis challenging. The silicon sample holders, with their rigid walls and reproducible positioning, largely eliminate these sources of irreproducibility. However, Figure 2.13a shows that differences in loaded solution volume also cause variations in the scattering profile. These differences occur even at room temperature, and even though the fill level in all cases is well above the X-ray beam's location so there can be no variation in path length. These variations are likely due to differences in how much parasitic scatter is absorbed by different volumes: larger solution volumes should absorb more parasitic scatter, yielding lower measured scatter, in agreement with the trend in Figure 2.13a. The use of the modified microliter syringe and syringe pump described in Section 2.3.3 enable reproducible loaded volumes. Figure 2.13b shows that, using this loading method, excellent agreement between scattering profiles from three different trials of the same buffer is observed down to $q \approx 0.02 \text{ \AA}^{-1}$. Buffer was used for these experiments because the lower signal is more sensitive to any changes in the scattering.

2.5.5 Origin of cooling-dependent variations in data below $q \sim 0.02 \text{ \AA}^{-1}$

For samples of a given composition, the 100 K scattering profiles obtained using silicon holders are reproducible down to $q \approx 0.02 \text{ \AA}^{-1}$. However, as seen in Figure 2.13b, discrepancies are

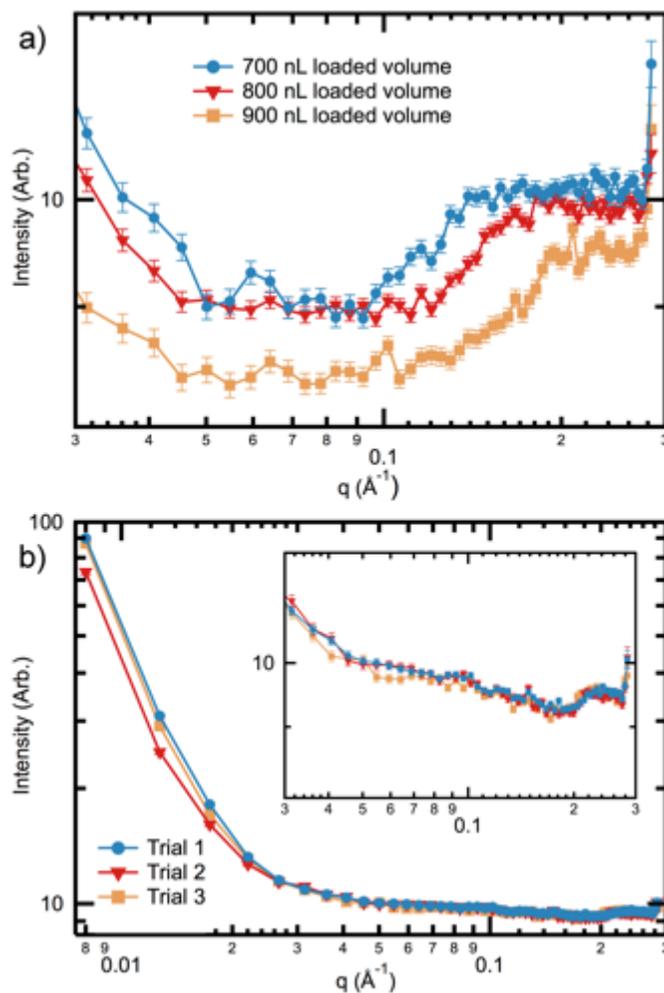


Figure 2.13 – a) Scattering profiles from three different loaded sample volumes. The q range is limited to emphasize differences in profile at high q . Lines added to guide eye. b) Scattering profiles from three trials of separately loaded and cooled but otherwise nominally identical buffer. These profiles agree well down to $q \sim 0.02 \text{ \AA}^{-1}$. Inset shows the same q and intensity ranges as a). Shapes of buffer in a) and b) are different because the data were taken in different sample holders. Lines added to guide eye.

observed at lower q values, resulting in imperfect background subtraction. This effectively limits the accessible q range to $q \gtrsim 0.02 \text{ \AA}^{-1}$. The limitations imposed by this apparent minimum q value can be calculated according to Shannon's sampling theorem, which states that full information about $I(q)$ is obtained as long as a minimum q value is measured that is less than the Shannon increment:

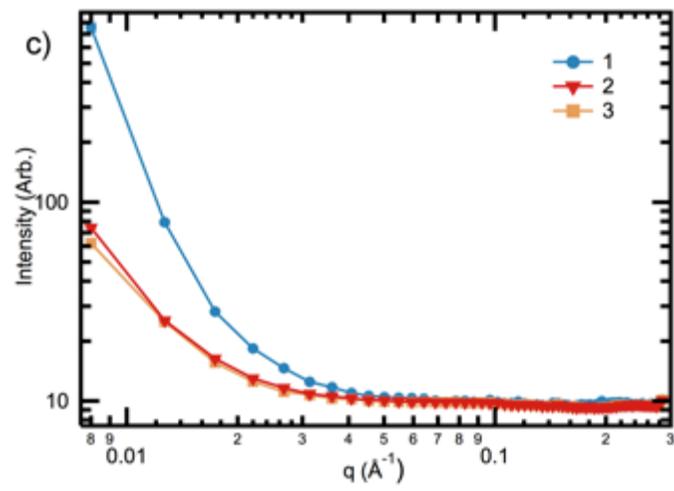
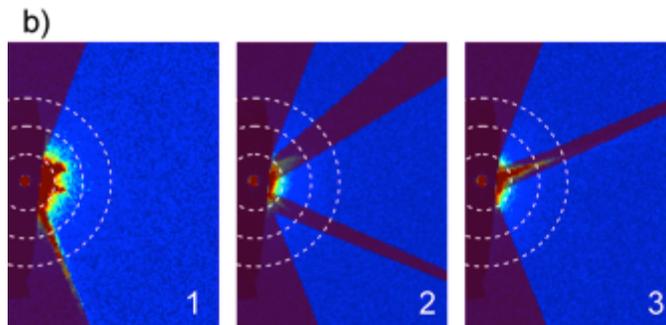
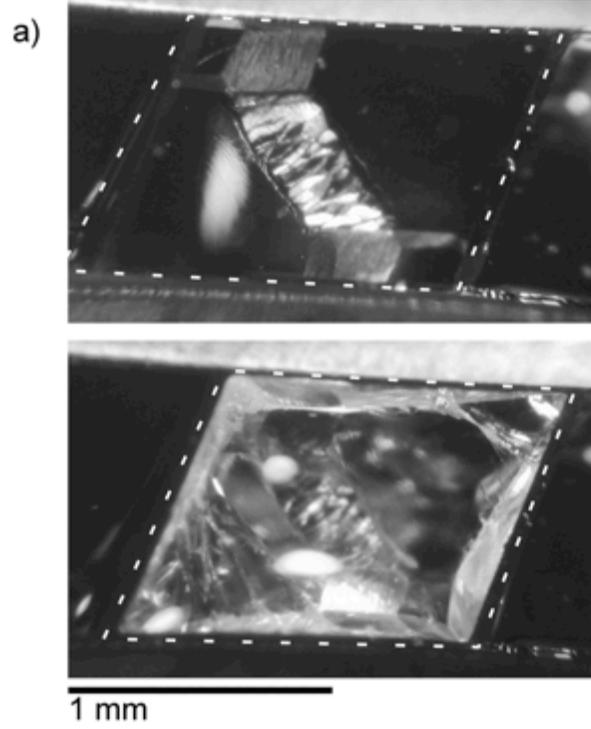
$$q_{min} = \frac{\pi}{D_{max}} \quad (2.1)$$

where D_{max} is the maximum dimension of the protein (Svergun & Koch, 2003). For $q_{min} = 0.02 \text{ \AA}^{-1}$, the largest value of D_{max} for which the scattering profile contains full information is 157 \AA . Some degree of oversampling may be necessary to accurately determine D_{max} , so the practical limit may be somewhat smaller than 157 \AA . Thus, the available q range for cryoSAXS using the silicon sample holders is adequate for macromolecules within this size range, including large multi-subunit enzymes such as glucose isomerase ($D_{max} \sim 90 \text{ \AA}$). However, most synchrotron beamlines for macromolecular SAXS reliably measure data below $q = 0.01 \text{ \AA}^{-1}$, and this extended q range can be useful for ascertaining the quality of the sample via detection of large aggregates, or for examining very large macromolecular complexes. Thus, an extended q range for cryoSAXS is desirable.

The discrepancies at low q are most likely due to sample fracturing, shown in Figure 2.14a. The fractures probably occur to relieve stress due to differential contraction of the sample and silicon holder that accumulates upon cooling. In cryocooled aqueous solutions, fracturing is observed in many geometries of rigid sample holders (Rall, 1987; Fahy et al., 1990; Steif et al., 2005; Rabin et al., 2006; Yavin & Arav, 2007).

Observed sample fractures are generally oriented with their interfaces parallel to the X-ray windows and thus perpendicular to the beam, minimizing their impact on the scattering profile.

Figure 2.14 – a) Optical images of a filled sample holder at room temperature (top) and the fractured sample after cooling to 100 K (bottom), top down view. Dashed lines show the boundaries of the sample holder. X-ray beam direction is vertically in the plane of the page. b) Detector images showing bright scatter with jets (1) and isolated jets (2 and 3). The red overlay shows the mask used to integrate each image for the curves in c). The dashed white circles correspond, with increasing radius, to $q = 0.01, 0.02, \text{ and } 0.03 \text{ \AA}^{-1}$. The color scale of each image was normalized by the counts in the direct beam so that the displayed intensities are comparable. c) Scattering profiles from images shown in b). Curve 1 was integrated even though the scatter was visibly anisotropic in the detector image. Lines added to guide eye.



The magnitude and q range of the excess scattering due to fractures depends on the size, interfacial texture, and orientation of the fractures. Fractures oriented along the beam may generate significant grazing incidence scatter. In general, scatter from fractures will be anisotropic.

Fractured samples show three general patterns of scatter in detector images: isotropic scattering similar to that observed at room temperature, excess isotropic scattering concentrated at low q , and strongly anisotropic “jets”. The latter two patterns are shown in Figure 2.14b. These features are not mutually exclusive, and jets can appear with either type of underlying isotropic scattering.

A major concern is determining the extent to which an image or scattering profile has been affected by fractures. As separately cryocooled but otherwise identical samples fracture in different ways, variation in scattering from different trials should be due to the fractures. The extent to which two different scattering profiles agree indicates which part of the profiles reflects the underlying scatter of the sample and which part has additional scattering from the fractures. Thus, collection of scattering profiles from at least two separately cryocooled but otherwise identical samples is needed.

For detector images with isolated jets, as shown in images 2 and 3 of Figure 2.14b, masking of the jets can result in scattering profiles that agree well to $q \sim 0.02 \text{ \AA}^{-1}$, as shown in Figure 2.14c. This indicates that the unmasked regions of the images are unaffected by the fractures and can yield accurate scattering profiles to low q . However, for images with many layered jets and/or excess isotropic scatter at low q , such as image 1 of Figure 2.14b, typically no mask is sufficient to achieve agreement with profiles from other samples at low q . The lack of agreement between

otherwise identical scattering profiles below $q \sim 0.02 \text{ \AA}^{-1}$ likely stems from isotropic scattering from the fractures, which cannot be removed by masking.

2.5.6 Attempts to mitigate fracturing

Fracturing in cryocooled samples can be reduced or eliminated by reducing the cryoprotectant concentration, increasing the final temperature, and reducing the cooling rate (Yavin & Arav, 2007). All of these changes should have the effect of reducing the magnitude of the elastic stresses that accumulate due to differential contraction between the sample and holder or within the sample (due to thermal gradients) during cooling.

The geometry and dimensions of the sample may also strongly influence fracturing (Yavin & Arav, 2007). For rigid sample holders used in biological cryopreservation of 100 μL and larger volumes, fractures occur in many holder geometries (Rall, 1987; Fahy et al., 1990; Steif et al., 2005; Rabin et al., 2006; Yavin & Arav, 2007), so eliminating fractures by changing geometry is not trivial.

In an attempt to eliminate fractures in the cryoSAXS samples, the sample cooling conditions and sample holder dimensions were varied. As discussed in Section 2.6, these did affect the fracturing behavior, but, for the range of experimental parameters explored, the best cryoSAXS scattering profiles were obtained with the sample holder dimensions, final temperature, and cooling conditions described in Sections 2.2 and 2.3 and used for data presented in this paper. Reductions in cryoprotectant concentration could not be explored because these experiments already used the minimum required to avoid ice formation.

By raising the final temperature, visible fracturing in these sample holders was eliminated. However, as discussed in Section 2.6, the scattering was irreproducible at both low and high q . Thus, elimination of fractures is not necessarily sufficient to improve scattering.

Reproducible scattering from unfractured samples was observed in cryoSAXS samples cooled in the windowless thin-wall polymer capillary sample holders of Meisburger et al. (2013). These holders freely allow sample contraction in the beam direction, and have flexible walls that may minimize stresses upon cooling. Therefore, reduction in the overall stresses beyond that required to prevent fracturing might be necessary to achieve reproducible scattering.

Numerical analysis of stress profiles is routinely used in mechanical design to optimize part shapes for minimal stresses, to eliminate fracturing, or to concentrate stresses and fracturing in non-critical regions. It has also been used in study of fracturing during cryopreservation (Steif et al., 2005; Rabin et al., 2006; Steif et al., 2007). Consequently, optimization of geometry, cooling protocols and solution compositions via modeling and experiment should allow high quality, isotropic scattering to low q values to be routinely achieved in fixed path-length sample holders for cryoSAXS. Even without further optimization, the current sample holders provide complete small angle scattering information over a q range sufficient for study of a wide range of biological macromolecules.

2.6 Conclusion

Small volume, fixed path length sample holders were fabricated for cryoSAXS experiments. The use of a fixed path length sample holder eliminates the need to measure three distinct scattering curves (buffer, sample, and instrument background) and to normalize by the transmission of the X-ray beam through the sample, as was required in previous work (Meisburger et al., 2013).

Using these sample holders, one can measure sample and buffer curves and normalize by transmitted beam intensity as in standard SAXS experiments, making cryoSAXS easier and more accessible.

Measurements in these holders are reproducible and agree with previous cryoSAXS results down to $q \approx 0.02 \text{ \AA}^{-1}$. Reliable background subtraction was demonstrated throughout the SAXS region. The lowest accessible q value is most likely limited by fractures that occur in the sample upon cryocooling. However, this is only prohibitive for proteins with $D_{\text{max}} \gtrsim 150 \text{ \AA}$ and may be improved with future designs. Consequently, these sample holders should facilitate cryoSAXS experiments on a wide range of macromolecules.

2.7 Supporting information

2.7.1 Efforts to mitigate fractures

A range of cryocooling conditions were tested in an attempt to eliminate fracturing. The magnitude of the elastic stresses that accumulate due to differential contraction between the sample and holder or within the sample during cooling is determined in part by the temperature range over which the sample is vitreous and “solid.” This temperature range can be reduced by reducing the cryoprotectant concentration (which lowers the vitrification temperature), by increasing the final temperature, and by reducing the cooling rate (which also lowers the effective vitrification temperature) (Fahy et al., 1990; Steif et al., 2005; Yavin & Arav, 2007).

The temperature to which the sample was cooled was varied from 90 K to 190 K. Samples cooled to final temperatures above ~ 140 K showed no visible fractures. However, scattering from samples cooled to ~ 140 K to 170 K was irreproducible, and was often anisotropic. Speculatively, this irreproducibility and anisotropy is due to nonuniform elastic stresses and to

microscale defects/fractures that give structure in the q range of interest. These likely arise as the sample cools through the region of extreme viscosity near the glass transition. Scattering from samples cooled to above ~ 170 K was isotropic, but at low q displayed a power law that in (Meisburger et al., 2013) was characteristic of ice formation. Increasing the cryoprotectant concentration in an attempt to find a temperature near 170 K at which scatter was reproducible and isotropic without the power law characteristic of ice was unsuccessful.

Two types of slow cooling were used in an attempt to reduce or eliminate fracturing. In the first, samples with larger cryoprotectant concentrations (to prevent ice formation) were cooled at 0.1 K/s from 250 K (to prevent initial evaporation) to the final temperature. In the second, samples with standard cryoprotectant concentrations were first cooled at ~ 25 K/s to an intermediate temperature such as 180 K and then cooled at 0.1 K/s to the final temperature. Both of these techniques slightly lowered the temperature at which fractures appeared, but did not alter the observed scattering behaviors.

Reducing the dimensions of the sample holder should reduce thermal gradients and peak sample stresses, but it also increases the sample cooling rate, which should increase the effective vitrification temperature and the accumulated stress at the final temperature (Yavin & Arav, 2007). Sample holders with X-ray path lengths from 0.1-1.0 mm, lateral dimensions from 0.5-2.0 mm, and thus sample volumes of ~ 10 -1500 nL were tested. The trade-off between reduced volume and increased cooling rates caused first a lowering and then a raising of the fracturing temperature as the volume was decreased, but did not change the general scattering behaviors.

Based on these observations, for these sample holders visibly fractured samples gave the best scattering profiles, perhaps because they relieve thermal stress with large scale features that

scatter at low q , minimally perturbing the q range of interest in SAXS. The conditions used in the main body of this work were the optimal conditions identified within the limited search space.

The tendency of fractures to form parallel to the walls of the sample holders suggests that the holder's geometry affects the pattern of fracture. Homogeneously stressed uniform samples can often sustain very large elastic stresses. Asymmetries in the current holders due to the design and to the presence of interior surfaces defined by $\{110\}$, $\{111\}$ and $\{311\}$ crystallographic planes result in nonuniform stresses in cryocooled samples. Fracture typically occurs at much lower average sample stresses due to geometry- or temperature-gradient-related stress concentration, and/or due to the presence of pre-existing defects like bubbles or inclusions (Anderson, 2004). Thus, the geometry may be chosen either to produce maximally uniform stresses or to concentrate stress and promote fracturing in non-critical regions so as to reduce fracturing in critical regions. Therefore, controlling the location and preferred orientation of fractures by sample holder geometry is a promising future research direction.

Elimination of fracturing is not a sufficient criterion to guarantee good scattering, it is likely that the overall stresses in the sample upon cooling must be greatly reduced from the current levels. Good scattering profiles have been collected from unfractured samples at 100 K (Meisburger et al., 2013). These profiles were collected from windowless, flexible, plastic sample holders, which likely minimized stresses from cooling (flexible sample holder materials also reduce fracturing in much larger samples (Rall, 1987)). Thus careful optimization of the sample holder design to eliminate or reduce fracturing will be carried out hand-in-hand with modeling of the stress effects upon cooling to minimize overall stress in the sample.

REFERENCES

- Anderson, T. L. (2004) *Fracture Mechanics*. (CRC Press, Boca Raton, FL).
- Bean, K. (1978). *Electron Devices, IEEE Trans.* **25**, 1185–1193.
- Blundell, T. L., Jhoti, H., & Abell, C. (2002). *Nat. Rev. Drug Discov.* **1**, 45–54.
- Debenedetti, P. G. (2003). *J. Phys. Condens. Matter.* **15**, R1669–R1726.
- Dwivedi, V. K., Gopal, R., & Ahmad, S. (2000). *Microelectronics J.* **31**, 405–410.
- Ezoe, Y., Koshiishi, M., Mita, M., Mitsuda, K., Hoshino, A., Ishisaki, Y., Yang, Z., Takano, T., & Maeda, R. (2006). *Appl. Opt.* **45**, 8932–8938.
- Fahy, G. M., Saur, J., & Williams, R. J. (1990). *Cryobiology.* **27**, 492–510.
- Franke, D. & Svergun, D. I. (2009). *J. Appl. Crystallogr.* **42**, 342–346.
- Heuberger, A. (1990). *J. Electrochem. Soc.* **137**, 3612–3626.
- Hölke, A. & Henderson, H. (1999). *J. Micromechanics Microengineering.* **9**, 51–57.
- Hong, X. & Hao, Q. (2009). *Rev. Sci. Instrum.* **80**, 014303.
- 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.
- Iwamoto, H. (2009). *J. Synchrotron Radiat.* **16**, 336–345.
- Kam, Z., Koch, M. H. J., & Bordas, J. (1981). *Proc. Natl. Acad. Sci. U. S. A.* **78**, 3559–3562.
- Kaminsky, G. (1985). *J. Vac. Sci. Technol. B.* **3**, 1015–1024.
- Kendall, D. L. (1979). *Annu. Rev. Mater. Sci.* **9**, 373–403.
- Kern, W. (1990). *J. Electrochem. Soc.* **137**, 1887–1892.
- Khattab, I. S., Bandarkar, F., Khoubnasabjafari, M., & Jouyban, A. (2012). *Arab. J. Chem.*
- Kim, B. & Cho, D. D. (1998). *J. Electrochem. Soc.* **145**, 2499–2508.
- Kozak, M. (2005). *J. Appl. Crystallogr.* **38**, 555–558.

- Kozak, M. & Taube, M. (2009). *Radiat. Phys. Chem.* **78**, S125–S128.
- Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Crystallogr.* **34**, 33–41.
- Kraft, P., Bergamaschi, A., Broennimann, C., Dinapoli, R., Eikenberry, E. F., Henrich, B., Johnson, I., Mozzanica, A., Schlepütz, C. M., Willmott, P. R., et al. (2009). *J. Synchrotron Radiat.* **16**, 368–375.
- Meisburger, S. P., Warkentin, M., Chen, H., Hopkins, J. B., Gillilan, R. E., Pollack, L., & Thorne, R. E. (2013). *Biophys. J.* **104**, 227–236.
- Mertens, H. D. T. & Svergun, D. I. (2010). *J. Struct. Biol.* **172**, 128–141.
- Muller, E. A. & Rasmussen, P. (1991). *J. Chem. Eng. Data.* **36**, 214–217.
- Nakasako, M. (2002). *J. Biol. Phys.* **28**, 129–137.
- 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.
- Pal, P., Sato, K., & Hida, H. (2011). *2011 Int. Symp. Micro-NanoMechatronics Hum. Sci.* 55–59.
- Rabin, Y., Steif, P. S., Hess, K. C., Jimenez-Rios, J. L., & Palastro, M. C. (2006). *Cryobiology.* **53**, 75–95.
- Rall, W. F. (1987). *Cryobiology.* **24**, 387–402.
- Roessle, M. & Svergun, D. I. (2011). *Present. “XDL-2011 Sci. Hard X-Ray Diffr. Limit.”*
- Steif, P. S., Palastro, M., Wan, C.-R., Baicu, S., Taylor, M. J., & Rabin, Y. (2005). *Cell Preserv. Technol.* **3**, 184–200.
- Svergun, D., Barberato, C., & Koch, M. H. J. (1995). *J. Appl. Crystallogr.* **28**, 768–773.
- Svergun, D. I. (1992). *J. Appl. Crystallogr.* **25**, 495–503.
- Svergun, D. I. & Koch, M. H. J. (2003). *Reports Prog. Phys.* **66**, 1735–1782.
- Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Crystallogr.* **36**, 860–864.
- Wu, B., Kumar, A., & Pamarthy, S. (2010). *J. Appl. Phys.* **108**, 051101.
- Yavin, S. & Arav, A. (2007). *Theriogenology.* **67**, 81–89.

CHAPTER 3

QUANTIFYING RADIATION DAMAGE IN BIOMOLECULAR SMALL-ANGLE X-RAY SCATTERING

3.1 Preface

In order to properly discuss how much cryoSAXS increases the radiation tolerance of a sample, we must first be able to properly quantify radiation damage at room temperature. While there are a few published studies (summarized in Section 3.2), we felt that these did not provide a precise method for quantifying radiation damage. The rest of this chapter is a manuscript detailing our methods for properly quantifying dose in biological SAXS, and the application of those methods to three standard molecules. The manuscript has been submitted to the Journal of Applied Crystallography, and is currently out for review. We found that different molecules show different modes of damage at room temperature, and that no single metric properly captures all of the behavior observed.

3.2 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. 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, to minimize dose (Fischetti et al., 2003). 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 are 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 develop 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), accounting for variables including beam shape, diffusive exchange of protein into and out of the beam, and beam heating. 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 3.11.1; 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.

3.3 Overview of a radiation damage experiment

The general steps required to quantify radiation damage in a 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.

3.4 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 in general is calculated as

$$\text{Dose} = \frac{ftAE_{\gamma}}{\rho l}, \quad (3.1)$$

where f is the flux density (ph/s/ μm^2), 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.

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.

3.5 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 3.1 and 3.2.

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 unavailable or not fixed, the sample transmission can be measured 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). For solutions with modest protein, salt and buffering agent concentrations, they report densities 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/cm^3 .

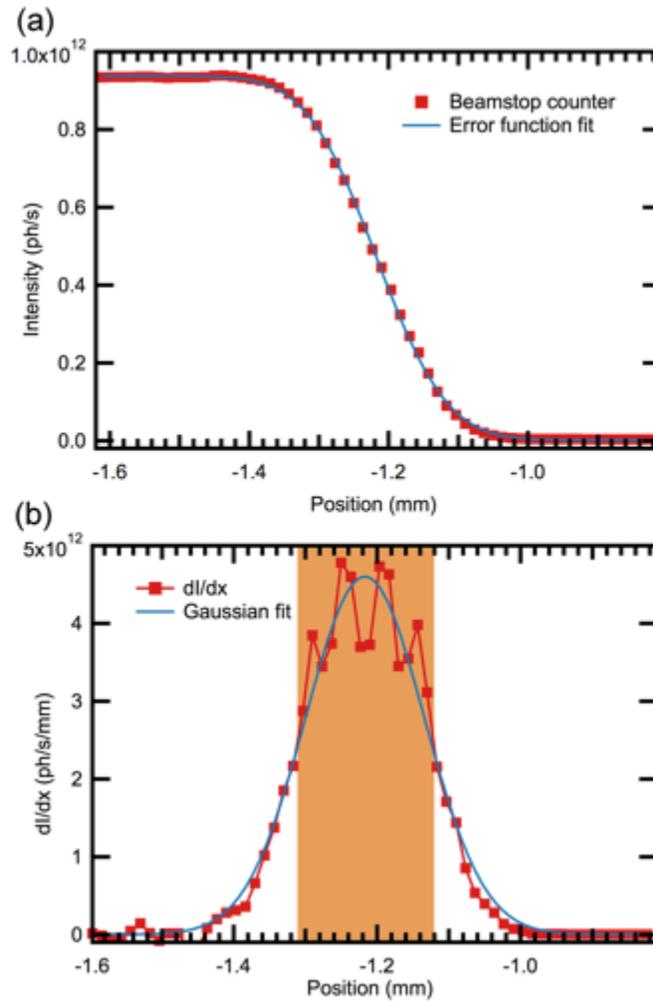


Figure 3.1 – a) The beamstop counter output (converted to flux) as a function of the position of a knife edge being scanned through the beam in the x direction. A Gaussian beam should give an error function shape. The blue line shows an error function fit with FWHM of 188 μm . b) The numerical derivative 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 .

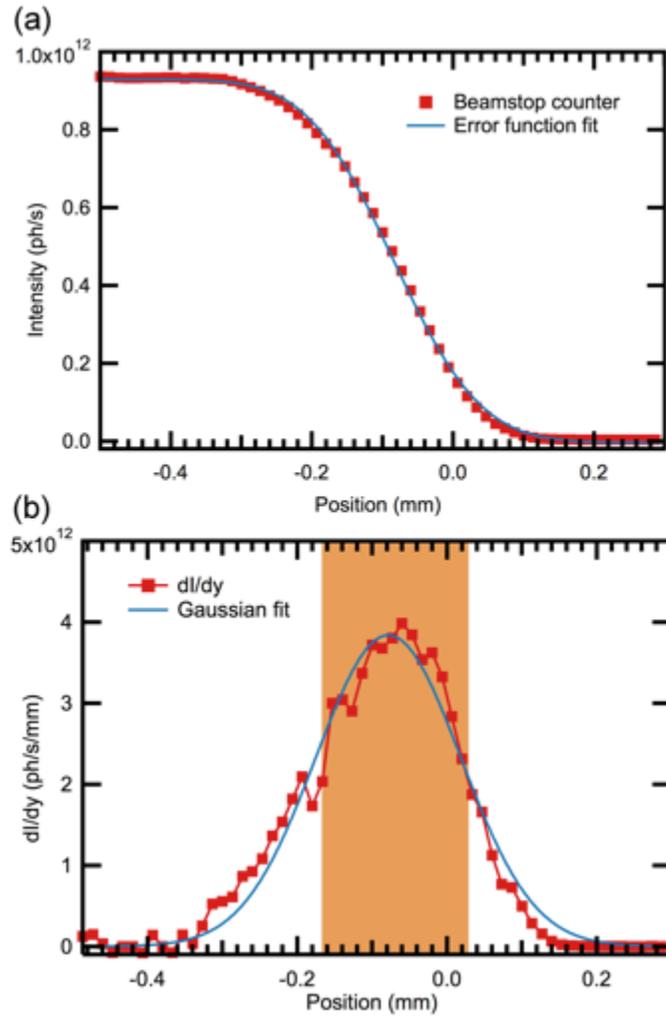


Figure 3.2 – a) The beamstop counter output (converted to flux) as a function of the position of a knife edge being scanned through the beam in the y direction. A Gaussian beam should give an error function shape. The blue line shows an error function fit with FWHM of $229 \mu\text{m}$. b) The numerical derivative of part a, giving the beam shape in the y direction. A Gaussian fit, FWHM $228 \mu\text{m}$, is shown. The shaded orange region is the actual FWHM of the data, $196 \mu\text{m}$.

3.6 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 frames, 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 3.3 shows sample SAXS scattering profiles at increasing accumulated dose (calculated as described in Section 3.7) for lysozyme, glucose isomerase and xylanase, 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 (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 . 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.

To reduce the effects of radiation damage on SAXS data, most current SAXS experiments either oscillate or flow the sample in the beam, to spread the flux over a larger total sample volume and lower the dose (Fischetti et al., 2003; Lipfert et al., 2006; Skou et al., 2014). This complicates the dose calculation, as the sample mass illuminated by the X-rays is harder to determine, so static experiments are recommended for accurate quantification. Even in static experiments, the sample

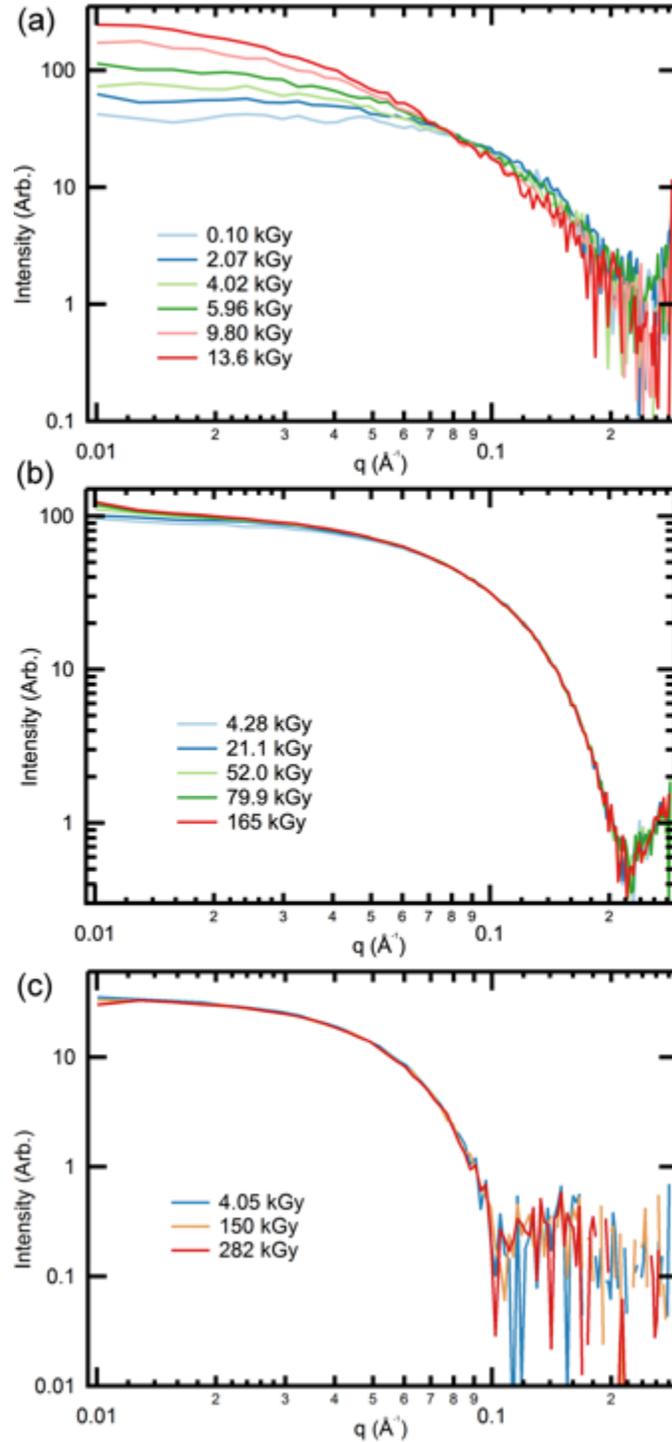


Figure 3.3 – Select scattering profiles of (a) 4.1 mg/mL lysozyme, (b) 4.9 mg/mL xylanase, and (c) 1.2 mg/mL glucose isomerase as a function of dose. For lysozyme and xylanase, the damage manifests primarily as an increase in intensity at low q , indicating aggregation. For glucose isomerase, only a slight downturn is observed at the lowest q , which could be attributable to damage or charging.

plug may slowly creep or drift, bringing fresh sample into the beam; samples should be carefully observed or imaged before and after each measurement.

3.7 Calculation of dose

Accurate calculation of dose using Eq. (3.1) requires accurate determination of all parameters. The flux density and absorbed energy fraction can be subject to a variety of corrections, as outlined below.

3.7.1 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. (3.1) is given by

$$\bar{f} = \frac{\int f(x,y)f(x,y)dA}{\int f(x,y)dA}, \quad (3.2)$$

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

For a Gaussian beam,

$$f(x,y) = \frac{f_0}{\sqrt{2\pi\sigma_x\sigma_y}} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right], \quad (3.3)$$

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. (3.2) gives

$$\bar{f} = \frac{f_0}{4\pi\sigma_x\sigma_y}. \quad (3.4)$$

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. (3.2) is

$$\frac{f_{FWHM}}{\bar{f}} = \frac{\pi}{2\ln 2} \approx 2.27. \quad (3.5)$$

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. (3.2) 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.

3.7.2 Determining the absorbed energy fraction

In Eq. (3.1), the fraction of the incident energy that is absorbed, A , can be calculated from Beer's law,

$$A = 1 - \exp\left(-\left(\frac{\mu}{\rho}\right)\rho l\right), \quad (3.6)$$

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 3.11.2, 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. (3.6) to estimate A .

3.7.3 Correcting for diffusive turnover

Eq. (3.1) 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 frames), 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

$$t_d = \frac{l_{FWHM}^2}{4D} . \quad (3.7)$$

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 3.11.3 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. (3.1) 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 3.11.1 for details) — the maximum diffusion corrections determined by the calculations in section 3.11.3 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 3.4 shows how much diffusion reduces the dose, calculated as in Section

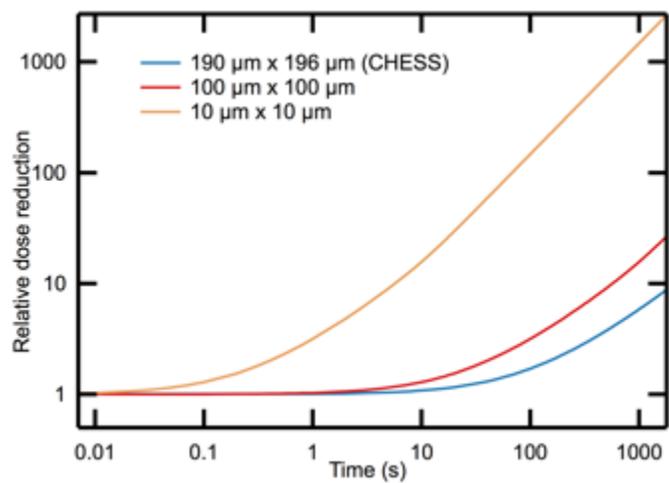


Figure 3.4 – Dose reduction by diffusive exchange of lysozyme versus time for three beam sizes: $190 \times 196 \mu\text{m}^2$ (used for the experiments in this paper), $100 \times 100 \mu\text{m}^2$, and $10 \times 10 \mu\text{m}^2$. The vertical axis is the (unitless) factor by which diffusion reduces the dose, relative to the expected dose without diffusion. These reduction factors were calculated as described in Section 3.11.3.

3.11.3 relative to the expected dose without diffusion, for lysozyme irradiated by different beam sizes.

3.8 Quantifying radiation damage data

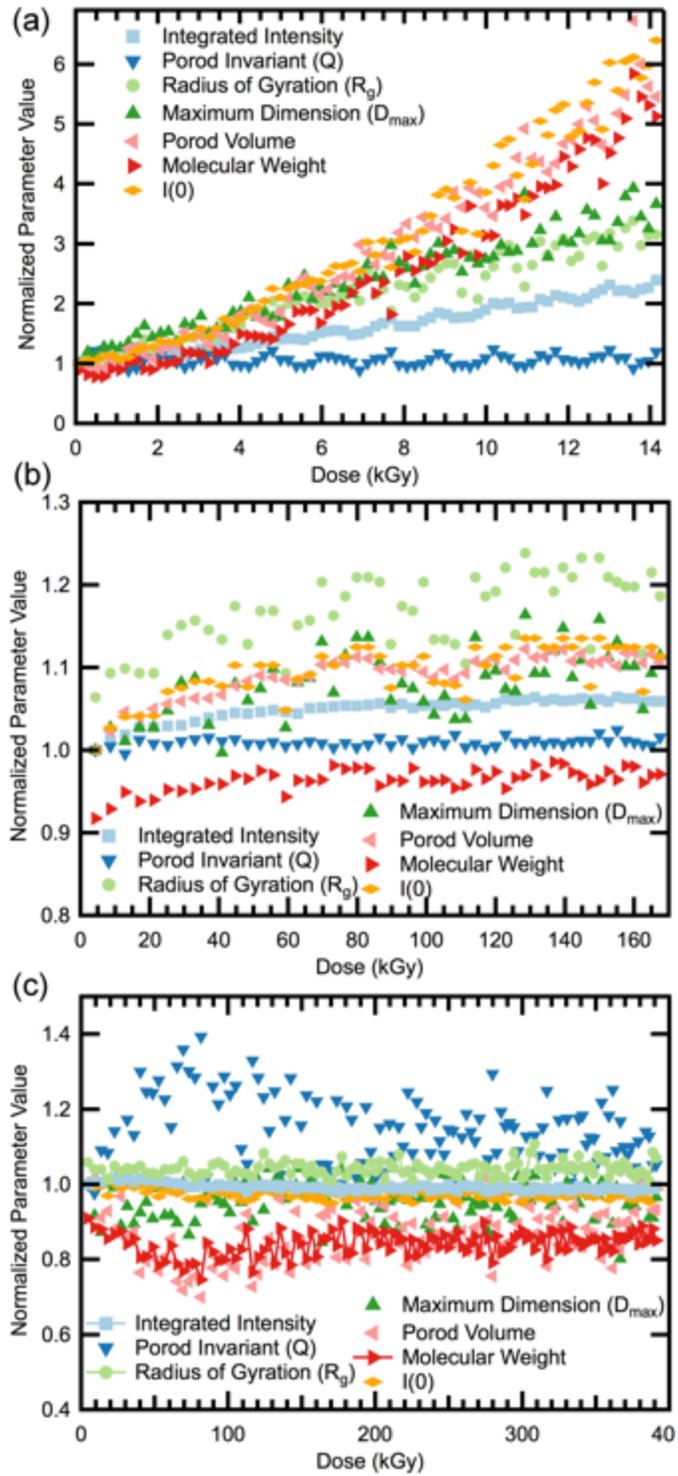
3.8.1 Calculate parameters from scattering profiles, normalize

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 frames, the parameter can be normalized by its value in the first frame, 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 additional metrics for a total of nine, including the Porod Invariant, the Porod volume, the molecular mass, the maximum dimension, and the integrated intensity. Details of these calculations are given in Section 3.11.1.

Figure 3.5 shows representative dose curves calculated from our sample data sets for lysozyme, xylanase, and glucose isomerase, for these nine metrics. These dose curves show how radiation damage changes the information available from the SAXS scattering profile.

Figure 3.5 – 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.



3.8.2 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 3.1 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 3.11.4. Figure 3.6 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 3.11.4 gives a detailed discussion of the behavior of these samples vs. dose, including possible sources of sample-to-sample variability.

3.8.3 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 3.7-3.9. For lysozyme, Figure 3.7, 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 3.5a, the Porod invariant doesn't change for these samples, so in this case, any other parameter would characterize radiation damage in the system.

Table 3.1 – 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

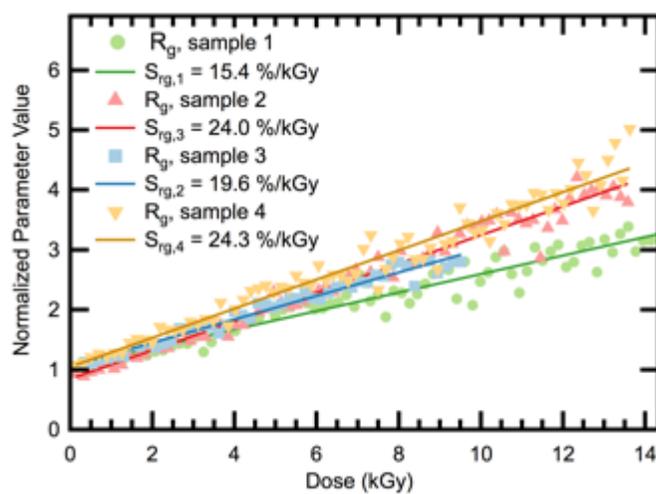


Figure 3.6 – The change in normalized R_g as a function of dose for four identically prepared lysozyme samples. The lines represent the best linear fit, and the values of the slopes are reported in the legend. For nominally identical samples, sensitivities ranged from 15.4 %/kGy to 24.3 %/kGy, emphasizing the necessity of measuring and averaging sensitivities from multiple samples.

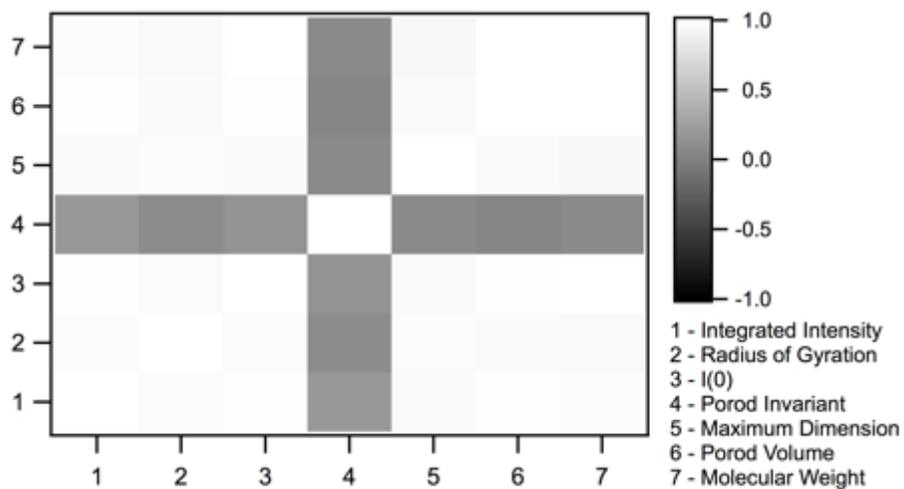


Figure 3.7 – Plot of the average pairwise Pearson's correlation coefficient, r , for seven of the parameters calculated for each scattering profile for 4.1 mg/mL lysozyme. This shows that every parameter except the Porod invariant is strongly correlated.

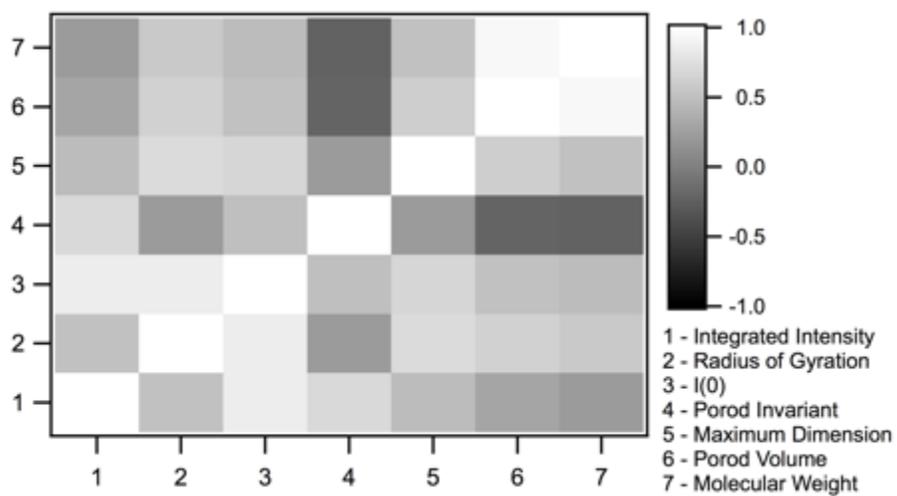


Figure 3.8 – Plot of the average pairwise Pearson's correlation coefficient, r , for seven of the parameters calculated for each scattering profile for 4.9 mg/mL xylanase. This shows much less correlation than Figure 3.3.

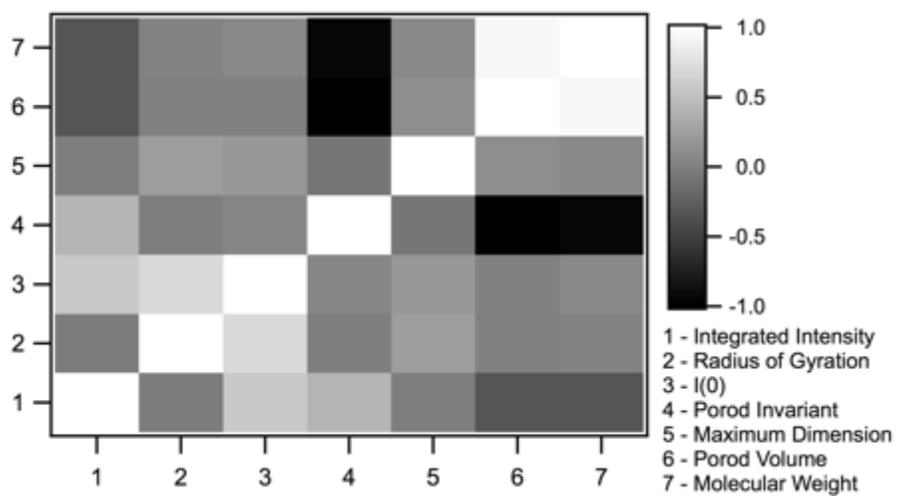


Figure 3.9 – Plot of the average pairwise Pearson’s correlation coefficient, r , for seven of the parameters calculated for each scattering profile for 1.2 mg/mL glucose isomerase. The only strong correlations seen are for the molecular weight and Porod volume.

For xylanase, Figure 3.8, there was less overall correlation. $I(0)$ correlated strongly with the integrated intensity and R_g ($r \approx 0.84$), and the molecular weight correlated strongly with the Porod volume ($r = 0.93$). D_{max} correlated with R_g , $I(0)$, and Porod volume ($r \approx 0.7$), 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 3.9 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, ($r \approx -0.95$). 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 3.11.1) 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 3.11.5. 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.

3.8.4 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.

3.9 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.

3.9.1 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 3.11.4) 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 3.11.4, 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 3.11.4 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^{*-} , 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).

3.9.2 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.

3.9.3 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 3.11.6, 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 3.2 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. 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.

3.9.4 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. Our data suggests that one particular definition of a critical dose will not work for all proteins. For example, if we apply the definition of (Jeffries et al., 2015) to our glucose isomerase data, we find a critical dose of ~66,000 kGy. However, the molecular weight shows a significant (~13%) change after just ~75 kGy.

The critical dose for a system will depend upon the metrics / macromolecular properties of interest. For example, R_g is typically determined to an accuracy of ~0.1-0.2 Å (in the best

Table 3.2 – 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 (3.1), 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); BM29, assuming 1.8 mm path length (ESRF); 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 (MacCHESS)	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

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 ~10%, and our lysozyme data then gives a dose limit of ~0.2-0.3 kGy. The intrinsic "noise" and accuracy 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.

3.10 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.

3.11 Supporting information

3.11.1 Methods

3.11.1.1 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 μ m 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). 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.

3.11.1.2 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 ($d = 58.38 \text{ \AA}$ (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 \AA^{-1} to 0.2822 \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 frame 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. 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 3.1 and 3.2. 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).

3.11.1.3 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.

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 frame. 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).

Pseudo R_g and initial damage rate – The pseudo R_g (R_g^{ps}) defined in (Jeffries et al., 2015) was calculated by manually finding $R_{g,0}$, the R_g for the first frame in every data set, and then calculating the R_g^{ps} for each subsequent frame using a Guinier fit in the q range $0.8 < qR_{g,0} < 1.3$.

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 frames 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 frames 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 mass – The molecular mass 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).

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 mass: 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 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.

3.11.2 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:

$$A = 1 - \exp(-\mu l) \quad (3.8)$$

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

$$A = 1 - \exp(-(\mu_{pe} / \rho)(l\rho)). \quad (3.9)$$

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 (3.9) to calculate A , or the transmission itself can be adjusted according to the difference between the mass attenuation and photoelectric absorption coefficient as

$$A_{meas} = 1 - T_{meas} \exp\left[(\mu / \rho - \mu_{pe} / \rho)\rho l_m\right]. \quad (3.10)$$

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

$$A = 1 - \exp(-(\mu / \rho)(l\rho)) \approx (\mu / \rho)(l\rho). \quad (3.11)$$

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.

3.11.3 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

$$\frac{\partial F(t,x,y,z)}{\partial t} = D \left(\frac{\partial^2 F(t,x,y,z)}{\partial x^2} + \frac{\partial^2 F(t,x,y,z)}{\partial y^2} + \frac{\partial^2 F(t,x,y,z)}{\partial z^2} \right) + B(t,x,y,z) \quad (3.12)$$

where $F(t,x,y,z)$ is the fluence ($\text{ph}/\mu\text{m}^2$) (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 was used Equation (3) from the main paper, was used. These assumptions give

$$\frac{\partial F(t,x,y)}{\partial t} = D \left(\frac{\partial^2 F(t,x,y)}{\partial x^2} + \frac{\partial^2 F(t,x,y)}{\partial y^2} \right) + B(x,y). \quad (3.13)$$

Equation (3.13) 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,

$$W(t) = \frac{\int B(t,x,y)F(t,x,y) dx dy}{\int B(t,x,y) dx dy} . \quad (3.14)$$

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 $W_0(t)$,

$$C(t) = \frac{W(t)}{W_0(t)} . \quad (3.15)$$

Figure 3.4 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, Worchestershire, 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 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 $<0.2 \mu\text{m}$ 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.

3.11.4 Effect of radiation on lysozyme, xylanase, and glucose isomerase

3.11.4.1 Measured Sensitivities

Scattering profiles at several doses are shown in Figure 3.3 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 3.11.1 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 3.5 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. The rate of change in molecular weight was mirrored by 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, the rate of change in the radius of gyration and the maximum dimension were quite similar, 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 same mirroring of dose curves was seen here as for lysozyme.

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 3.10. This plot demonstrates two things. First, for lower concentrations the dose curves tended to be noisier than at higher

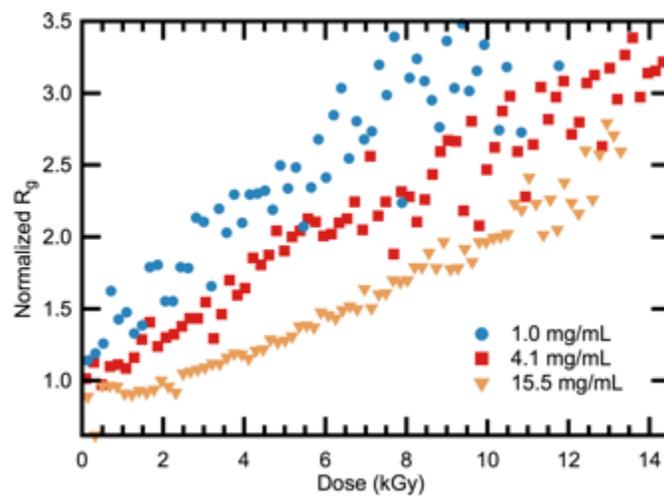


Figure 3.10 – Plot of the normalized R_g dose curves for lysozyme at 1.0, 4.1, and 15.5 mg/mL. The lower concentration shows more scattering in the value of the R_g . The 15.5 mg/mL curve has an initial plateau, a delay to the onset of observable damage.

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 3.8.2 (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 3.1 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 3.3 and clearly show that there can be large sample-to-sample variations in damage rate for identically prepared samples.

3.11.4.2 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 mass, 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

Table 3.3 – 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

(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 3.3 and 3.5. 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 14.0, 10.8 and 7.5 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 3.5. 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 GI samples showed similar decreases, though the change in the molecular weight was not quite as large. 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, $S_{rg}^{bys} / S_{rg}^{gi} = -532000$, $S_{mw}^{bys} / S_{mw}^{gi} = -504$, and $S_I^{bys} / S_I^{gi} = -907$. 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.

3.11.4.3 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 3.11 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

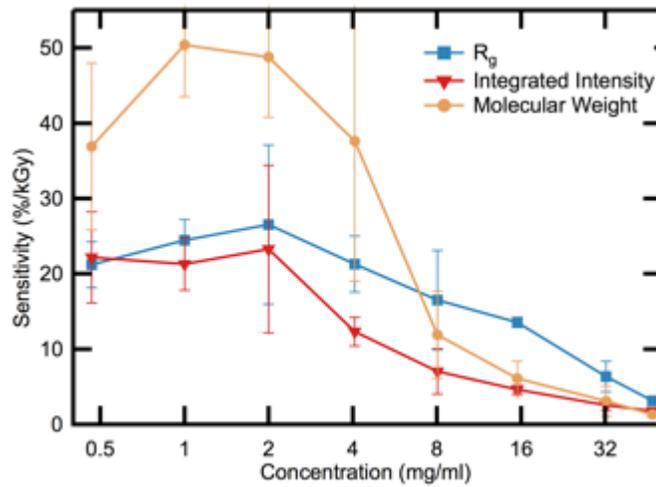


Figure 3.11 – The R_g , molecular weight, and integrated intensity sensitivity as a function of lysozyme concentration. There is a region of either increasing or flat sensitivity at low concentration, followed by a decrease in sensitivity with concentration.

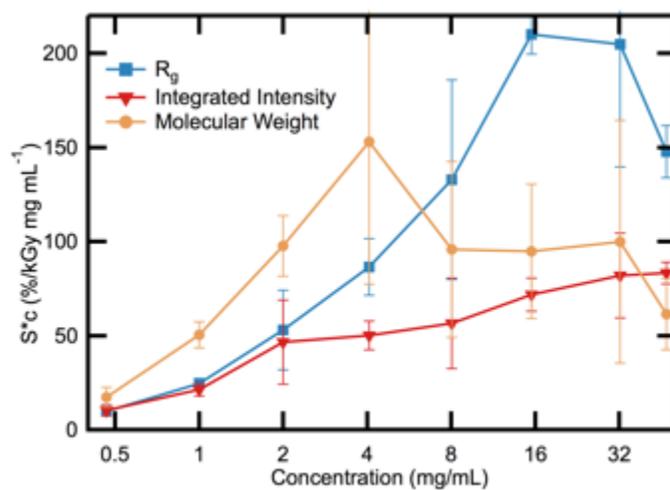


Figure 3.12 – Plot of sensitivity times concentration vs. concentration for S_{rg} , S_{mw} , and S_I for lysozyme. No clear plateau is seen in the plot.

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 $S_l c$ are not constant for our data, Figure 3.12, 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 3.10. 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.

3.11.4.4 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 3.11.7 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, $O_2^{\bullet-}$, 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).

3.11.5 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 3.7-3.9. Some analysis is given in the main body of the paper, Section 3.8.3.

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.

3.11.6 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

$$\delta T_{ad} = \frac{DR}{c}, \quad (3.16)$$

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

$$t_d = \frac{L^2}{\alpha}, \quad (3.17)$$

where L is a characteristic sample length and $\alpha \approx 0.14 \cdot 10^{-6} \text{ m}^2/\text{s}$ 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

$$\Delta T_{ad} = \delta T_{ad} t_d = \left(\frac{DR}{c} \right) \left(\frac{L^2}{\alpha} \right), \quad (3.18)$$

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:

$$T(r) = \frac{DRr_1^2}{2k} \log\left(\frac{r_2}{r}\right) + \frac{DRr_1^2}{2r_2h} + T_{ambient} . \quad (3.19)$$

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

$$\Delta T_{ss} = T(r_1) - T_{ambient} . \quad (3.20)$$

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 mm x 0.2 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

$$t^* = 43 \left(\frac{k\nu}{g\gamma\alpha^{1/2}q'} \right)^{2/3}, \quad (3.21)$$

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

$$t^* = 43 \left(\frac{0.533 \text{ s}^{3/2} \text{ W/m}}{q} \right)^{2/3}. \quad (3.22)$$

Using our dose rate of 5.1 kGy/s, and assuming that all the 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 3.2 (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 3.2 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.

3.11.7 Rate of oxygen absorption in degassed samples

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

$$\frac{1}{A} \frac{dW}{d\theta} = k_L (C_g - C_L) , \quad (3.23)$$

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 $5 \text{ }^\circ\text{C}$ is $12.8 \text{ mg/L} = 1.28 \cdot 10^{-5} \text{ g/cm}^3$ (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

$$\frac{dW}{d\theta} = Ak_L \left(C_g - \frac{W}{V} \right) . \quad (3.24)$$

This differential equation has the solution

$$W(\theta) = \frac{C_g V}{V} \left(1 - e^{-\frac{Ak_L \theta}{V}} \right) , \quad (3.25)$$

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

$$\frac{C_L(\theta)}{C_g} = \left(1 - e^{-\frac{Ak_L\theta}{V}} \right). \quad (3.26)$$

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

References

- Acerbo, A., Cook, M., & Gillilan, R. (2015). *J. Synchrotron Radiat.* **22**, 1–7.
- 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.
- 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**, 1–13.
- 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 v. 1*, 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. Eng. 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. (1987). *J. Biol. Chem.* **262**, 9895–9901.
- Davies, K. & Delsignore, M. (1987). *J. Biol. Chem.* **262**, 9908–9913.
- Davies, K., Delsignore, M., & Lin, S. (1987). *J. Biol. Chem.* **262**, 9902–9907.
- Davies, M. J. (2012). *Encyclopedia of Radicals in Chemistry, Biology and Materials v. 3*, C. Chatgililoglu, & A. Studer, edited by, pp. 1425–1457. Hoboken, N.J.: John Wiley & Sons.
- Diamond Technical Specifications.
- 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., Hammel, M., & Rambo, R. (2014). *Struct. Genomics Gen. Appl.* **1091**,.

ESRF BM29 Beamline Specifications.

ESRF 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. (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**, 1–7.
- 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., Cowieson, N., Samardzic-Boban, V., Felzmann, U., Mudie, N., & Dwyer, J. (2013b). *ACA Trans. 2013.* 27–39.
- Kmetko, J., Husseini, N., Naides, M., Kalinin, Y., & Thorne, R. (2006). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 1030–1038.
- Kmetko, J., Warkentin, M., English, U., & Thorne, R. E. (2011). *Acta Crystallogr. 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**, 10–13.
- 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., & Carman, 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**, 1–3.
- 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.

CHAPTER 4

CRYOSAXS INCREASES RADIATION TOLERANCE BY THREE ORDERS OF MAGNITUDE

4.1 Preface

CryoSAXS is a technique with a large potential upside. In the original paper we reported that the radiation tolerance of the sample increased by 2 to 5 orders of magnitude when it was cooled to 100 K. We wanted to make this more precise, so we developed a set of methods for quantifying radiation damage (Chapter 3), and applied them to cryoSAXS experiments. Our goal was to compare two proteins, lysozyme, which is very sensitive at room temperature, and glucose isomerase, which is very insensitive at room temperature. Unfortunately, the lysozyme samples did not yield high quality cryocooled results, despite having previously done so. This chapter focuses on the results for glucose isomerase, the lysozyme data is more briefly presented in Section 4.7. Supporting figures have been placed in the main text as in previous chapters. We refer to Chapter 3 as (Hopkins & Thorne, 2016), a submitted paper draft.

4.2 Introduction

Biological small angle X-ray scattering (SAXS) is a solution based structural technique that provides low-resolution information on biological macromolecules. The biomolecules need only be soluble and monodisperse, and effects of buffer conditions, including pH and ionic strength, as well as the presence of ligands, substrates, or other additives can be examined. This versatility makes SAXS a powerful structural technique even though the resolution is limited, as it can investigate macromolecules and systems that are difficult or impossible to study with other

methods. SAXS can provide static and time resolved information on oligomeric states, flexible macromolecules, complex formation, domain motion, interconversion between multiple allosteric states, weakly interacting systems, macromolecular folding and unfolding, and conformational changes (Jacques & Trehwella, 2010; Blanchet & Svergun, 2013; Petoukhov & Svergun, 2013; Kirby & Cowieson, 2014; Vestergaard & Sayers, 2014; Chaudhuri, 2015).

Conventional SAXS experiments are carried out between 277 and 293 K, and sample preparation, handling, and measurement at these temperatures is routine (Dyer et al., 2014; Skou et al., 2014). However, at these temperatures samples can be damaged by the X-ray beam, corrupting the data obtained from the sample (Kuwamoto et al., 2004; Jeffries et al., 2015; Hopkins & Thorne, 2016). The need to avoid radiation damage limits the minimum sample volume (~10 μ L) and exposure times that can be used in SAXS (Dyer et al., 2014; Skou et al., 2014). Standard macromolecule concentrations for SAXS are ~1 mg/mL, so radiation limits the samples that can be studied to those that can be expressed and purified in relatively large quantities (typically several 10s of micrograms). Sample requirements for time resolved SAXS, an increasingly popular technique, are even larger, anywhere from 1-300 mg per time series (Pollack, 2011; Graceffa et al., 2013).

Following the lead of cryo crystallography and cryo electron microscopy, cryoSAXS is being developed as a method to prevent radiation damage in SAXS (Meisburger et al., 2013; Hopkins et al., 2015). CryoSAXS samples are prepared by rapidly cooling small volumes of macromolecule solution with added chemical cryoprotectants. The cryoprotectants and rapid cooling are required to prevent crystalline ice formation, instead obtaining a vitrified sample with repeatable scattering. The scattering from the samples is then measured at cryogenic temperatures. Initial rough measurements suggested that the radiation tolerance of the sample is

increased by two to five orders of magnitude relative to room temperature (Meisburger et al., 2013).

It is necessary to put tighter bounds on this gain from cryocooling for particular macromolecules. This will allow realistic estimation of the reduction in sample volume possible from cryoSAXS, as well as the possible increases in signal to noise from longer exposures on samples of a given volume. An increased understanding of the mechanisms of damage in cryoSAXS may also allow prediction of damage rates for macromolecules at 100 K, further simplifying experiments. Properly quantifying damage requires: measuring changes in SAXS profiles, and deciding how much change is tolerable.

Here we quantify radiation-induced changes to SAXS profiles of glucose isomerase at 100 K, and compare with changes observed at room T. We used the sample holders and modified cryoSAXS methodology developed by (Hopkins et al., 2015). In these sample holders, the sample fractured when cooled to 100 K, which led to anisotropic scatter, and that is dealt with as discussed in (Hopkins et al., 2015). We found that, unlike in standard room temperature SAXS, most or all of the radiation induced change at 100 K occurs to the cryoprotectant-containing sample buffer, rather than to the macromolecule. We measured a gain upon cryocooling of ~3500x for glucose isomerase at 100 K compared to the same sample at room temperature (i.e. both in cryoprotected buffers). Compared to glucose isomerase in a buffer without cryoprotectant at room temperature, the radiation tolerance of the cryoprotected and cryocooled samples was ~60x larger. We measured a decrease in radiation tolerance of room temperature glucose isomerase samples when the cryoprotectant is added to the buffer, which accounts for the radical differences in the ratio of 100 K to room temperature radiation tolerances noted above. We also

observed that the damage rate at 100 K depends upon the buffer constituents. We finish by discussing possible sources for the damage at 100 K.

4.3 Methods

4.3.1 Sample preparation

Radiation damage was evaluated by examining SAXS scattering profiles of glucose isomerase from *Streptomyces rubiginosus* (Hampton Research, Aliso Viejo, CA, HR7-100). The protein was buffer exchanged in a spin column (Amicon 30 kDa MW cut-off, EMD Millipore, Billerica, MA) into 100mM Tris pH 8.0 and 1 mM magnesium chloride. The buffer used for cryoSAXS was 36% w/w propylene glycol (PG), 64% w/w the protein buffer given above. To prepare the protein and matching buffer solutions with 36% w/w PG, a solution was made with 730 mg/mL of PG and the same molality of other components as in the non PG-containing buffer. This PG solution was combined with non PG-containing protein and buffer solutions in a 1:1 volume ratio using positive displacement pipettes, creating samples with a 36% w/w PG concentration (Meisburger et al., 2013; Hopkins et al., 2015). Room temperature samples were created in the same way, with a 1:1 combination with the buffer, to minimize any difference in protein concentration the mixing might create between the two solutions. Prior to 1:1 mixing, protein concentration was measured using a Cary 50 UV-Vis Spectrophotometer (Agilent, Santa Clara, California). The final concentration of glucose isomerase after 1:1 mixing was assumed to be 2.57 mg/mL, half of the measured pre-mixing concentration. We also measured radiation induced changes on an additional buffer: 40 mM sodium acetate pH 4.0, 150 mM sodium chloride with either 36% or 50% w/w PG. This buffer was chosen as it was on hand, and is

different from the glucose isomerase buffer in pH, buffering agent, salt type, and salt concentration. This allowed us to check if buffer composition was relevant to damage at 100 K.

In order to avoid confusion we will use the following terms: protein/protein samples/protein solutions refers to protein in buffer (either at 100 K or room temperature), protein-free buffer refers to buffer without any protein in it. The different buffers will be referred to as: buffer 1 (the buffer used for glucose isomerase: 100 mM Tris pH 8.0 and 1 mM magnesium chloride), or buffer 2 (40 mM sodium acetate pH 4.0 and 150 mM sodium chloride). Buffers without cryoprotectant will be referred to as PG-free buffers, buffers with cryoprotectant will be referred to as 36% w/w or 50% w/w PG buffers as appropriate.

4.3.2 Beamline setup and data collection

SAXS measurements were carried out at the BioCAT beamline (18-ID-D) at the Advanced Photon Source (Fischetti et al., 2004), chosen for its relatively high flux density. For these measurements the energy (wavelength) was 12 keV (1.033 Å). The sample-to-detector distance was 2506 mm, found using silver behenate powder ($d = 58.38$ Å (Huang et al., 1993)). 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 0.0079 Å⁻¹ to 0.3526 Å⁻¹ (range in figures is slightly reduced by binning). Data was collected on a Pilatus3 1M detector (Dectris, Baden, Switzerland). Room temperature exposures were collected in a shutterless mode with individual frame exposure times from 0.02 to 0.05 s and total exposure times per sample up to 8.2 s. 100 K data was collected in a shuttered data collection mode with individual frame exposure times of 2 s and total sample exposure times up to 360 s. The overhead for shuttering and readout was 1 s, so 360 s of exposure took 540 s of experiment time. An active beamstop was used for normalization and an upstream

nitrogen ion chamber monitored the incident flux. The delay between the start of framing in the detector and the opening of the x-ray shutter was measured to be 34 ± 2 ms, and was compensated for in dose calculations. In order to accommodate our small sample holders, the incident beam was collimated, so the flux and beam size values reported here differ from the standard values at BioCAT.

The beamline was set up with a ~ 20 cm air gap at the sample position. For room temperature measurements, acrylic sample holders with polystyrene windows, described in (Hopkins et al., 2015), were used. These sample holders had a 1.8 mm path length, a total sample volume of ~ 3 μL , and were attached to a kinematic mount that ensured reproducible sample positioning when the holder was removed and then remounted. Fresh sample holders were first centered in the X-ray beam, the holder was then removed, loaded, replaced in the beam, exposed, removed from the beam, and then cleaned before loading the next sample (Hopkins et al., 2015). Protein-free buffer and protein were measured in the same sample holder, and the scattering from the empty holder was measured periodically to check for damage or protein adsorption on its windows. When a change in scattering from the empty holder was detected, the sample holder was changed out for a new one. In order to test sample-to-sample variation of damage rates, at least three nominally identical samples of each of the protein, protein-free PG-free buffers, protein-free 36% w/w PG buffers, and protein-free 50% w/w PG buffers were measured. Due to an oversight, only two nominally identical samples of the PG-free and 36% w/w PG buffer 2 were measured at room temperature.

For 100 K measurements, the silicon sample holders, sample loading, and kinematic mounting scheme described in (Hopkins et al., 2015) was used. The samples were cooled using a 100 K nitrogen gas cryostream (700 series, Oxford Cryosystems, Oxford, United Kingdom) placed at

the sample position, equipped with an air-blade shutter. The sample holders were open-topped silicon parallelepiped boxes, with ~ 27 μm thick smooth silicon X-ray windows (Hopkins et al., 2015). The path length in these sample holders was 1 mm and the total volume was ~ 800 nL. First, a sample holder would be scanned into the X-ray beam. The sample holder would then be removed and filled with a precise sample volume using a syringe pump and high precision syringe. The cryostream would be shuttered, then the sample holder was replaced in the X-ray beam. The cryostream was then unshuttered, giving consistent *in situ* cooling. After data collection, the sample holder was removed from the beam, cleaned, and blown dry. In order to test sample-to-sample variation of damage rates, at least three identically prepared samples of each of the protein, protein-free 36% w/w PG buffers, and protein-free 50% w/w protein-free buffer were measured.

In order to accurately calculate absorbed dose, the flux, sample transmission, and beam size/shape at the sample position were measured. The flux was measured in the upstream nitrogen ion chamber, and the transmission of the windows between the chamber and the sample position was accounted for, as was the attenuation of the air gap (~ 16.5 cm of air upstream of the sample). The transmission of the sample cell window material was also corrected for: for the cell used at room temperature each window was 25 μm polystyrene, 0.997 transmission; for the 100 K cell each window was 27 μm silicon, 0.889 transmission. At the time of the measurement, the incident flux on the sample was $3.30 \cdot 10^{12}$ ph/s (accounting for all attenuation except the sample holder). This provided a calibration value for measurements taken by the upstream ion chamber and the beamstop, allowing adjustment for changing ring currents (and any other effects that altered the incident flux. Sample transmission was calculated using the ratio of the beamstop counter for empty and full sample holders. 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 4.1 and 4.2. The FWHM of the beam was 142 (H) x 170 (V) μm^2 .

4.3.3 Calculating dose

It is best practice in both SAXS and crystallography to measure radiation damage as a function of the dose delivered to the sample (Holton, 2009; Jeffries et al., 2015; Hopkins & Thorne, 2016). Dose is the energy absorbed per unit mass, and has units of Gray (Gy), where 1 Gy = 1 J/kg. The necessary details for calculating an accurate dose in SAXS are provided in (Hopkins & Thorne, 2016). This procedure required knowledge of the solution densities, along with other variables already described in this paper. At room temperature, it was assumed that the solution density was 1 g/mL. At 100 K, solution density was measured as described in (Shen et al., 2016) and a density of 1.091 g/mL was found for a 36% w/w PG 64% w/w water solution. For the 50% w/w PG solutions, this density was scaled by a factor of 1.01, the ratio of the room temperature solution densities (Khattab et al., 2012). The dose calculations gave dose rates of roughly 29 kGy/s at 100 K and 32 kGy/s at room temperature, though the exact rate depended on the particular sample and ring current. Over the course of 12 hours, the ring current decreased by ~20%, causing a corresponding decrease in dose rate. Sample based variation was a few percent due to changes in the absorption coefficient with sample and buffer components.

4.3.4 Data processing

Initial data processing was done using python scripts provided by the beamline, Albula (Dectris), and PRIMUS (Konarev et al., 2003). Further data processing was carried out in the lab using a custom version of BioXTAS RAW (Nielsen et al., 2009), which was modified in house to read in

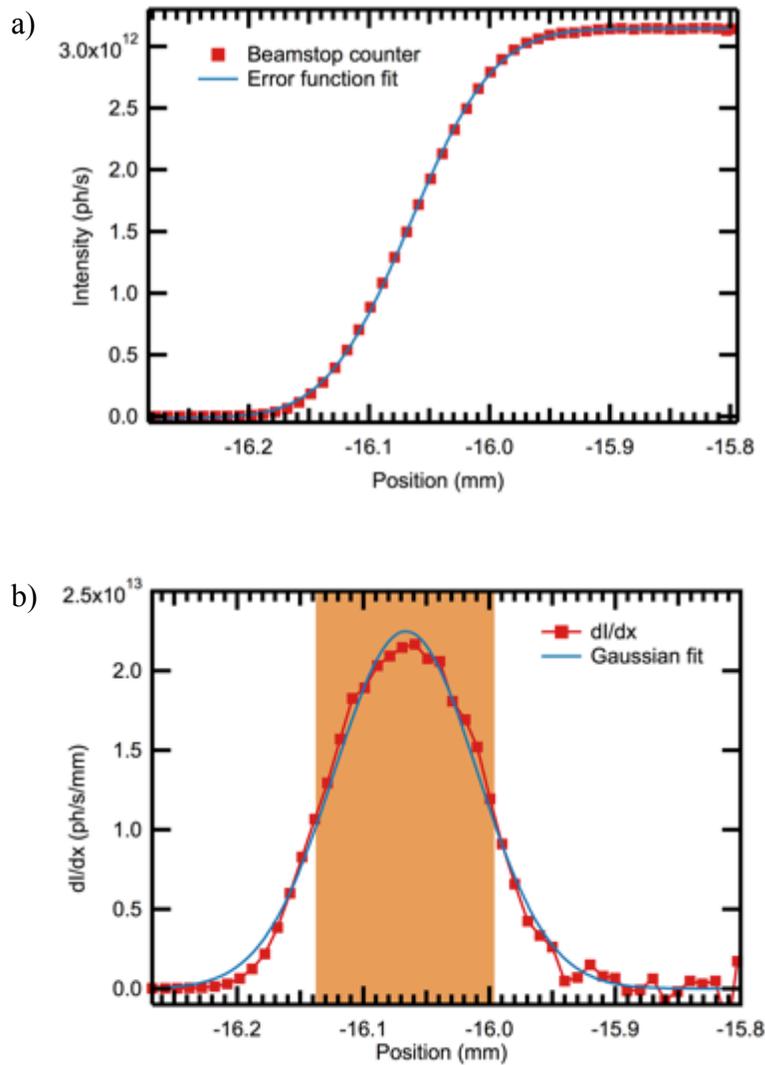


Figure 4.1 – a) Scan showing measured intensity vs. position of a knife-edge being scanned through the beam in the x direction. If the beam is Gaussian, the functional form should be an error function. b) Scan showing the derivative of the measured intensity vs. position. A Gaussian fit is also shown, and the orange area highlights the actual FWHM region. The measured FWHM was $142 \mu\text{m}$.

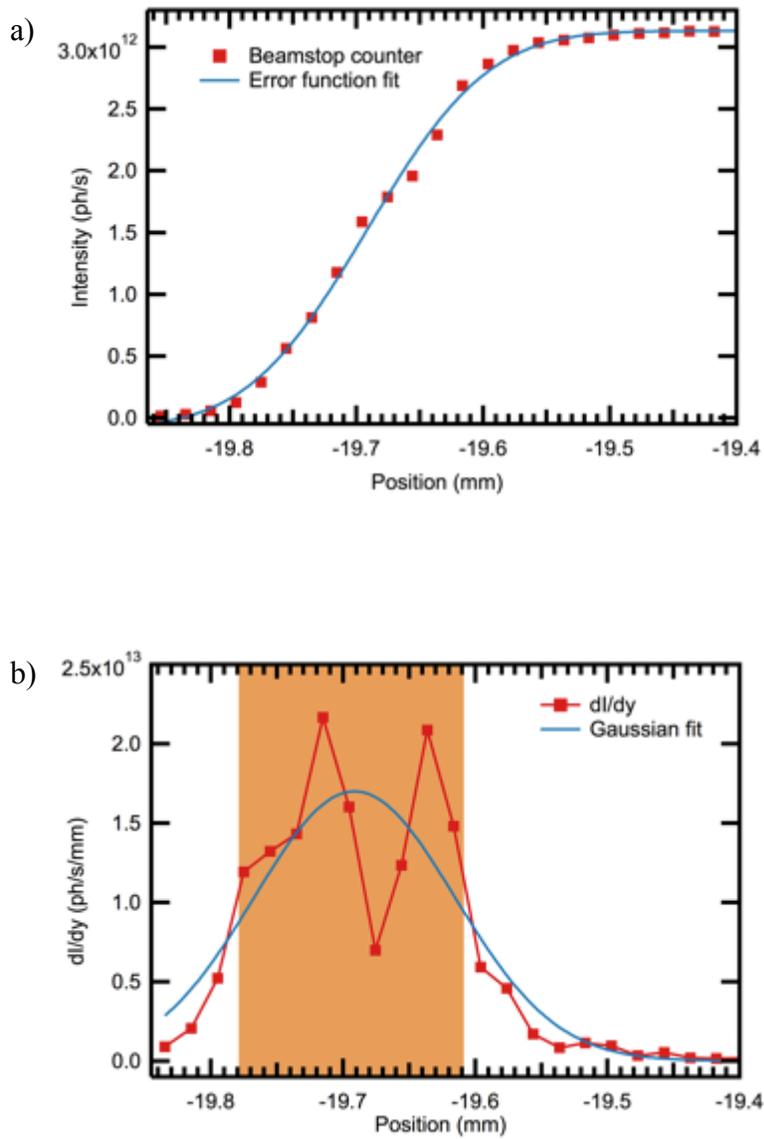


Figure 4.2 – a) Scan showing measured intensity vs. position of a knife-edge being scanned through the beam in the y direction. If the beam is Gaussian, the functional form should be an error function. b) Scan showing the derivative of the measured intensity vs. position. A Gaussian fit is also shown, and the orange area highlights the actual FWHM region. The measured FWHM was 170 μm .

the data from BioCAT. As discussed in (Hopkins et al., 2015), at 100 K some images displayed anisotropic scattering from fractures in the sample. This was masked out prior to integration, to minimize their effect on the scattering profile.

Radiation damage is typically quantified by looking at the change in scattering profile derived parameter vs. dose (Kuwamoto et al., 2004; Jeffries et al., 2015; Hopkins & Thorne, 2016). We followed the protocol in (Hopkins & Thorne, 2016), and the processing provided all of the parameters described there. Following their work, we report the metrics derived from the radius of gyration, the molecular weight, and the absolute integrated intensity. Radiation damage processing was done using the custom python program described in (Hopkins & Thorne, 2016), which utilizes: the ATSAS package (Petoukhov et al., 2007, 2012), the pseudo R_g and initial damage rate definitions from (Jeffries et al., 2015), and the methods for calculating molecular weight described in (Fischer et al., 2009; Rambo & Tainer, 2013). Sensitivities were calculated from the parameters as detailed in (Hopkins & Thorne, 2016). At 100 K, some dose curves (normalized parameter value vs. dose) became non-linear, in these cases only the initial linear region was fit (typically the first 2-4 MGy). At room temperature, the protein samples and cryoprotectant buffer showed a delay to the onset of radiation-induced changes in the absolute intensity metric. Fits were done after this initial delay. In some cases, noted where appropriate, both positive and negative sensitivities were measured from identically prepared samples. For purposes of reporting, the average of the absolute sensitivities is given.

Two additional processing methods were applied to the scattering profiles. First, to accurately compare the change in integrated intensity between protein-free buffer curves at different temperatures (and in different sample holders), the contribution to the scattered intensity from

the background (instrument and sample holder) were removed, this is indicated as $I_B(q)$, Equation (4.10). Second, the initial contribution of the protein had to be removed from the protein scattering, allowing comparison between the rates of change of the integrated intensity of buffer and protein solutions at the same temperature, indicated as $I_{nom,B}$, Equation (4.12). Further details on this are given in Section 4.7.1.

4.4 Results

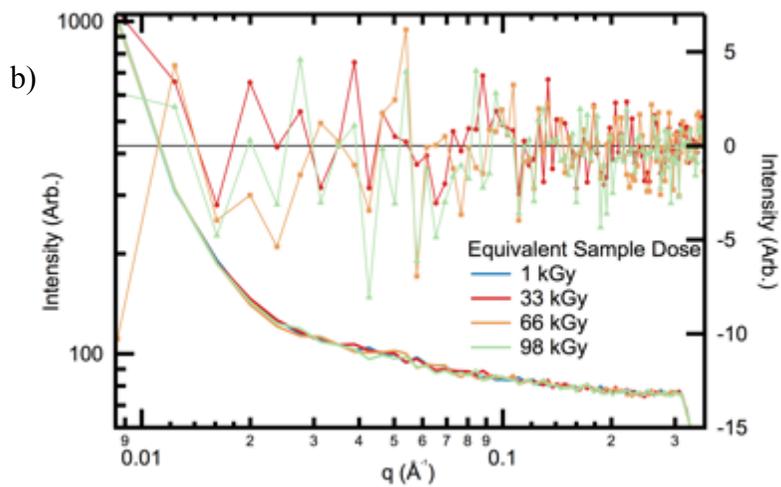
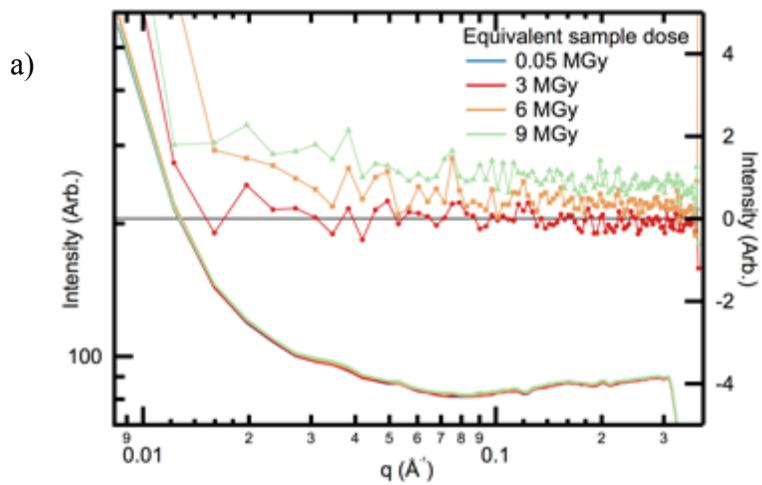
4.4.1 Time or dose dependent changes to background scattering

Because we anticipated collecting data to large doses and for long times, we needed to quantify time or dose dependent changes unrelated to sample damage. To do so, exposures of equal time to the sample exposures at each temperature were taken of empty sample holders. At 100 K, the silicon sample holder, some change was observed, Figure 4.3a. This could be due to damage to the sample holder (though it is not clear what would be damaging), gradual frost accumulation due to slight misalignment of the sample holder in the cryostream, or drift in the beamline. The integrated intensity sensitivity was calculated, using the dose that would have been absorbed by the sample at a given exposure time (this is different from the dose absorbed by the sample holder), and found to be an order of magnitude slower than the slowest measured sample damage rates, so this change was ignored. At room temperature, the polystyrene sample holder, Figure 4.3b, no changes were seen in the scattering profile (note that the total exposure time is much lower at room temperature).

4.4.2 Damage at 100 K

There are several distinct qualitative features to the damage observed at 100 K. Figure 4.4 shows scattering profiles as a function of dose for glucose isomerase. There was a clear increase with

Figure 4.3 – a) This shows scattering profiles from an empty sample holder at 100 K, the lower curves, plotted on left axis. These were taken at the same set of exposure times as the sample, and the legend lists the equivalent sample dose for a given exposure. The actually dose for the sample holder was not calculated. The upper curves, plotted on the right axis, show the difference between the initial scattering profile and the others shown. There is a slow upward trend visible in the exposures, along with some larger variation at low q . The rate of this change was, on average, at least an order of magnitude slower than any other damage rates reported. The source of the change is not clear, as silicon is not expected to damage. There may be some residue left from the fabrication process that is damaging. It is also possible that this is not damage. This could be slow frost build-up on the sample holder due to poor centering of the sample holder in the cryostream, though ice typically has a power law scattering profile. It may also be some slow change in the beamline. The total exposure time (including dead time between images) was ~ 540 s. b) This shows scattering profiles from an empty sample holder at room temperature, the lower curves, plotted on left axis. These were taken at the same set of exposure times as the room temperature sample, and the legend lists the equivalent sample dose for a given exposure. The actually dose for the sample holder was not calculated. The upper curves, plotted on the right axis, show the difference between the initial scattering profile and the others shown. As can be seen, there is no visible change in the room temperature sample holders. The total exposure time was ~ 8 s, so if the changes in a) were due to some kind of slow drift in beamline components it would not be expected to affect these images.



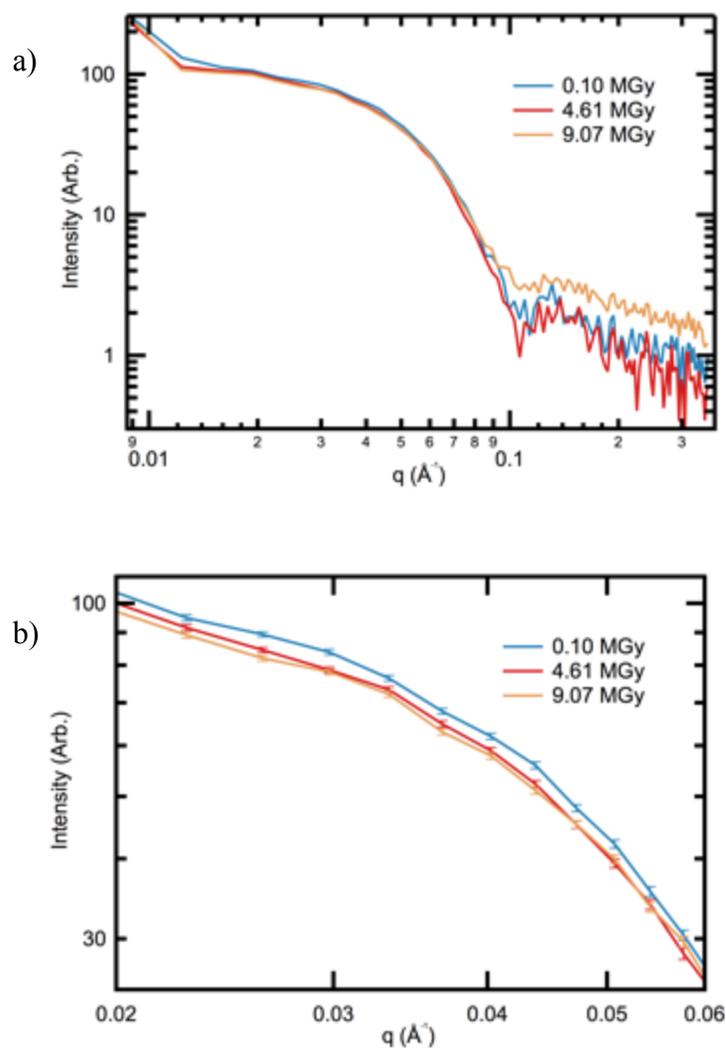


Figure 4.4 – a) Scattering profiles for glucose isomerase in 36% w/w PG buffer 1 at 100 K as a function of dose. An increase at high q with dose is clearly visible. The uptick at the lowest q is scattering from fractures that could not be removed by masking. b) The same scattering profiles as in part a, but from $q = 0.02$ to 0.06 \AA^{-1} . The restricted q range allows clear observation of a decrease in the intensity at low q with dose.

dose at high q , Figure 4.4a, while at low q there was a decrease, Figure 4.4b. Figure 4.5 shows that radiation induced changes occurred in the protein-free buffers as well as the protein samples. Both the protein-free 36% w/w buffer 1 and protein-free 36% w/w buffer 2 showed a clear decrease at low q , Figure 4.5a, but only the protein-free buffer 2 increased at high q , Figure 4.5b. In addition to the relative minor changes shown in Figures 4.4 and 4.5, in 5 of the 16 measured samples with 36% w/w PG and all 3 of the 50% w/w PG protein-free buffer 2 samples a catastrophic failure mode appeared, Figure 4.6. This was characterized by a dramatic, anisotropic, increase in the low- q scattering that occurs much faster than the changes shown in Figures 4.4 and 4.5. The dose of first observation varied from ~ 3.5 - 9.6 MGy, and the failure typically took ~ 3 - 5 seconds or ~ 100 kGy of dose (it is unclear what the relevant scale is for this change) to appear at full strength. Finally, the radiation-induced changes in the protein-free buffers depended on the buffer constituents, both the cryoprotectant concentration, Figure 4.7, and the other components, Figure 4.5. These observations are distinct from the radiation-induced changes seen for these solutions at room temperature, which will be discussed in the following section. We show additional scattering profiles from nominally identical samples in Figures 4.8-4.10. A sample dose curve for the protein is shown in Figure 4.11.

The rate of radiation damage at 100 K was quantified using the R_g sensitivity, S_{rg} , the molecular weight sensitivity S_{mw} , the integrated intensity sensitivity, S_I , and the associated standard deviations σ_{rg} , σ_{mw} , and σ_I . The reported S_{rg} is for the R_g calculated from the $p(r)$ function, but all methods of determining R_g used in (Hopkins & Thorne, 2016) yielded average sensitivities within 2% of this value, so our choice of a particular one did not significantly influence the results. Both negative and positive values were measured for S_{rg} . As we were

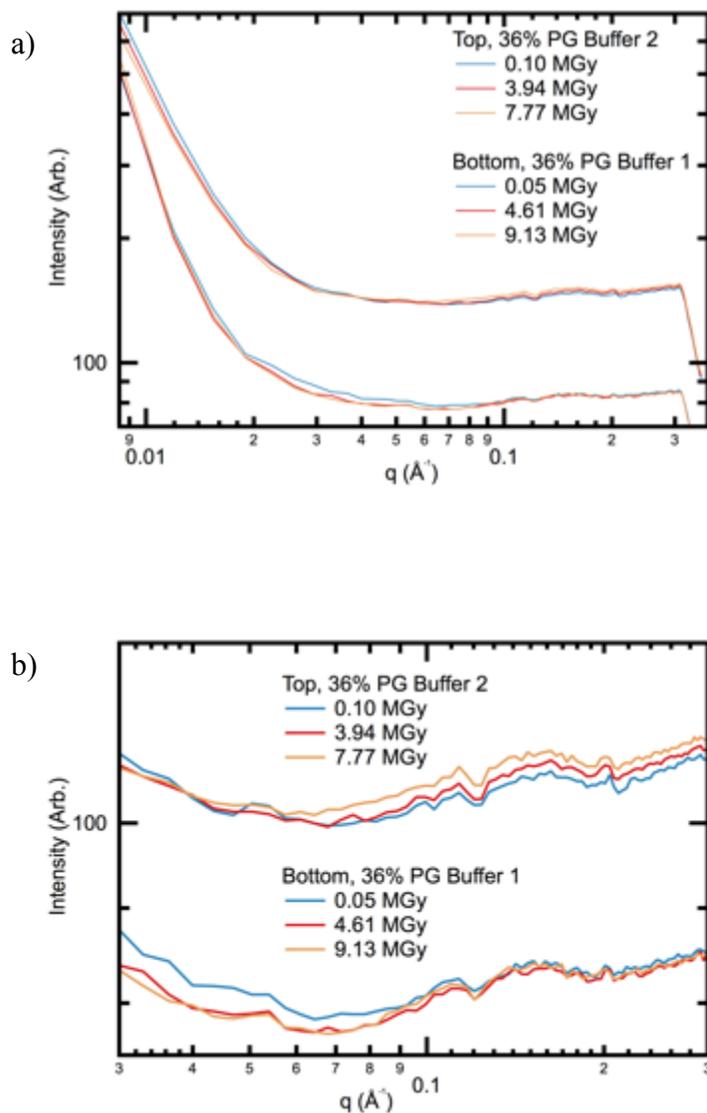


Figure 4.5 – a) Scattering profiles from protein-free 36% w/w PG buffer 1 and buffer 2. Lower scattering profiles are buffer 1, upper scattering profiles are the buffer 2. A decrease at low q is seen for buffer 2, while a decrease at mid q is seen for the buffer 1. The buffer 2 scattering profiles were offset for clarity by multiplication by 1.25. b) The same scattering profiles, from $q = 0.03 - 0.3 \text{ \AA}^{-1}$, to allow easier observation of changes with dose. An increase at high q is seen in buffer 2, which is not present in buffer 1. The offset has been changed between a) and b) to a multiplication of buffer 2 by 0.9.

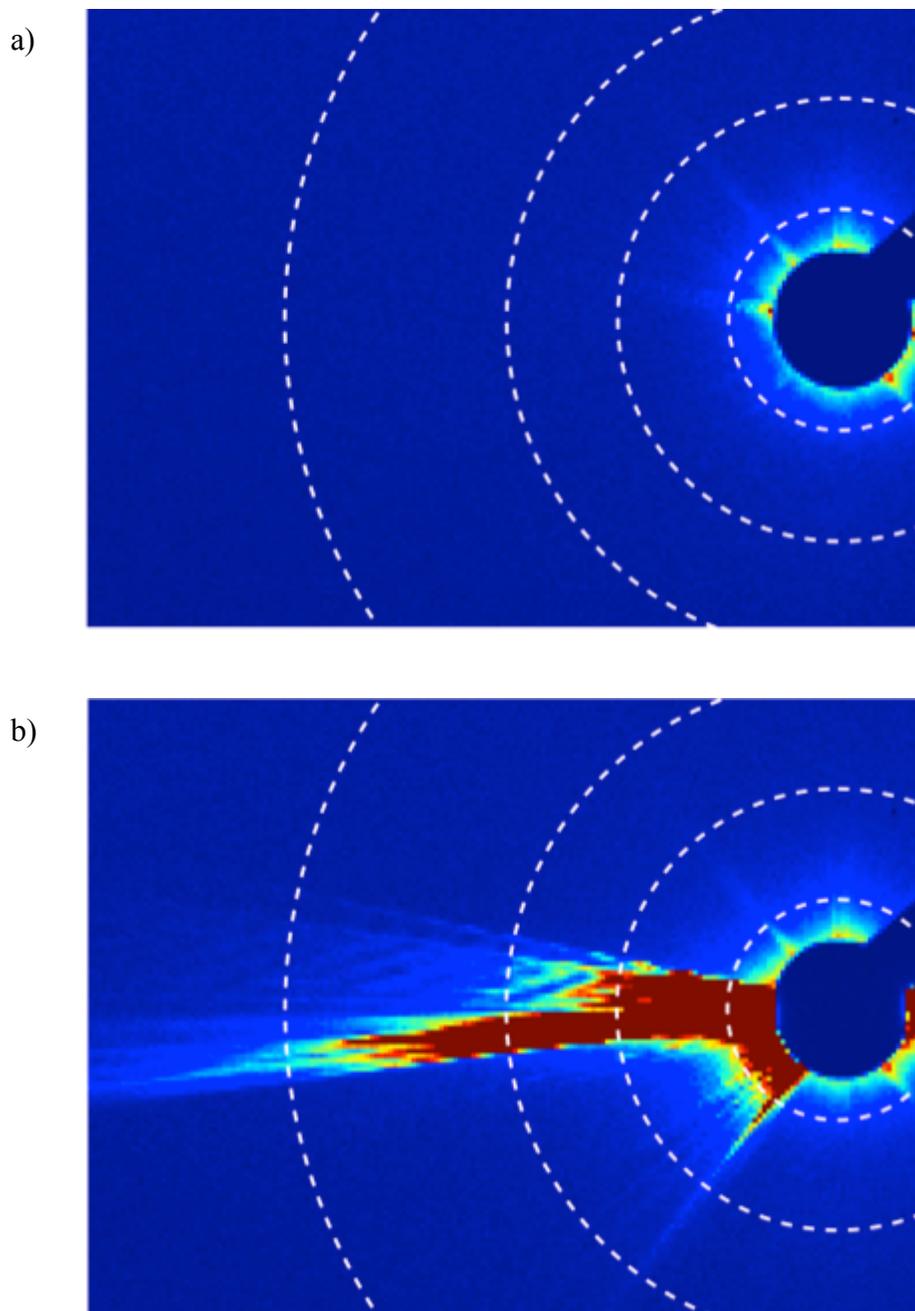


Figure 4.6 – Parts a) and b) show a portion of the detector image near the beamstop. The white dashed lines show resolution rings (in increasing distance from the beamstop) of $q = 0.01$, 0.02 , 0.03 , and 0.05 \AA^{-1} . The intensity scale is the same in both images. These images are from protein-free 50% w/w PG buffer 2, though the same behavior was also observed in 36% w/w PG solutions with and without protein. a) Detector image before catastrophic failure. Some small jets are visible from sample cracking, as discussed in (Hopkins et al., 2015). This image is at a dose of 4.29 MGy, though the failure has been observed at lower and higher doses. b) Detector image post-catastrophic failure. The images were taken 3 seconds, $\sim 100 \text{ kGy}$, apart. This image is at a dose of 4.39 MGy.

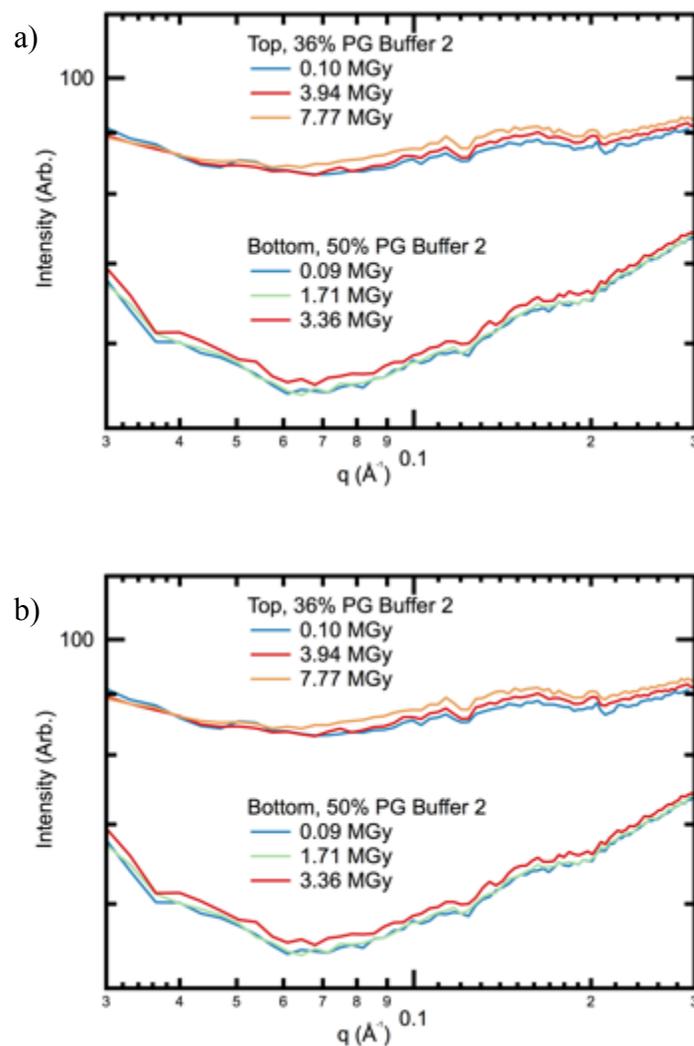


Figure 4.7 – a) Scattering profiles as a function of dose for protein-free 36% and 50% w/w buffer 2. The lower scattering profiles are the 50% w/w PG buffer, upper scattering profiles are the 36% w/w PG buffer. The 36% w/w PG shows the same decrease at low q visible in Figure 4.5. After 3.36 MGy, the 50% w/w PG buffer underwent a catastrophic failure. Those curves are not shown, as they would obscure all other changes. b) The same scattering profiles as a), from $q=0.03-0.3 \text{ \AA}^{-1}$, to allow easier observation of changes with dose. An increase at all q shown is now visible for the 50% w/w PG buffer. The 36% w/w PG were offset for clarity by multiplication by 0.75.

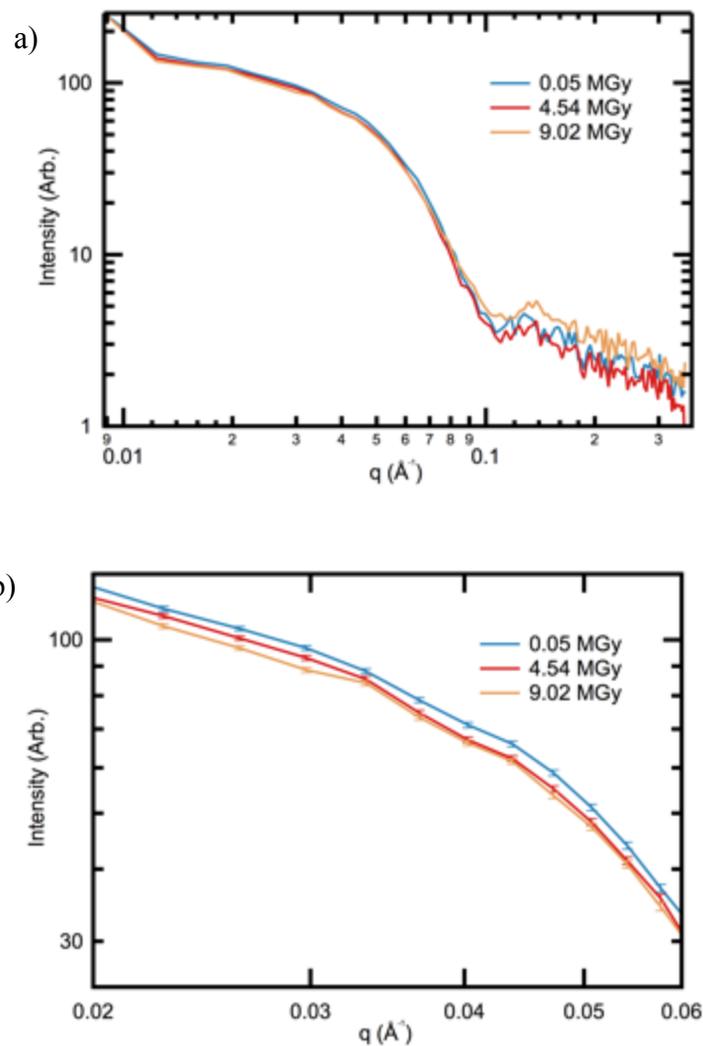


Figure 4.8 – This shows data for a glucose isomerase sample nominally identical to that in Figure 4.4. a) Scattering profiles for glucose isomerase in 36% w/w PG buffer 1 at 100 K as a function of dose. An increase at high q with dose is clearly visible. The uptick at the lowest q is scattering from fractures that could not be removed by masking. b) The same scattering profiles as in a), but from $q = 0.02$ to 0.06\AA^{-1} . The restricted q range allows clear observation of a decrease in the intensity at low q with dose.

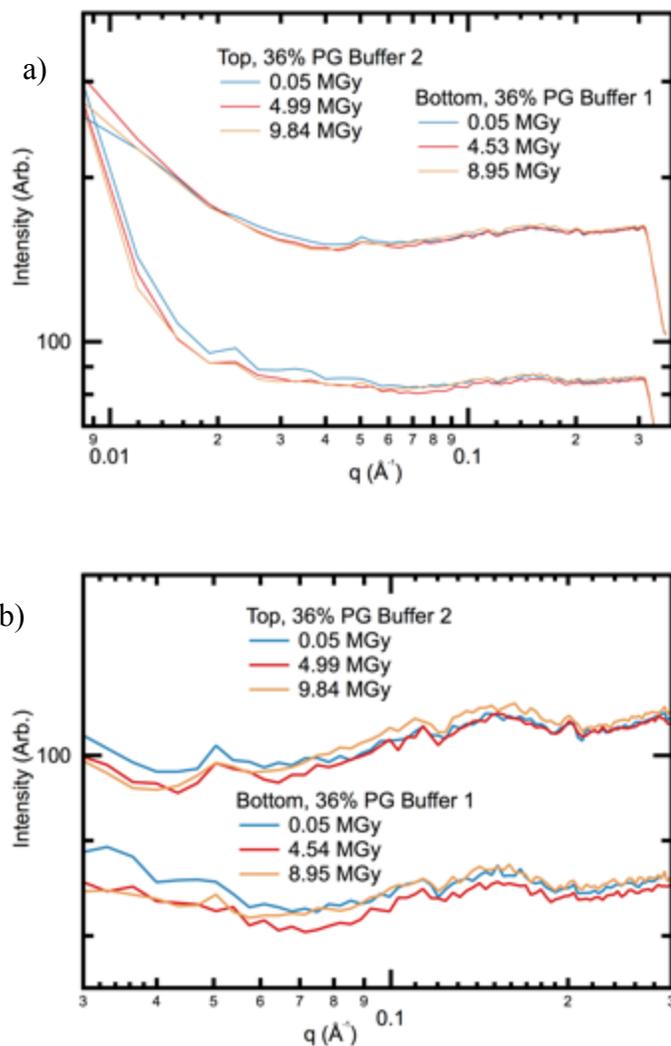


Figure 4.9 – This shows data for buffer samples nominally identical to those in Figure 4.5. a) Scattering profiles from protein-free 36% w/w PG buffer 1 and buffer 2. Lower scattering profiles are buffer 1, upper scattering profiles are the buffer 2. A decrease at low q is seen for buffer 2, while a decrease at mid q is seen for the buffer 1. The buffer 2 scattering profiles were offset for clarity by multiplication by 2. b) The same scattering profiles, from $q = 0.03$ - 0.3 \AA^{-1} , to allow easier observation of changes with dose. An increase at high q is seen in buffer 2, which is not present in buffer 1. The offset has been changed between a) and b) to a multiplication of buffer 2 by 1.3.

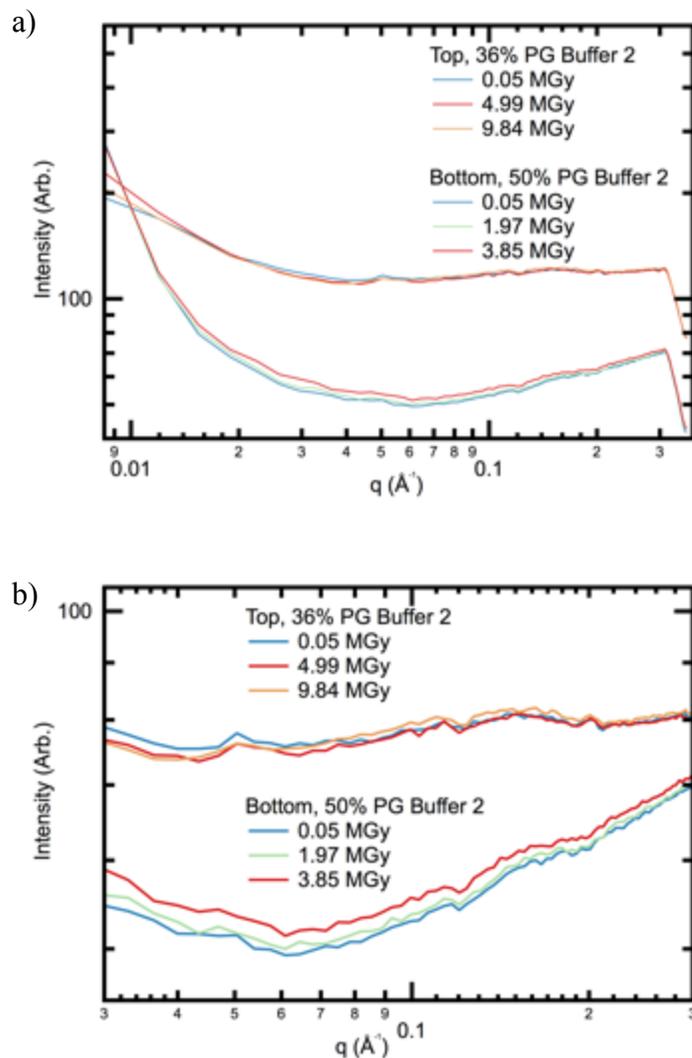


Figure 4.10 – This shows data for buffer samples nominally identical to those in Figure 4.7. a) Scattering profiles as a function of dose for protein-free 36% and 50% w/w buffer 2. The lower scattering profiles are the 50% w/w PG buffer, upper scattering profiles are the 36% w/w PG buffer. The 36% w/w PG shows the same decrease at low q visible in Figure 4.5. After 3.36 MGy, the 50% w/w PG buffer underwent a catastrophic failure. Those curves are not shown, as they would obscure all other changes. The 36% w/w PG buffer 2 scattering profiles were offset for clarity by multiplication by 1.5. b) The same scattering profiles as a), from $q = 0.03 - 0.3 \text{ \AA}^{-1}$, to allow easier observation of changes with dose. An increase at all q shown is now visible for the 50% w/w PG buffer. Unlike in a), there is no offset applied to these profiles.

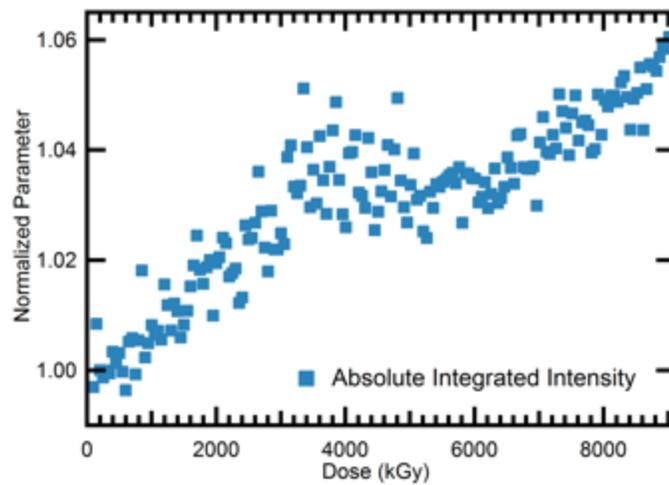


Figure 4.11 – An example dose curve from glucose isomerase in the 36% w/w PG buffer 1 at 100 K. This shows the normalized absolute integrated intensity vs. dose. Each point represents the value from a single scattering profile.

interested in how quickly the radius of gyration changes with dose, the average and standard deviation was calculated from the absolute values of the measured sensitivities. This is reported in Table 4.1. The initial R_g for glucose isomerase ranged from 33-36 Å (with uncertainties of 5 Å or more from AUTORG), in good agreement with the previously reported value in 36% w/w PG of 33.3 Å (Hopkins et al., 2015).

(Hopkins & Thorne, 2016) found that molecular weight was one of the few parameters for glucose isomerase that showed any significant change with dose at room temperature. We calculated S_{mw} , and found that, as for the radius of gyration, not all of the sensitivities have the same sign. Again, the reported value is the average and standard deviation of the absolute values, and is given in Table 4.1. The initial calculated molecular weight ranged from 112 to 159 kDa, in poor agreement with the literature value of 172 kDa. As the radius of gyration agreed with expected values, and the initial scattering profiles at 100 K and room temperature agree, it is more likely the automated determination of molecular weight was inaccurate than that the molecular weight changed upon cooling. This failure may come in part from the change in contrast from the cryoprotectant and cryocooling, as the estimation method depends on $I(0)$, and so must be sensitive to contrast (Rambo & Tainer, 2013).

In order to compare the rate of change in protein-free buffer to protein at 100 K, data was processed as described in Section 4.3.4. We report the average integrated intensity sensitivity, S_I , and standard deviations, σ_I for $I_B(q)$ (protein-free buffer) and $I_{nom,B}(q)$ (protein) in Table 4.2. In this case, all of the measured sensitivities had the same sign. The measured sensitivities show that the glucose isomerase solutions changed at a rate similar to the protein-free buffer 1.

Table 4.1 – Comparison between S_{rg} and S_{mw} for glucose isomerase at 100 K and at room temperature in the PG-free and 36% w/w PG buffer 1. The sensitivity ratios are given for the room temperature measurements to the 100 K measurement. For S_{rg} , the same sample at 100 K is 4370 times less sensitive than at room temperature. For S_{mw} the room temperature to 100 K ratio is 650. Results are also reported for the analysis of radiation damage data from (Meisburger et al., 2013), glucose isomerase in 45% w/w PEG 200 at 100 K.

Sample	S_{rg}	σ_{rg}	S_{mw}	σ_{mw}	S_{rg}	S_{mw}
	(%/kGy)	(%/kGy)	(%/kGy)	(%/kGy)	ratio:	ratio:
					RT/100	RT/100
					K	K
G.I. (100 K), 36% PG	$1.5 \cdot 10^{-4}$	$0.6 \cdot 10^{-4}$	$1.7 \cdot 10^{-3}$	$0.7 \cdot 10^{-3}$	-	-
G.I., PG-free (RT)	$1.1 \cdot 10^{-2}$	$0.5 \cdot 10^{-2}$	-0.11	0.05	74	-65
G.I., 36% PG (RT)	$6.6 \cdot 10^{-1}$	$1.5 \cdot 10^{-1}$	1.1	1.3	4400	650
G.I., 45% PEG 200 (100 K)	$8.6 \cdot 10^{-4}$	-	$-6.3 \cdot 10^{-3}$	-	-	-

Table 4.2 - Radiation induced changes to glucose isomerase and protein-free buffers at 100 K as measured by S_I . S_I and standard deviation σ_I are for $I_{nom,B}(q)$ (protein) and $I_B(q)$ (protein-free buffer). The protein is glucose isomerase (G.I.) in the 36% w/w PG buffer 1. The buffers are the protein-free 36% w/w PG buffer 1 and buffer 2 and protein-free 50% w/w PG buffer 2. For glucose isomerase, the sensitivity ratio of protein to protein-free buffer 1 is 0.7. S_I for buffer 1 has a large standard deviation, so we cannot reliably state that the protein sample damages more slowly than the protein-free buffer.

Sample	S_I (%/kGy)	σ_I (%/kGy)
G.I., 36% PG	$1.3 \cdot 10^{-4}$	$0.6 \cdot 10^{-4}$
Buffer 1, 36% PG	$2.0 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$
Buffer 2, 36% PG	$2.6 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$

The ratio of S_I for protein and protein-free buffer 1 is 0.7, but S_I for the protein-free buffer 1 has a large standard deviation, so we cannot reliably state whether the protein-free buffer 1 and protein samples change at the same or different rates. Examining the sensitivities of the protein-free 36% w/w PG buffer 1 and buffer 2 showed that the non-cryoprotectant solution components may influence the rate of damage. There is a measured 30% difference in S_I , but again the standard deviations are large. However, there is a stronger effect of changing cryoprotectant concentration, as demonstrated by the protein-free 50% w/w PG buffer 2, which was twice as sensitive as the protein-free 36% w/w PG buffer 2.

In order to enable additional comparison with room temperature, S_I was calculated for protein solution using the standard subtracted intensity, I_{sub} (as in Equation (4.11)), rather than for $I_{nom,B}$, and is reported in Table 4.3.

4.4.3 Damage at room temperature

Figure 4.12 shows room temperature scattering profiles vs. dose for glucose isomerase in the PG-free and 36% w/w PG buffer 1. The scattering profiles for protein in the PG-free buffer 1 slowly decreased in intensity at low q , similar to what was seen by (Hopkins & Thorne, 2016). The protein in the 36% w/w PG buffer 1 had a sharp uptick in the scattering profiles at low q , at much lower doses than the protein in the PG-free buffer 1. We applied the same metrics for damage at room temperature as we did at 100 K. Calculated S_{rg} values are listed in Table 4.1, and show that the addition of 36% w/w PG to buffer 1 increased S_{rg} sensitivity by a factor of 59. For glucose isomerase in 36% PG, initial R_g ranged from 33-34 Å (uncertainties were 5 Å or

Table 4.3 – Comparison of the integrated intensity sensitivity, S_I , for glucose isomerase at 100 K and room temperature using the standard subtracted intensity, I_{sub} . Note that due to the different processing method these values should not be comparable to those listed in Tables 4.1 and 2. The sensitivity ratio between the same sample at 100 K and room temperature is 5500, larger than to the value found using S_{rg} (Table 4.1). Results are also reported for the analysis of radiation damage data from (Meisburger et al., 2013), glucose isomerase in 45% w/w PEG 200 at 100 K.

Sample	S_I (%/kGy)	σ_I (%/kGy)	S_I ratio: RT/100 K
G.I. (100 K), 36% PG	$1.30 \cdot 10^{-3}$	$0.04 \cdot 10^{-3}$	-
G.I., PG-free (RT)	$4.3 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$	33
G.I., 36% PG (RT)	7.1	1.2	5500
G.I., 45% PEG 200 (100 K)	$4.3 \cdot 10^{-3}$	-	-

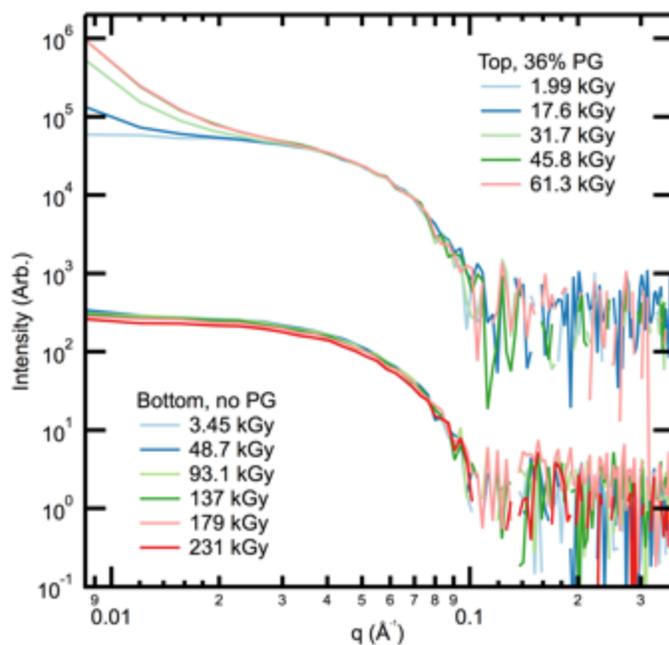


Figure 4.12 – Room temperature scattering profiles glucose isomerase as a function of dose. The bottom set is in the PG-free buffer 1, while the top set is in the 36% w/w PG buffer 1. Glucose isomerase in the PG-free buffer 1 show a graduate decrease in intensity at low and mid q , while in the 36% w/w PG buffer 1 it shows a sharp increase in intensity at low q at much lower doses. The 36% w/w PG profiles were offset for clarity by multiplication by 250.

larger), in good agreement with the previously reported value of 32.7 Å (Hopkins et al., 2015). In the PG-free buffer 1, initial R_g took values of 33-34 Å (5 Å or greater uncertainties), also in good agreement with the expected 32.7 (Kozak, 2005).

The molecular weight sensitivity was also calculated, and is listed in Table 4.1. The addition of 36% w/w PG to buffer 1 increased S_{mw} by a factor of 10. The initial calculated molecular weight in the PG-free buffer 1 ranged from 143 to 198 kDa, while in the 36% w/w PG buffer 1 it ranged from 151 to 179 kDa. Both of these are in reasonable agreement with the literature value of 172 kDa, but show some scatter.

Figure 4.13 shows the change in the protein-free buffers with absorbed dose. There was little or no visible change for the protein-free PG-free buffer 1, but the protein-free 36% w/w PG buffer 1 scattering profiles had a low q uptick similar to that observed for the protein solution. The rate of change in the buffers was quantified as described above. We report the average integrated intensity sensitivity and standard deviations for $I_B(q)$ (protein-free buffer) and $I_{nom,B}(q)$ (protein) in Table 4.4. These show that glucose isomerase in 36% w/w PG buffer 1 damaged more quickly than the protein-free buffer 1 changed by itself, indicating that there was some actual damage to the protein, in addition to the changes to the buffer. This was also true for glucose isomerase in the PG-free buffer 1, as expected.

As noted previously, a delay in the onset of radiation-induced changes was seen in the integrated intensity for protein samples (both with and without cryoprotectant), and for the cryoprotectant buffer samples. The average delay for each condition is reported in Table 4.5, and an example dose curve is shown in Figure 4.14. Though the standard deviations are relatively large, the delay

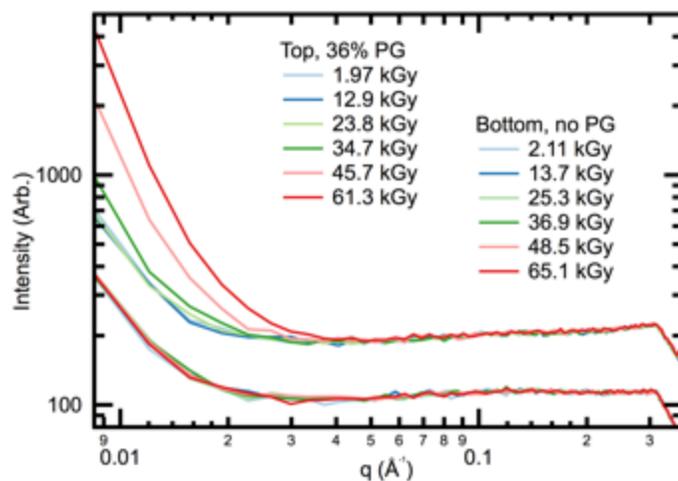


Figure 4.13 – Comparison between protein-free buffers at room temperature. The top scattering profiles are for 36% w/w PG buffer 1 while the bottom scattering profiles are for PG-free buffer 1. There is no visible change to the PG-free buffer 1, while the 36% w/w PG buffer 1 shows an increase at low q similar to that observed for the protein (Figure 4.12). The 36% w/w PG buffer 1 profiles were offset for clarity by multiplication by 2.

Table 4.4 – Radiation induced changes to glucose isomerase and protein-free buffers at room temperature as measured by S_I . S_I and standard deviation σ_I are for $I_{nom,B}(q)$ (protein) and $I_B(q)$ (protein-free buffer). The protein is glucose isomerase in the PG-free and 36% w/w PG buffer 1. The buffers are the protein-free PG-free and 36% w/w PG buffer 1 and the protein-free PG-free, 36% w/w PG and 50% w/w PG buffer 2. The ratio of S_I for protein to protein-free buffer 1 is 3.4 for glucose isomerase in the PG-free buffer 1 and 1.7 for glucose isomerase in the 36% w/w PG buffer 1.

Sample	S_I (%/kGy)	σ_I (%/kGy)
G.I., PG-free	$1.1 \cdot 10^{-2}$	$0.5 \cdot 10^{-2}$
G.I., 36% PG	1.62	0.26
Buffer 1, PG-free	$3.2 \cdot 10^{-3}$	$4.5 \cdot 10^{-3}$
Buffer 1, 36% PG	0.92	0.34
Buffer 2, PG-free	$5.6 \cdot 10^{-4}$	$2.5 \cdot 10^{-4}$
Buffer 2, 36% PG	1.24	0.51
Buffer 2, 50% PG	1.61	0.30

Table 4.5 – At room temperature, cryoprotectant containing solutions of protein and protein-free buffer showed delays in the onset of radiation induced changes. This table reports the delays and standard deviation in delays for the integrated intensity, using $I_{nom,B}(q)$ (protein) and $I_B(q)$ (protein-free buffer). The protein is glucose isomerase (G.I.) in the 36% w/w PG buffer 1. The buffers are the protein-free 36% w/w PG buffer 1 and buffer 2 and protein-free 50% w/w PG buffer 2.

Sample	Delay (kGy)	σ_{Delay} (kGy)
G.I., 36% PG (RT)	10.4	3.3
Buffer 1, 36% PG (RT)	17.6	1.6
Buffer 2, 36% PG (RT)	14.0	5.4
Buffer 2, 50% PG (RT)	12.5	6.8

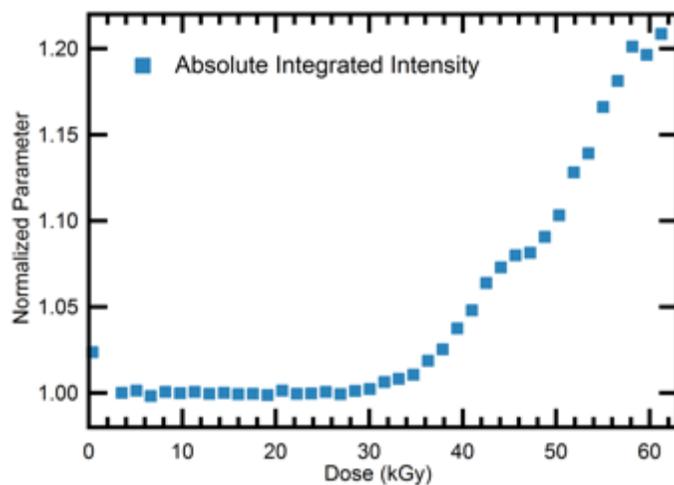


Figure 4.14 – An example dose curve from protein-free 36% w/w PG buffer 1 at room temperature. This shows the normalized absolute integrated intensity vs. dose. Each point represents the value from a single scattering profile.

observed for the protein sample is smaller than that for the protein-free buffer, which agrees with the observation that there is damage to the protein, not just changes to the buffer. Increasing the concentration of cryoprotectant increased the sensitivity of the protein-free buffer 2, and the buffer components affected the rate of change of the buffer with PG. This is consistent with the results at 100 K, though the manifestation of the radiation-induced changes in the scattering profile is quite different. Both protein-free PG-free buffers were relatively insensitive.

In order to enable additional comparison with 100 K, S_I calculated for protein solution using the standard subtracted intensity, I_{sub} , is given in Table 4.3. The addition of 36% w/w PG to buffer 1 increased S_I by 165x.

4.5 Discussion

4.5.1 Comparison between room temperature and 100 K sensitivities

Table 4.1 compares S_{rg} and S_{mw} for glucose isomerase at 100 K, in the 36% w/w PG buffer 1, and at room temperature in both cryoprotectant free and 36% w/w PG buffer. We start with S_{rg} and S_{mw} because change in R_g is the typical metric for radiation damage at room temperature, while S_{mw} was found to be a good metric for glucose isomerase in particular (Hopkins & Thorne, 2016). The direct comparison is for protein in the 36% w/w PG buffer 1 at room temperature and 100 K, which gave a 4400x decrease in S_{rg} and 650x decrease in S_{mw} for the cryocooled protein solution. However, at room temperature glucose isomerase in the 36% w/w PG buffer 1 was more sensitive than glucose in the PG-free buffer 1, so the conservative comparison is between

the protein in the PG-free buffer 1 at room temperature and the 100 K protein in the 36% w/w PG buffer. This showed a 74x decrease in S_{rg} and 65x decrease in S_{mw} due to cryocooling.

It is not clear what would cause a change in either S_{rg} or S_{mw} at 100 K, as the protein is surrounded by solid solvent that will prevent significant changes in size, including the aggregation or fragmentation and dissociation that would lead to changes in molecular weight. This is discussed further in Section 4.5.2. The fact that we measured both positive and negative sensitivities for S_{rg} and S_{mw} for identically prepared samples at 100 K also suggests that these are not good metrics for cryocooled samples. The integrated intensity sensitivity is a more suitable metric, as it will detect any change in the scattering profiles. Table 4.3 compares S_I for the standard subtracted intensity for samples at 100 K and room temperature. The sensitivity decreased on cooling from room temperature to 100 K by a factor of 5500 for glucose isomerase in 36% w/w PG buffer 1, and by a factor of 33 when comparing glucose isomerase in the PG-free buffer 1 at room temperature with the 100 K sample in 36% w/w PG buffer 1. These sensitivity ratios are comparable to those calculated from S_{rg} and S_{mw} .

For glucose isomerase in the 36% w/w PG buffer 1 the ratio of S_I in protein samples to protein-free 36% w/w PG buffer 1 was 0.7 at 100 K and 1.7 at room temperature. As such, the change in sensitivity of the protein-free buffer 1 may also be a good measure of gain from cryocooling for glucose isomerase. Table 4.6 lists S_I for the protein-free buffers at room temperature and 100 K. Upon cryocooling, both of the protein-free 36% w/w PG buffers decreased in sensitivity by roughly a factor of 4700 (4600 for the buffer 1, 4800 for the buffer 2), while the protein-free 50% w/w PG buffer 2 decreased in sensitivity by 3200x. The decrease in sensitivity for the

Table 4.6 – Comparison of protein-free buffer integrated intensity sensitivities, S_I , at 100 K and room temperature (calculated from $I_B(q)$). All of the buffers at room temperature change more quickly than those at 100 K. The ratio of the buffer sensitivity for the protein-free glucose isomerase buffer is 4600, which is 20% lower than that for the protein sample (Table 4.3).

Sample	S_I (%/kGy)	σ_I (%/kGy)	S_I ratio: RT/100 K
Buffer 1, 36%	$2.0 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$	-
PG (100 K)			
Buffer 2, 36%	$2.6 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$	-
PG (100 K)			
Buffer 2, 50%	$5.1 \cdot 10^{-4}$	$1.1 \cdot 10^{-4}$	-
PG (100 K)			
Buffer 1, 36%	0.92	0.34	4600
PG (RT)			
Buffer 2, 36%	1.24	0.51	4800
PG (RT)			
Buffer 2, 50%	1.61	0.30	3200
PG (RT)			

protein-free 36% w/w PG buffers is similar to the decrease in sensitivity upon cryocooling of the protein, 5500x. Regardless of what buffer, sample, or sensitivity you compare, cryocooling significantly increased the radiation tolerance. The minimal decrease in sensitivity (and hence gain in radiation tolerance) upon cryocooling we measured was a factor of 33, while the maximum was a factor of 5500. Glucose isomerase at room temperature in the PG-free buffer 1 is much less radiation sensitive than other standard proteins (Hopkins & Thorne, 2016). Most of the observable changes in the scattering profile at 100 K for this protein occur in the buffer, discussed in Section 4.5.2. If this is the case for most proteins, the minimum gain may be closer to a factor of 3000 than 30. We also found that increasing the cryoprotectant concentration increased the rate of change at 100 K. Thus, as the technique moves to faster cooling and lower cryoprotectant concentration the sensitivity at 100 K should decrease.

4.5.2 Origin of damage at 100 K

Radiation damage mechanisms in room temperature SAXS has been discussed in detail elsewhere (Kuwamoto et al., 2004; Meisburger et al., 2013; Jeffries et al., 2015; Hopkins & Thorne, 2016). Briefly, a ~10 keV X-ray is absorbed by protein or solvent, creating a shower of secondary electrons (Singh & Singh, 1982). These secondary electrons generate a cascade of radiochemical reactions and additional electrons within ~3 μm of the absorption event, creating various radical species (Cole, 1969; Garrison, 1987; Nave & Hill, 2005; Holton, 2009; Bobrowski, 2012). These radical species diffuse through the solution, reacting with both solvent and macromolecules (Davies, 1987, 2012; Davies & Delsignore, 1987; Davies et al., 1987; Saha et al., 1995; Houée-Levin & Sicard-Roselli, 2001; Stadtman & Levine, 2003; Bobrowski, 2012). The macromolecules undergo radical induced damage processes, including conformational

changes, unfolding, aggregation, and fragmentation (Garrison, 1987; Houée-Levin & Sicard-Roselli, 2001; Stadtman & Levine, 2003; Davies, 2012).

By cooling the sample to 100 K, the diffusion of the radiochemically generated radicals, of other non-radical products of radiolysis, and of the biomolecules themselves is essentially stopped (Kempner, 1993), preventing damage due to diffusing radicals, formation of hydrogen bubbles, formation of aggregates, and also molecular fragmentation. For the macromolecules, the frozen solvent around the molecule inhibits large structural relaxations following damage, such as conformational changes, unfolding, or fragmentation. Thus, cryoSAXS should be successful at preventing radiation-induced changes to macromolecular scattering profiles because cryocooling prevents the main cause of damage, radical reactions, and results of damage - structural changes, aggregation, and fragmentation, from occurring. For protein crystals, where the ratio of protein to solvent is roughly 1:1, cryocooling typically reduces damage per unit dose of X-rays by a factor of ~30-50 (Warkentin et al., 2013).

Figure 4.15 shows scattering profiles as a function of dose for glucose isomerase at 100 K, with the background scattering and the first buffer-subtracted protein scattering profile subtracted, $I_{nom,B}$, and for the protein-free 36% w/w PG buffer 1, I_B . These both show very similar changes. This, combined with the similar sensitivities, Table 4.2, shows that most or all of the changes measured by cryoSAXS occur to the buffer rather than the protein. Based on the relative concentrations of the buffer components, and the fact that most radiation products will be immobile at 100 K, we expect the changes are occurring primarily to the water and/or propylene glycol (an alcohol), and not the buffering agent or salt. Figure 4.5 and Table 4.2 show that may be some influence due to changes in pH, buffering agent, and salt, though the change in the sensitivity is small.

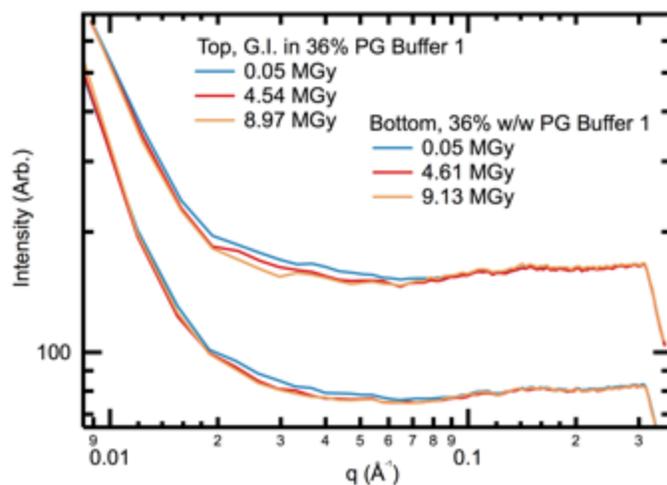


Figure 4.15 – Scattering profiles as a function of dose for glucose isomerase in free 36% w/w PG buffer 1 and the protein-free 36% w/w PG buffer 1 at 100 K. The scattering profiles for glucose isomerase are with the first buffer-subtracted protein scattering profile removed, and with the instrument background scattering subtracted, $I_{nom,B}$. The protein-free 36% w/w PG buffer 1 is with the instrument background subtracted, I_B . The glucose isomerase profiles have been offset by a factor of 2 for clarity. The two sets of scattering show very similar dose dependent changes.

Experiments on pure water systems have shown that at low temperatures the primary radiation generated radicals are *OH (ice and glassy systems) and *HO_2 (only glassy systems) (Pikaev et al., 1970; Bednarek et al., 1996, 1998a; b, 1999b). In binary aqueous glass systems, electrons and atomic hydrogen can also be trapped (some references include (Kroh & Plonka, 1977; Riederer et al., 1983; Plonka et al., 1984), but these indicate that this is commonly accepted and do not provide further references). The electrons act as color centers, giving the glass a characteristic blue color, something we observed in our samples during and after irradiation. The final products of irradiation of pure crystalline ice and vitreous aqueous solutions include molecular hydrogen and oxygen (H_2 and O_2), and hydrogen peroxide (H_2O_2) (Sharpatyi, 1963; Zheng et al., 2006, 2007).

Irradiation of neat alcohols at low temperature results in the production of electrons, alkoxy radicals and hydroxyl alkyl radicals (the latter two are often called macroradicals) (e.g. Lund et al., 1986; Awadelkarim et al., 1988 and references therein). When cryocooled aqueous alcohol solutions are irradiated, as in our experiments, the trapped electrons react with alcohol molecules to form hydroxyl alkyl radicals (Awadelkarim et al., 1988). For polymers of alcohols in water at 77 K, similar macroradicals are observed, along with the radicals expected from pure water (Zakurdaeva et al., 2005, 2010). There is also evidence that molecular hydrogen is formed at 77 K in glassy ethanol, and this may be a general feature of irradiation of solid alcohols (Fletcher & Freeman, 1967; Makarov et al., 1969; Hase & Ishioka, 1992; Ichikawa & Yoshida, 1996).

Changes are seen at most q values in both protein and buffer. It is well known that changes in the minor components of the buffer, such as salt type or concentration, can affect not just the protein but also the buffer scattering. Enough small radiation induced changes, such as production of H_2 or O_2 molecules, radiation induced strain the sample, or generation of large quantities of radicals,

may produce the measured changes in the scattering profile. Radiation generated radicals from the water and macroradicals from the propylene glycol may also react with the propylene glycol or water molecules immediately adjacent, allowing further structural changes.

It has been claimed that radiolytically produced hydrogen gas is responsible for the majority of global radiation damage in protein crystals at 100 K (Meents et al., 2010). The authors suggested that build-up of hydrogen gas results in stresses in the crystal lattice, decreasing the diffraction quality. This gas formation is spectacularly manifested in protein crystallography: when crystals that are irradiated at 100 K are warmed, the crystal bubbles and explodes. The authors also report large production of hydrogen gas, measured by evolution upon warming, from irradiated ethylene glycol at low temperatures, they measured production rates of gas of ~ 35 nmol/MGy for irradiation at 100 K. The authors also report that production due to direct radiolysis of water is less rapid, ~ 0.01 nmol/MGy. As propylene glycol has a very similar composition and structure to ethylene glycol, irradiation may also produce significant amounts of H_2 . Similar to the case for protein crystals, H_2 gas formation in our samples could lead to localized build-ups of gas, followed by increased strain in the sample and a change in the scattering.

In crystalline water ice, irradiation-produced hydrogen gas starts to evolve from samples near 90 K, and is completely evolved by 140 K (Zheng et al., 2006). In amorphous water ice films, diffusion of hydrogen gas through 26 monolayers of D_2O took up to ~ 100 s at 100 K (Petrik & Kimmel, 2004). On this basis, it has been suggested in protein crystals that there must be an observable timescale to global radiation damage between 90-140 K (Warkentin et al., 2011). This has not been observed, which the authors suggest is because hydrogen gas contributes to radiation damage in ways equivalent to all the other radiation products. It is also hard to justify why hydrogen gas would have the specific effects we see from the scattering profile, on the

timescales we have observed the damage. Thus, while the theory of (Meents et al., 2010) is attractive in that it provides a simple mechanism to explain our observed damage, we believe hydrogen gas is unlikely to be the sole source radiation damage.

While changes to the buffer are the dominant component, there may also be damage to the protein. In particular, there is a small increase at high q in the protein scattering profiles, Figure 4.4a, that is not readily apparent in the buffer. Damage to proteins in frozen solution is typically attributed to direct absorption of a photon by the macromolecule or the solution immediately around the macromolecule, and the subsequently generated energetic electrons (Kempner, 1993, 2011; Bednarek et al., 1999a; Houée-Levin & Sicard-Roselli, 2001; Staluszka et al., 2003, 2008). Damage in cryocooled aqueous solutions is typically observed as radiolytic cleavage of N-C bonds in either the protein backbone or side chains (Garrison, 1987; Kempner, 1993). This results in fragmentation of the protein without dissociation, by far the most commonly observed type of change to proteins in frozen solutions (Filali-Mouhim et al., 1997; Miller et al., 1998; Belloni et al., 2000; Shalaev et al., 2003; Audette-Stuart et al., 2005; Terryn et al., 2005). Measurements on proteins in solutions warmed subsequent to damage at cryo temperatures yield sizes equal to undamaged samples when measured with non-denaturing gels or size exclusion chromatography, while SDS-PAGE shows multiple fragments (Kempner, 1993, 2011; Filali-Mouhim et al., 1997; Miller et al., 1998; Audette-Stuart et al., 2005). It is thought that denaturation is required to remove stabilizing interactions within the protein and allow dissociation. A loss of activity is also seen for protein solutions irradiated at 100 K. This damage primary located at the N-C bonds is not seen in site specific damage observations in protein crystallography at 100 K (Burmeister, 2000; Ravelli & McSweeney, 2000; Weik et al., 2000; Warkentin et al., 2012). It is unclear why there are differences in these observations, though it

has been suggested that it is due to the difference in water content and total dose delivered in the observations (Terry et al., 2005). In systems of nucleosides, nucleotides, and nucleobases in vitreous water, some damage to the macromolecules has been observed from radicals produced in the solvent (Bednarek et al., 1999a; Staluszka et al., 2003, 2008), which may also contribute to damage in cryoSAXS.

The source of the catastrophic failure, described in Section 4.4.2 and shown in Figure 4.6, is uncertain. There were no clear visual changes in the samples corresponding with the large change in scattering. In the $\sim 1/3^{\text{rd}}$ of samples that showed this behaviour, the failure happened at doses between $\sim 3.5\text{-}9.6$ MGy (times $\sim 180\text{-}420$ s after initiation of exposure, exposure started ~ 120 s after initial cooling). In order to test whether dose or time since cooling was the important variable, a sample was left sitting at 100 K without X-ray exposure for over 1 hour. This sample showed no catastrophic failure, either after sitting without X-rays or after the standard X-ray exposure. Due to limited beamtime, we were unable to repeat this test, so while it is suggestive that dose is the important variable, it is not conclusive.

In general, radiation damage involves breaking of bonds and an increase in average atomic separations, and so irradiated samples tend to expand, as seen by expansion of lattice parameters with dose in crystallography (Holton, 2009; Garman & Weik, 2015). In the confined geometry of our sample cell, irradiation will thus increase sample stress. If a critical stress is reached, plastic failure via formation of fractures or other defects occurs. The presence of a fracture or other defect in the beam could have strong anisotropic scattering (depending on size, orientation, etc.), as observed. Fracturing at a size scale smaller than the pre-existing optically visual fractures would account for the higher intensity and larger q range of the scattering observed during catastrophic failures, relative to the pre-existing fractures. The lack of consistent

occurrence and dose threshold for catastrophic failure could be attributed to different locations of the initial fractures created by cooling, affecting when/where the stress in the sample becomes critical. This explanation does not account for the observed difference in rate of occurrence in 36% w/w PG buffer, 5 or 16 samples, and 50% w/w PG buffer, 3 of 3 samples.

A possible alternative explanation is formation of hydrogen bubbles in the sample. Hydrogen bubble formation has been observed in cryo electron microscopy (cryo-EM) of samples with high organic content (such as 50% glycerol/water solutions), albeit at doses near 45 MGy (Dubochet et al., 1988; Leapman & Sun, 1995; Baker & Rubinstein, 2010). It has been claimed that hydrogen bubbles are the primary cause of observed strong increases in the SAXS scattering from protein crystals near 70 MGy of dose (Meents et al., 2010). The formation of these bubbles in cryo-EM is only weakly dose rate dependent (Dubochet et al., 1988), suggesting that total absorbed dose is the important variable. If cryoSAXS is a more sensitive detection mechanism for these bubbles or fractures it would explain the discrepancy in observation dose between cryoSAXS and cryo-EM. However, due to the difference in damage mechanisms (electron irradiation can create atomic voids inside solid materials such as Silicon from the direct collisions with highly energetic electrons), these bubbles may be harder to create in cryoSAXS, and so may not be the cause.

4.5.3 Comparison to previous results

It was previously reported that for 2 mg/mL glucose isomerase in a 55% w/w 100 mM Tris pH 8.0 and 1 mM MgCl₂, 45% w/w polyethylene glycol 200 (PEG 200) buffer no damage was observed after 3.7 MGy of dose (Meisburger et al., 2013). The beam profiles and scattering profiles for glucose isomerase as a function of dose from that paper were obtained and

reanalysed using the methods described in this paper. The dose rate adjusted for the beam shape (Hopkins & Thorne, 2016) is decreased by a factor of 1.9, so the data reported is for a total of 1.95 MGy of dose. The results of this analysis are reported in Tables 4.1 and 4.3 with the corresponding measurements from this paper. We found $S_{rg} = 8.55 \cdot 10^{-4}$ %/kGy for the R_g calculated by DATGNOM (the AUTORG and pseudo- R_g sensitivities were $4.02 \cdot 10^{-4}$ %/kGy and $6.76 \cdot 10^{-4}$ %/kGy respectively). The molecular weight sensitivity was $S_{mw} = -6.3 \cdot 10^{-3}$ %/kGy. For the standard subtracted scattering profile, $S_I = 4.3 \cdot 10^{-3}$ %/kGy. Radiation damage was reported for only one sample in (Meisburger et al., 2013), so standard deviations could not be calculated. This sample was in a 45% w/w PEG 200 buffer, compared with our 36% w/w PG buffer. Both cryoprotectant concentration and presence/concentration of other buffer components seem damage rates, so we expect the difference in buffers altered the sensitivities.

In this work, we found S_{rg} , S_{mw} , and S_I to be $1.52 \cdot 10^{-4} \pm 0.6 \cdot 10^{-4}$ %/kGy, $1.7 \cdot 10^{-3} \pm 0.7 \cdot 10^{-4}$ %/kGy, and $1.3 \cdot 10^{-3} \pm 0.04 \cdot 10^{-3}$ %/kGy respectively. The integrated intensity sensitivity will depend on the q range being integrated over, and so is not directly comparable. Recall that for the samples measured in this work, both positive and negative values were measured for S_{rg} and S_{mw} , and we reported the average of the absolute value. Thus, the sign change of S_{mw} is not necessarily meaningful. All of the sensitivities for the data from (Meisburger et al., 2013) are larger than those measured in this work, suggesting that damage to either the protein-free 45% w/w PEG 200 buffer or protein in the PEG 200 buffer progresses more quickly than damage to our samples.

4.5.4 Comparison to expected results

In the work of (Meisburger et al., 2013) it was speculated that the dose limit for cryoSAXS would be between ~50-200 MGy. The upper limit of 200 MGy was proposed based on the observation that the damage rate for biological samples in x-ray crystallography and x-ray diffraction microscopy is very similar, on a per dose basis, at 100 K (Howells et al., 2009). This lead Howells et al. to propose a maximum tolerable dose to obtain a given spatial resolution of data (in Å) of $\sim 10 \text{ MGy}/\text{Å}$. For $q = 0.3 \text{ Å}^{-1}$ the resolution is $\sim 20 \text{ Å}$, yielding a maximum tolerable dose of 200 MGy (Meisburger et al., 2013).

There are several reasons this assessment may be inaccurate for cryoSAXS. First, while for overall structural comparisons, the “resolution” of the data set may be comparable to or less than the naïve conversion offered above (Rambo & Tainer, 2013), the radius of gyration, a structural parameter, is often determined to an Angstrom or better (certainly it is more well determined than 20 Å). In general, SAXS is highly sensitive to some structural parameters, such as overall size of the molecule, but that does not translate into resolution for the full three dimensional structure. Thus, if the radiation damage affects parameters to which SAXS is highly sensitive, the effective spatial resolution may be much better than the $\sim 20 \text{ Å}$ used in (Meisburger et al., 2013).

Second, in crystallography the dose limit used to derive the $10 \text{ MGy}/\text{Å}$ rule was based on the dose required to decrease the intensity of diffraction spots at a given resolution by a factor of 2. In SAXS, a 50% change in intensity represents a large amount of damage, a more reasonable dose limit might be a 1% change in the integrated intensity. A functional form of the dose limit from (Howells et al., 2009) in terms of the percentage decay of intensity at a given resolution was proposed by (Holton, 2009),

$$\frac{I}{I_0} = \exp\left[-\ln(2)\frac{D}{Hd}\right], \quad (4.1)$$

where I/I_0 is the fractional change in diffraction spot intensity, D is the dose in MGy, d is the d -spacing in Å ($q=2\pi d$), and H is the Howells limit of 10 MGy/Å. Simply applying this formula to a 1% intensity change and 20 Å resolution yields a dose limit of 2.9 MGy. Table 4.7 gives the dose required to the integrated intensity by 1%, and finds a range of 0.8-6.7 MGy (depending on the metric used), in good agreement with the estimated 2.9 MGy.

The proposed lower limit of ~50 MGy was based on the observation of the formation of hydrogen bubbles in cryo electron microscopy around 45 MGy of dose (Leapman & Sun, 1995; Baker & Rubinstein, 2010) and the observation of a strong increase in the SAXS scattering profile of insulin crystals attributed to hydrogen bubble formation near 70 MGy (Meents et al., 2010). It is not clear that either of these observations is relevant for cryoSAXS. Hydrogen bubbles would appear as strong anisotropic scatter, rather than an overall increase in the scattering profile, inconsistent with the non-catastrophic damage modes observed. Therefore, these observations do not provide a reasonable damage limit for cryoSAXS.

4.5.5 Room temperature damage

Damage to glucose isomerase at room temperature has been characterized previously (Jeffries et al., 2015; Hopkins & Thorne, 2016). The damage at room temperature to glucose isomerase in the PG-free buffer 1 is similar in form to that reported in (Hopkins & Thorne, 2016), however our measured S_{rg} , $1.1 \cdot 10^{-2} \pm 0.5 \cdot 10^{-2}$, was ~250 times greater than their reported value of $S_{rg} = -0.5 \cdot 10^{-4} \pm 7 \cdot 10^{-4}$. We found $S_{mw} = -0.11 \pm 0.05$ %/kGy, while (Hopkins & Thorne, 2016)

Table 4.7 – The dose required for a 1% change in the metric ($1\%/S$) for samples at 100 K calculated as. Doses are given for glucose isomerase in 36% w/w PG buffer 1 and protein-free 36% w/w buffer 1 and protein-free 36% w/w and 50% w/w PG buffer 2. From left to right, the columns are the dose for a 1% change in $S_{I,nom,B}$ (protein only), $S_{I,B}$ (buffer only), $S_{I,sub}$ (protein only), S_{rg} , and S_{mw} .

Sample	$D_{1\%,I,nomB}$ (kGy)	$D_{1\%,I,B}$ (kGy)	$D_{1\%,I,sub}$ (kGy)	$D_{1\%,rg}$ (kGy)	$D_{1\%,mw}$ (kGy)
G.I., 36% PG	6700	-	770	6600	570
Buffer 1, 36% PG	-	5000	-	-	-
Buffer 2, 36% PG	-	3700	-	-	-
Buffer 2, 50% PG	-	1900	-	-	-

reported $S_{mw} = -0.10 \pm 0.08$, which agree well. The large difference in S_{rg} may simply indicate that the radius of gyration is not a useful measurement of the radiation damage to glucose isomerase, as suggested previously (Hopkins & Thorne, 2016). Alternatively, the difference may be due to differences in buffer, concentration, dose rate, or X-ray energy, though these factors seem unlikely to cause a change of that magnitude. The integrated intensity sensitivity for the standard subtracted protein scattering profile depends on the q ranges measured and the instrument background and so is not comparable to the value from (Hopkins & Thorne, 2016).

Jeffries et al. (2015) measured the dose to change the pseudo- R_g by 1 Å, CD_J , for glucose isomerase. They report 6.5-7 kGy (their reported doses do not account for the actual beam shape, and so may be underestimated by a factor of 2 or more (Hopkins & Thorne, 2016)). The uncertainty in our measurement of the pseudo- R_g is ~ 1.5 Å, so we cannot determine CD_J with any accuracy, making it hard to compare damage rates. Naively carrying out the calculation for our data gives, on average, $CD_J = 2.9$ kGy. (Hopkins & Thorne, 2016) discuss why the reported glucose isomerase data from (Jeffries et al., 2015) may be unreliable, making this comparison less useful.

The addition of propylene glycol to the buffer increased the damage rate of the protein by one to two orders of magnitude, depending on the sensitivity examined. The changes that occurs in the protein-free PG buffers are likely dominated by radiation induced polymerization and crosslinking of the propylene glycol monomers in solution, as observed for other polymers (Alexander & Charlesby, 1957; Rosiak & Ulański, 1999; Buwalda et al., 2014). The formation of larger units in solution fits with the increase in intensity at low- q . Differences in the rate of crosslinking has been observed in polymer solutions with different pHs, different polymer

concentrations, and different small molecule compounds (Alexander & Charlesby, 1957; Wenxiu et al., 1985; Wang et al., 1998), which may explain the observed differences between the sensitivities of our different PG buffers at room temperature.

We were unable to find reports of radiolysis experiments on propylene glycol. Ethylene glycol is quite similar in structure and composition, and has been reported to generate a wide number of radiation products, including relatively small amounts of diethylene and triethylene glycol (Barker et al., 1964). Additionally, longer chain poly(ethylene glycol) is often used to make hydrogels via radiation induced crosslinking (Keys et al., 1998). This supports the possibility of polymerization and crosslinking in our propylene glycol solutions, but is not conclusive, and there may be alternative explanations. Other reaction products may aggregate or polymerize, leading to the observed damage.

4.5.6 Implications for minimum sample volume

The reduction of radiation sensitivity allows a reduction of the total sample volume used to obtain high quality SAXS scattering profiles. It has previously been demonstrated that high quality cryoSAXS scattering profiles could be collected from ~0.8-1 μL total sample volumes with nominal X-ray illuminated volumes as small as 2.5 nL (Meisburger et al., 2013; Hopkins et al., 2015). Total sample volumes of ~50-100 nL with nominal illuminated volumes of ~15-30 nL have been used to measure somewhat noisy cryoSAXS scattering profiles (Meisburger et al., 2013). Using our measurement of the 100 K radiation sensitivity, we estimated the minimum sample volume that can be used in cryoSAXS.

The scattering intensity from a sample is given by

$$I(q) \propto I_0 A l T (\Delta\rho)^2 P(q) , \quad (4.2)$$

where I_0 is the total fluence at the sample, A is the area illuminated by the beam, l is the X-ray path length in the sample, T is the sample transmission, $\Delta\rho$ is the contrast, and $P(q)$ is the single particle form factor (note that we have ignored factors such as the detector efficiency that do not affect our estimate) (Orthaber et al., 2000; Dreiss et al., 2006). As background scattering varies from beamline to beamline and currently cannot be well predicted, we will ignore it. Equation (4.2) can be used to find a ratio of the necessary illuminated volumes at room temperature and 100 K to achieve equal measured intensity. The details of this are given in Section 4.7.2. For the simplest case of equal path length (1.5 mm) samples at room temperature and 100 K, the volume ratio for equal measured intensity is given by

$$\frac{V_{cryo}}{V_{rt}} = 0.029 \quad (4.3)$$

It is more intuitive to look at the inverse ratio, $V_{rt}/V_{cryo} = 35$. So, for glucose isomerase cryocooling will allow a reduction in illuminated volume by a factor of 35. Because glucose isomerase is extremely radiation tolerant at room temperature (Hopkins & Thorne, 2016), we expect this ratio will be much greater for other proteins. If the same calculation is done comparing the sample in 36% PG buffer 1 at both temperatures, $V_{rt}/V_{cryo} \sim 5000$.

The question of minimum possible sample volume also must consider the constraints imposed by the sample holder. In order to avoid reflection of the beam tails off of the side walls of the sample holder, sample holder dimensions must be significantly larger than the beam dimension, placing a limit on the fraction of the total volume that is illuminated by X-rays. Section 4.7.2

gives details of our calculation, which found that in an optimally sized room temperature flow cell, the ratio of illuminated volume, V_{il} , to total sample volume, V_s , is

$$\frac{V_{il}}{V_s} = 0.52. \quad (4.4)$$

This ratio represents that maximum possible sample usage, given our assumption of a Gaussian beam and our conditions to avoid reflections from the sample holder. This maximal ratio is reduced by ~10% (the exact value depends on the sample holder geometry, beam size and shape, and sample volume), due to the inability to use data when the beam is close to the sample meniscus.

For standard setups, this optimum value is not usually achieved. For example, at beamline P12 at PETRAIII a 1.7 mm ID capillary is used as the sample holder, and the beam size is 0.2 (H) x 0.11 (V) mm FWHM (Jeffries et al., 2015). Assuming the beam is Gaussian and the flow direction is horizontal gives a volume fraction ratio of $V_{il}/V_s = 0.19$ (if the flow direction is vertical, $V_{il}/V_s = 0.34$), not considering the meniscus effect. The volume fraction ratio is less clear in oscillation at room temperature, where mixing must be accounted for when calculating the illuminated volume.

For a 100 K sample, take a cylinder with path length l as the sample holder, such as that used by (Meisburger et al., 2013). The maximum illuminated volume ratio for this sample holder is

$$\frac{V_{il,g}}{V_s} = 0.12 \quad (4.5)$$

The small volume holder used by (Meisburger et al., 2013) could also be considered a cylinder for calculation purposes. Recalculating their illuminated volumes as described in Section 4.7.2,

we find $V_{il}/V_s \sim 0.5$ ($V_{il}/V_s \sim 0.3$ for their reported illuminated volumes calculated using the FWHM), in excess of the maximum ratio of Equation (4.5). We believe this was possible because the small volume holder had a very thin polyimide support, minimizing/eliminate beam reflections. In their 1 μL cylindrical holder, the illuminated volume fraction was ~ 0.09 . For the work in this paper, using the silicon sample holders, the illuminated volume fraction at 100 K was ~ 0.08 .

Based on the maximum illuminated volume fractions, Equations (4.4) and (4.5), a cylindrical cryoSAXS sample holder will need 4.7 times more volume than a flow cell to achieve the same illuminated sample volume with beam reflections. Based on the actual illuminated volume fractions, this ratio is closer to 2-3, and could possibly be reduced to 1 (or less than 1) with the right sample holder geometry. For example, an optimized rectangular prism sample holder gives an illuminated volume fraction of ~ 0.22 , comparable with our estimate of an actual illuminated volume fraction in RT SAXS, though still a factor of ~ 2.4 lower than the optimal system at room temperature.

Due to the effect of sample holder shape on fraction of illuminated, the maximum volume ratio, Equation (4.3), is actually 2-3 times greater for current standard practices. This means that for glucose isomerase, in terms of the total volume used, $V_{t,rt}/V_{t,crysto} \sim 10-20$. Note that due to the extreme radiation tolerance of glucose isomerase at room temperature we were able to collect a relatively high quality scattering profile (Figure 4.16) from $\sim 3 \mu\text{L}$ of solution, something that is not possible for most proteins. This suggests that for glucose isomerase, routine use of total volumes of $\sim 150-300 \text{ nL}$ should be possible. If the sensitivity of proteins at 100 K is limited by the buffer, all proteins will damage at similar rates at 100 K. For example, lysozyme is at least 500 times more sensitive (up to 50000 times more sensitive, depending on metric used) than

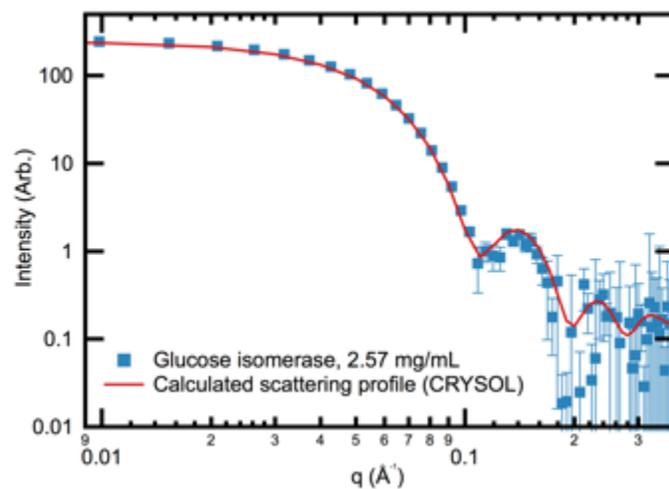


Figure 4.16 – Scattering profile for 2.57 mg/mL glucose isomerase in the PG-free buffer 1 at room temperature. The scattering profile calculated by CRY SOL (Svergun et al., 1995) from a high resolution structure (PDB 1XIB, with symmetry mates added in PyMOL) is also shown. The data is of relatively high quality for a sample of $\sim 3 \mu\text{L}$.

glucose isomerase (Hopkins & Thorne, 2016), so we would expect, at the very least,

$$V_{t,rt} / V_{t,cryo} \sim 5000 - 10000 .$$

4.6. Conclusion

We have found that radiation damage in cryoSAXS (at 100 K) is mostly due to changes that happen to the buffer, rather than to the protein. For two of the metrics of radiation damage considered, the radiation tolerance at 100 K increased by a factor of $\sim 5,000$ relative to identically prepared glucose isomerase samples at room temperature, while for the third metric the increase was a factor of 650. When compared to glucose isomerase in a standard room temperature buffer (i.e. without cryoprotectant), the radiation tolerance of the cryoprotected and cryocooled sample was $\sim 60x$ larger. This is much less than the $\sim 5,000x$ or $650x$ increase in identical samples, and the difference is due to the decrease in radiation tolerance at room temperature when cryoprotectant is added to the buffer. If radiation damage is limited by the buffer for other proteins, similar 100 K radiation sensitivities are expected for different proteins. Larger increase in radiation tolerance may be possible for proteins that are more sensitive at room temperature. For example, lysozyme is $\sim 1,000$ times more sensitive at room temperature than glucose isomerase (depending upon metric considered) (Hopkins & Thorne, 2016), so gains in radiation tolerance may be as large as $60,000x$. We also measured the rates of radiation-induced change for several protein-free buffers at 100 K and found that the rates depend upon both the cryoprotectant concentration and the other buffer components.

The radiation tolerance of samples in cryoSAXS may increase as improvements are made to the technique. Faster freezing, allowing use of less cryoprotectant, may result in slower damage rates. We saw that an increase from 36% to 50% w/w PG decreased the radiation tolerance of the

protein-free buffer 2 by a factor of 2. For 10 nL samples, similar to those used in (Meisburger et al., 2013), cooling rates near 10,000 K/s can be achieved by careful plunge cooling in liquid nitrogen or propane (Warkentin et al., 2006, 2008). The sample holders used in this work cool at roughly 20 K/s (Hopkins et al., 2015). For glycerol this would allow a reduction of necessary concentration of cryoprotectant by ~25% w/w (Warkentin et al., 2008). A similar reduction for PG (assuming the damage rate is linear in PG concentration) would increase the radiation tolerance by a factor of ~4. Careful choice of the other buffer components may also allow additional gain in the radiation tolerance.

The gain factors reported, ~60-5000x, represent extremely large increases in the radiation tolerance of the samples. The difference in damage rate for protein crystallography between room temperature and 100 K is typically a factor of 30-50 (Warkentin et al., 2013), and that was significant enough that it is estimated that more than 90% of the crystal structures are now determined at cryogenic temperatures (Garman & Owen, 2006). A 5000-fold increase in radiation tolerance allows a roughly commensurate decrease in sample volume, which could reduce the standard sample volumes in SAXS from near 10 μ L to ~5-10 nL. Even just a factor of ~30 increase in the radiation tolerance, the minimum found here, would allow use of 300 nL samples. While there is still work required to make cryoSAXS a routine technique, this shows the rewards will be significant.

4.7 Supporting information

4.7.1 Data processing

First a mask was applied and the detector images were azimuthally integrated to generate a scattering profile. Some data processing was done according to standard SAXS protocols. For

these, the scattering profiles were then normalized by the transmitted intensity, correcting for variations in the incident intensity or sample absorption. Protein-free buffer profiles were subtracted from scattering profiles of proteins in solution to generate the scattering profile of the proteins alone.

In general, the scattering profile from a sample may be decomposed into the contribution from three sources, the macromolecule, I_M , the macromolecule-free buffer, I_B , and the instrument background and sample holder I_{bkg} :

$$I_{total}(q) \propto I_0 T \left\{ \log \left(\frac{1}{T} \right) [I_M(q) + I_B(q)] + I_{bkg}(q) \right\} , \quad (4.6)$$

where T is the transmission of the sample (Meisburger et al., 2013). In order to compare the damage rate of protein-free buffers at 100 K and room temperature, the total integrated intensity was used as the parameter. To accurately compare the change in integrated intensity between protein-free buffer curves at different temperatures (and in different sample holders), the contribution to the scattered intensity from the background must be removed. Without the macromolecular contribution, the measured scattering from the protein-free buffer, I_{buff} , is

$$I_{buff}(q) \propto I_{0,buff} T_{buff} \left\{ \log \left(\frac{1}{T_{buff}} \right) I_B(q) + I_{bkg}(q) \right\} . \quad (4.7)$$

The scattering from just the empty sample holder is

$$I_{em}(q) \propto (I_{0,em} T_{em}) I_{bkg}(q) . \quad (4.8)$$

The leading factor is accounted for in standard SAXS by normalizing by the transmitted intensity, $I_{trans} = I_0 T$ (such as the measured intensity on an active beamstop). The transmission of the protein-free buffer can be directly measured as

$$T_{buff} = \left(\frac{I_{trans}^{buff}}{I_{in}^{buff}} \right) \times \left(\frac{I_{in}^{em}}{I_{trans}^{em}} \right), \quad (4.9)$$

where I_{in} is the incident intensity upstream of the sample. Knowing this, the contribution of just the protein-free buffer can be extracted:

$$I_B(q) \propto \left(\frac{I_{buff}(q)}{I_{trans}^{buff}} - \frac{I_{em}(q)}{I_{trans}^{em}} \right) \frac{1}{\log(1/T_{buff})}. \quad (4.10)$$

Use of Equation (4.10) allows comparison of rates of change of protein-free buffers at different temperatures collected in different sample holders. If the normalized integrated intensity is considered, as is the case in this paper, the $1/\log(1/T_{buff})$ factor will cancel out and so does not need to be calculated.

We also wanted to compare the damage rate of the protein to the rate of change of the protein-free buffer at the same temperature. In order to do this, the integrated intensity was again used, but in order to have accurate comparison the initial contribution of the protein had to be removed from the protein scattering. If this was not done, the initial intensity of the protein scattering would be greater than the protein-free buffer, and so for the same absolute change the relative rates of change of the protein and protein-free buffer scattering would be different. This was accomplished by first generating the standard subtracted protein scattering profile for the initial exposure of the sample

$$I_{sub,p0}(q) = \frac{I_{p0}(q)}{I_{trans}^{p0}} - \frac{I_{buff}(q)}{I_{trans}^{buff}}, \quad (4.11)$$

where $I_{p0}(q)$ is the first measured sample scattering profile (without the protein-free buffer subtracted). This was then subtracted from each subsequent scattering profile to generate a nominal protein-free buffer scattering profile $I_{nom,buff}(q) = I_p(q) / I_{trans}^p - I_{sub,p0}(q)$, where $I_p(q)$ is the unsubtracted sample scattering profile at some dose. Finally, in order to minimize the number of necessary comparisons and calculations, the background scattering was removed from $I_{nom,buff}(q)$, similar to Equation (4.10), giving $I_{nom,B}(q)$.

$$I_{nom,B}(q) = \left(I_{nom,buff}(q) - \frac{I_{em}(q)}{I_{trans}^{em}} \right) \frac{1}{\log(1/T)} \quad (4.12)$$

The transmission of protein-free buffer and protein samples agreed to better than 0.2%, so the transmission of the protein sample was used for ease of calculation. For this scattering profile, if the scatter from the protein is changing, then the change of $I_{nom,B}(q)$ will not be the same as the protein-free buffer. If there is no change in the protein scatter, it will show change identical to the protein-free buffer.

4.7.2 Illuminated volume

The reduction of radiation sensitivity allows a reduction of the total sample volume used to obtain high quality SAXS scattering profiles. It has previously been demonstrated that high quality cryoSAXS scattering profiles could be collected from ~0.8-1 μ L total sample volumes with nominal X-ray illuminated volumes as small as 2.5 nL (Meisburger et al., 2013; Hopkins et al., 2015). Total sample volumes of ~50-100 nL with nominal illuminated volumes of ~15-30 nL

have been used to measure somewhat noisy cryoSAXS scattering profiles (Meisburger et al., 2013). Using our measurement of the 100 K radiation sensitivity, we estimated the minimum sample volume that can be used in cryoSAXS.

The scattering intensity from a sample is given by

$$I(q) \propto I_0 A l T (\Delta\rho)^2 P(q), \quad (4.13)$$

where I_0 is the total fluence at the sample, A is the area illuminated by the beam, l is the X-ray path length in the sample, T is the sample transmission, $\Delta\rho$ is the contrast, and $P(q)$ is the single particle form factor (note that we have ignored factors such as the detector efficiency that do not affect our estimate) (Orthaber et al., 2000; Dreiss et al., 2006). As background scattering varies from beamline to beamline and currently cannot be well predicted, we will ignore it for the moment. The ratio between intensities at room temperature and cryo is then

$$\frac{I_{cryo}}{I_{rt}} = \frac{I_{0,cryo} A_{cryo} l_{cryo} T_{cryo} (\Delta\rho_{cryo})^2}{I_{0,rt} A_{rt} l_{rt} T_{rt} (\Delta\rho_{rt})^2}, \quad (4.14)$$

where we have assumed that the form factor does not change. In order to relate Equation (4.14) to our radiation damage measurements, we use the fact that $f t = I_0$ and the equation for calculating dose (Hopkins & Thorne, 2016),

$$\text{Dose} = \frac{f t A E_\gamma}{\rho l}, \quad (4.15)$$

giving

$$\frac{I_{cryo}}{I_{rt}} = \frac{(D_{cryo} \rho_{cryo} l_{cryo}) A_{cryo} l_{cryo} T_{cryo} (\Delta\rho_{cryo})^2 (1 - e^{-\mu l_{rt} \rho_{rt}})}{(D_{rt} \rho_{rt} l_{rt}) A_{rt} l_{rt} T_{rt} (\Delta\rho_{rt})^2 (1 - e^{-\mu l_{cryo} \rho_{cryo}})}, \quad (4.16)$$

where D is dose, and μ is the mass-energy attenuation coefficient (Hopkins & Thorne, 2016).

The illuminated volume is given by $V = Al$, so fixing the intensity ratio at 1 gives

$$\frac{V_{cryo}}{V_{rt}} = \frac{l_{rt} e^{-\mu l_{rt} \rho_{rt}} (1 - e^{-\mu l_{cryo} \rho_{cryo}})}{l_{cryo} e^{-\mu l_{cryo} \rho_{cryo}} (1 - e^{-\mu l_{rt} \rho_{rt}}} \left(\frac{\rho_{rt}}{\rho_{cryo}} \right) \left(\frac{\Delta \rho_{rt}}{\Delta \rho_{cryo}} \right)^2 \left(\frac{D_{rt}}{D_{cryo}} \right), \quad (4.17)$$

where we have rewritten the transmission using Beer's law.

The use of cryoprotectant and the cryocooling both change the contrast for cryoSAXS samples.

The average electron density of water is $0.334 \text{ e}^-/\text{\AA}^3$, adding 36% PG raises the bulk solvent electron density to $0.344 \text{ e}^-/\text{\AA}^3$ (Svergun et al., 1995; Hopkins et al., 2015). Cryocooling increases the density of a 36% w/w PG solution from 1.03 g/cm^3 (Khattab et al., 2012) to 1.09 g/cm^3 , further increasing the bulk solvent electron density by 6% to $0.354 \text{ e}^-/\text{\AA}^3$ at 100 K. The average electron density of a protein is $\sim 0.420 \text{ e}^-/\text{\AA}^3$ (Svergun & Koch, 2003), and observed radius of gyration and maximum dimensions from cryoSAXS are comparable or larger than those reported at room temperature so no increase in density of the protein is expected (Meisburger et al., 2013; Hopkins et al., 2015). This gives $\Delta \rho_{rt} / \Delta \rho_{cryo} = 1.3$. We can also evaluate the density ratio, $\rho_{rt} / \rho_{cryo} = 0.96$.

In order to estimate the maximum volume ratio, the maximum doses at each temperature must be used. As discussed in (Hopkins & Thorne, 2016), the dose limit is not particularly well defined as a metric for radiation damage in SAXS. However, for the same change in parameter at each temperature, the ratio of doses can simply be expressed as the ratio of the inverse sensitivities $1/S$, giving

$$\frac{V_{cryo}}{V_{rt}} = (1.6) \frac{l_{rt} e^{-\mu l_{rt} \rho_{rt}} (1 - e^{-\mu l_{cryo} \rho_{cryo}})}{l_{cryo} e^{-\mu l_{cryo} \rho_{cryo}} (1 - e^{-\mu l_{rt} \rho_{rt}}} \left(\frac{S_{cryo}}{S_{rt}} \right). \quad (4.18)$$

The illuminated volume depends upon the path length, so this is a transcendental equation.

The simplest case is for equal path length samples at room temperature and 100 K. In this case, a difference in the illuminated sample volume would result from changing the beam size. Using this assumption, Equation (4.18) reduces to

$$\frac{V_{cryo}}{V_{rt}} = (1.6) \frac{e^{-\mu\rho_{rt}}(1 - e^{-\mu\rho_{cryo}})}{e^{-\mu\rho_{cryo}}(1 - e^{-\mu\rho_{rt}})} \left(\frac{S_{cryo}}{S_{rt}} \right). \quad (4.19)$$

For a 36% w/w PG solution, at 10 keV, $\mu = 4.731 \text{ g/cm}^2$, as calculated by XCOM (Gerward et al., 2001; Berger et al., 2010). Using a 1.5 mm path length,

$$\frac{V_{cryo}}{V_{rt}} = 1.74 \left(\frac{S_{cryo}}{S_{rt}} \right) \quad (4.20)$$

For glucose isomerase, the sensitivity ratio between protein in the PG-free buffer 1 at room temperature and protein in the 36% w/w PG buffer at 100 K is $S_{cryo}/S_{rt} \sim 1/60$ (averaging all three different sensitivity ratios from Section 4.5.1). This gives

$$\frac{V_{cryo}}{V_{rt}} = 0.029 \quad (4.21)$$

It is more intuitive to look at the inverse ratio, $V_{rt}/V_{cryo} = 35$. So, for glucose isomerase cryocooling will allow a reduction in illuminated volume by a factor of 35. Because glucose isomerase is extremely radiation tolerant at room temperature (Hopkins & Thorne, 2016), we expect this ratio will be much greater for other proteins. If the same calculation is done comparing the sample in 36% PG buffer 1 at both temperatures, $V_{rt}/V_{cryo} \sim 5000$.

Equation (4.18) can also be naïvely applied in the unequal path length case, such as was tested in (Meisburger et al., 2013). Using, for example, $l_{rt} = 2$ mm and $l_{cryo} = 0.5$ mm gives $V_{rt} / V_{cryo} = 70$. However, with unequal path lengths, ignoring the background scatter becomes less and less realistic. For shorter path lengths, the background will have a larger contribution to the overall scattering intensity, relative to the protein and buffer scattering. That means that even for nominally the same scattered intensities, as calculated here, the signal to noise ratio will be worse.

The question of minimum possible sample volume also must consider the constraints imposed by the sample holder. In order to avoid reflection of the beam tails off of the side walls of the sample holder, sample holder dimensions must be significantly larger than the beam dimension. It has previously been estimated that the beam intensity at the side walls should be no more than the brightest scattering intensity, a decrease by a factor of 10^6 , to avoid these reflections (Nielsen et al., 2012). This distance will depend on the beam shape, in this discussion we will consider a Gaussian beam shape. Following (Nielsen et al., 2012), for a Gaussian beam with standard deviation σ ($\text{FWHM} = (2\sqrt{2\ln 2})\sigma$), this requires a sample holder dimension of $2\sigma\sqrt{2\ln(10^6)} \approx 10.5\sigma$ ($4.5 \cdot \text{FWHM}$).

In order to find the ratio of illuminated to total volume, a suitable definition of illuminated volume is needed. When the beam shape is not top-hat there is no clear definition for the illuminated volume, as volumes in the beam tails contribute little to the scattering but are technically illuminated. In the past, the FWHM of the beam has been used to find the illuminated volume (Nielsen et al., 2012; Meisburger et al., 2013; Hopkins et al., 2015), however this is a

conservative estimate. We define the illuminated volume as the volume giving the measured dose rate, assuming the incident flux is distributed uniformly across the surface of the volume.

For a Gaussian beam, (Hopkins & Thorne, 2016) found that the dose rate was reduced by a factor of 2.3 relative to a dose rate calculated assuming all the flux was uniformly within the beam FWHMs. From this, we define the illuminated volume as a rectangular prism with a face perpendicular to the direction of beam travel with area 2.3 times the area defined by the beam FWHMs, and the other dimension equal to the X-ray path length. Thus, for a Gaussian beam with standard deviations σ_x and σ_y , the illuminated volume $V_{il,g}$ is

$$V_{il,g} = (2.3) \left(2\sqrt{2\ln 2} \right)^2 \sigma_x \sigma_y l . \quad (4.22)$$

For a room temperature sample, consider a rectangular flow cell with X-ray path length l and dimension perpendicular to the flow direction $2\sigma_y \sqrt{2\ln(10^6)}$, the optimal size. In a cross section of width $2\sqrt{2\ln 2} \sigma_x$, there is a total sample volume V_s of

$$V_s = 2\sigma_x \sigma_y l \sqrt{2\ln(10^6)} \left(2\sqrt{2\ln 2} \right) . \quad (4.23)$$

The ratio of illuminated to total sample volume is

$$\frac{V_{il,g}}{V_s} = 0.52 . \quad (4.24)$$

For a flow system, if exposure is started before the sample is in the beam and ended after the sample has left the beam, such that the entire sample plug is exposed, this ratio represents that maximum possible sample usage, given a Gaussian beam and our conditions to avoid reflections from the sample holder. One additional consideration is that scattering from the sample within

$\sigma_x \sqrt{2 \ln(10^6)}$ of each meniscus will not be useable, reducing this ratio. How much of the volume this represents depends on the size of the sample holder and the loaded volume. For a 1.5 mm ID capillary, 20 μL of loaded volume and a 200 μm beam FWHM, the ratio is reduced by $\sim 8\%$, to ~ 0.48 .

For standard setups, this optimum value is not usually achieved. For example, at beamline P12 at PETRAIII a 1.7 mm ID capillary is used as the sample holder, and the beam size is 200 (H) x 110 (V) μm FWHM (Jeffries et al., 2015). Assuming the beam is Gaussian and the flow direction is horizontal gives a volume fraction ratio of $V_{il}/V_s = 0.19$ (if the flow direction is vertical, $V_{il}/V_s = 0.34$), not considering the meniscus effect. The volume fraction ratio is less clear in oscillation at room temperature, where mixing must be accounted for when calculating the illuminated volume.

For a 100 K sample, take a cylinder with path length l as the sample holder, such as that used by (Meisburger et al., 2013). In this geometry, both sides of the sample holder must be sufficiently far away from the beam, giving a total sample volume of

$$V_s = 4\sigma_y\sigma_x l \ln(10^6). \quad (4.25)$$

The ratio of this with Equation (4.22) gives

$$\frac{V_{il,g}}{V_s} = 0.12. \quad (4.26)$$

This has been exceeded by (Meisburger et al., 2013) in their small volume holder. Recalculating their illuminated volumes as above, we find $V_{il}/V_s \sim 0.5$ ($V_{il}/V_s \sim 0.3$ for their reported illuminated volumes calculated using the FWHM). We believe this was possible because the

small volume holder had a very thin polyimide support, which may minimize or eliminate beam reflections. In their 1 μL cylindrical holder, the illuminated volume fraction was ~ 0.09 . For the work in this paper, using the silicon sample holders, the illuminated volume fraction at 100 K was ~ 0.08 .

Based on the maximum illuminated volume fractions, Equations (4.4) and (4.5), a cylindrical cryoSAXS sample holder will need 4.7 times more volume than a flow cell to achieve the same illuminated sample volume with beam reflections. Based on the actual illuminated volume fractions, this ratio is closer to 2-3, and could possibly be reduced to 1 (or less than 1) with the right sample holder geometry. For example, an optimized rectangular prism sample holder gives an illuminated volume fraction of ~ 0.22 , comparable with our estimates of actual illuminated volume fractions in RT SAXS, though still a factor of ~ 2.4 lower than the optimal system.

Due to the effect of sample holder shape on fraction of illuminated, the maximum volume ratio, Equation (4.3), is actually 2-3 times greater for current standard practices. This means for glucose isomerase, in terms of the total volume used, $V_{t,rt} / V_{t,crysto} \sim 10-20$. Note that due to the extreme radiation tolerance of glucose isomerase at room temperature we were able to collect a relatively high quality scattering profile (Figure 4.16) from $\sim 3 \mu\text{L}$ of solution, something that is not possible for most proteins. This suggests that for glucose isomerase, routine use of total volumes of $\sim 150-300 \text{ nL}$ should be possible. If the sensitivity of proteins at 100 K is limited by the buffer, all proteins will damage at similar rates at 100 K. For example, lysozyme is at least 500 times more sensitive (up to 50000 times more sensitive, depending on metric used) than glucose isomerase (Hopkins & Thorne, 2016), so we would expect, at the very least, $V_{t,rt} / V_{t,crysto} \sim 5000-10000$.

4.7.3 Lysozyme data

In addition to the glucose isomerase used in the paper, data was also collected from hen egg white lysozyme (Affymetrix, Santa Clara, CA, 18645). This data was not presented in the main body of the paper as the lysozyme aggregated upon cryocooling. This will be discussed later. We present addendums to the relevant sections that discuss this additional data.

4.7.3.1 Additional Methods

Lysozyme was reconstituted from powder into a 40 mM sodium acetate pH 4.0, 150 mM sodium chloride buffer and filtered through a 0.22 μm syringe filter (Millex ethylene oxide sterilized, EMD Millipore) to remove undissolved powder. Note that the lysozyme buffer used here is the same as the buffer 2 used in the paper. We will continue to refer to it as buffer 2 to maintain consistency. Samples were mixed into cryoprotectant buffers as described in the paper, and the final concentration of lysozyme was 3.69 mg/mL.

We include data that was collected at the G1 beamline at the Cornell High Energy Synchrotron Source (CHESS). The setup used there was very similar to that described in the main methods, and almost identical to that described in (Hopkins et al., 2015). The energy (wavelength) was 11.17 keV (1.11 \AA), the flux was $5.23 \cdot 10^{11}$ ph/s, the beam size was set by beam defining slits as 150 μm x 150 μm , and the sample to detector distance was 1748.0 mm.

Data processing was carried out as described in Section 4.3.4. The known literature values used for lysozyme were: $R_g = 14.3 \text{ \AA}$ at room temperature (Mylonas & Svergun, 2007), molecular

mass = 14.3 kDa. For lysozyme at 100 K, the initial measured value was used to normalize all parameters.

4.7.3.2 Additional Results

4.7.3.2.1 100 K

All of the lysozyme samples at 100 K were aggregated, with initial R_g of ~20-220 Å depending on the sample. These values were calculated by DATGNOM, and agreed well with the pseudo- R_g . However, the values found by AUTORG range from just ~20-40 Å. The initial molecular weight of the samples was ~23-50 kDa, suggesting that the AUTORG range of 20-40 Å may be more accurate. Regardless of the method used to measure R_g , it was much larger than both the expected room temperature R_g , 14.3 Å (Mylonas & Svergun, 2007), and the measured initial room temperature range, 13-15 Å (with reported uncertainties of ~3-5 Å). The reason for this aggregation is not known, as nominally identical samples yielded unaggregated scattering profiles in previous cryoSAXS experiments at CHESS, Figure 4.17. The large range of initial R_g may be indicative of bad fracturing in some of the samples, or of variability in the behavior on cooling.

Despite this aggregation, the qualitative observations described Section 4.4.2 still applied. Damage to lysozyme in the 36% w/w PG buffer 2 at 100 K is shown in Figure 4.18. Damage appeared as a gradual increase at high q and decrease at low q . Both observations on protein-free buffers are still relevant, as the buffer 2 used in the paper is the buffer used for our lysozyme sample. This is shown in Figures 4.5 and 4.7 in the main paper. Catastrophic failures were also observed for the lysozyme samples.

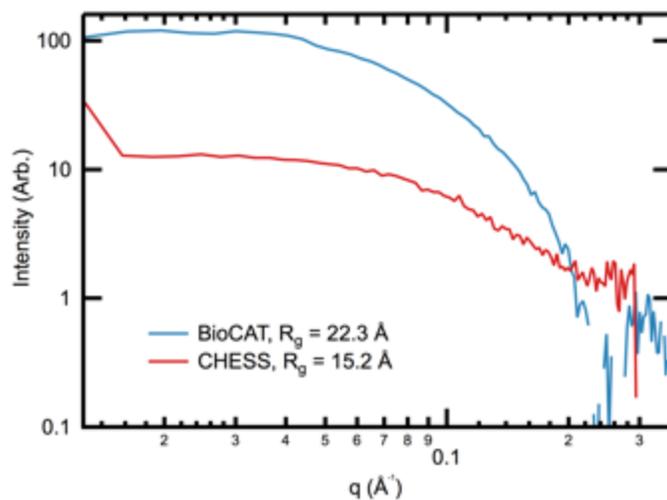


Figure 4.17 – Scattering profiles for lysozyme in nominally identical buffers (36% w/w PG buffer 2) at 100 K, taken at BioCAT and CHESS. The scattering profiles were scaled to match at $q=0.2 \text{ \AA}^{-1}$. The legend gives the R_g for each scattering profile. As can be seen, there is a significant difference between the two curves, the cause of which is unknown. The sharp increase in the scattering profile from CHESS at $q \sim 0.14 \text{ \AA}^{-1}$ is scatter from fractures in the sample that could not be masked out.

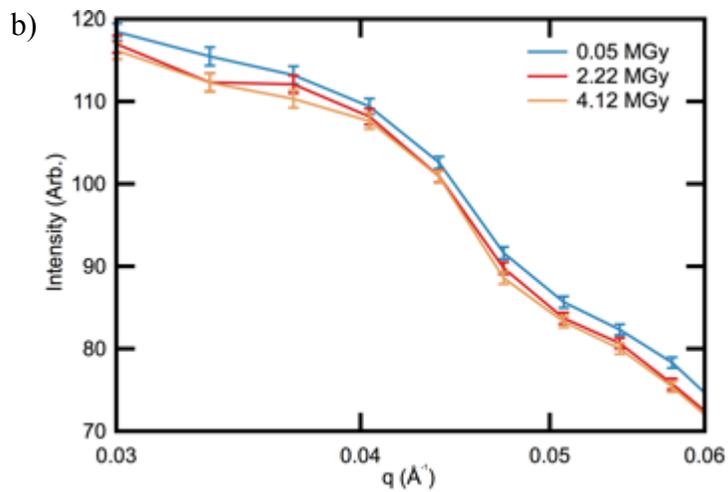
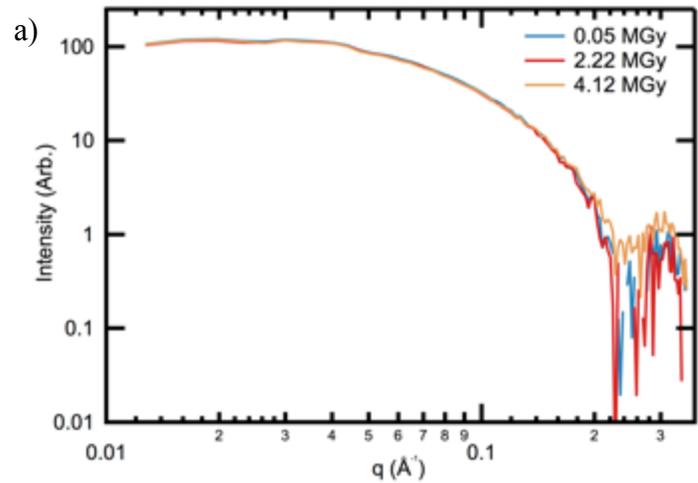


Figure 4.18 – Lysozyme damage at 100 K. a) Lysozyme scattering profiles as a function of dose at 100 K. A clear increase at high q is seen. b) This plot shows the same scattering profiles as a), from $q = 0.03$ to 0.06 . A clear decrease at low q is seen.

The lysozyme data were quantified in the same way as the glucose isomerase data. For $I_{nom,B}$, S_I is given in Table 4.8. The lysozyme sample in the 36% w/w PG buffer 2 was roughly 1.6 times more sensitive than the protein-free 36% w/w PG buffer 2. The damage rate at 100 K was also calculated for the standard subtracted intensity, I_{sub} , the calculated S_I is in Table 4.9. Lysozyme was half as sensitive as glucose isomerase using this metric.

For lysozyme, as for glucose isomerase, both negative and positive values were measured for S_{rg} and S_{mw} . As before, the average and standard deviation was calculated from the absolute values of the measured sensitivities. We found $S_{rg} = 1.8 \cdot 10^{-3}$ %/kGy with $\sigma_{rg} = 1.8 \cdot 10^{-3}$, and $S_{mw} = 1.0 \cdot 10^{-2}$ %/kGy with $\sigma_{mw} = 1.4 \cdot 10^{-2}$. The aggregation likely significantly increased the reported damage rate. If the initial R_g and molecular weight are used to normalize the data instead of 14.3 Å and 14.3 kDa, $S_{rg} = 2.6 \cdot 10^{-4}$ with $\sigma_{rg} = 1.8 \cdot 10^{-4}$ %/kGy and $S_{mw} = 1.8 \cdot 10^{-3}$ %/kGy with $\sigma_{mw} = 8.8 \cdot 10^{-4}$ %/kGy. These are the values listed in Table 4.10. The sensitivities are in relatively good agreement with those found for glucose isomerase: $S_{rg} = 1.5 \cdot 10^{-4}$ %/kGy and $S_{mw} = 1.7 \cdot 10^{-3}$ %/kGy.

It is not clear if it is better to treat the data relative to the measured initial values or the expected initial values (consistent with all of the other processing in the paper). The aggregation will also alter the other reported sensitivities, so any comparison between the 100 K and room temperature lysozyme samples or between 100 K samples and protein-free buffer is suspect. However, the sample does represent macromolecules in solution, so it is useful data for evaluating damage at 100 K, even though it is unlikely to be the true rate for an unaggregated lysozyme sample.

Table 4.8 – Radiation induced changes to lysozyme and protein-free buffer 2 at 100 K and room temperature. This table lists the average integrated intensity sensitivity, S_I , and standard deviation, σ_I , for $I_{nom,B}(q)$ (protein) and $I_B(q)$ (protein-free buffer). The protein is lysozyme in the PG-free and 36% w/w PG buffer 2. The buffers are the protein-free PG-free, 36%, and 50% w/w PG buffer 2. For lysozyme at 100 K, the ratio of protein to protein-free buffer sensitivity is 1.7. At 100 K, S_I for the buffer and protein have large standard deviations, so we cannot reliably state that the protein sample damages more slowly than the protein-free buffer.

Sample	S_I (%/kGy)	σ_I (%/kGy)
Lys. 36% PG (100 K)	$4.3 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$
Buffer 2, 36% PG (100 K)	$2.6 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$
Buffer 2, 50% PG (100 K)	$5.1 \cdot 10^{-4}$	$1.1 \cdot 10^{-4}$
Lys., PG-free (RT)	4.6	0.3
Lys., 36% PG (RT)	$1.3 \cdot 10^{-4}$	$0.6 \cdot 10^{-4}$
Buffer 2, PG-free (RT)	$5.6 \cdot 10^{-4}$	$2.5 \cdot 10^{-4}$
Buffer 2, 36% PG (RT)	1.24	0.51
Buffer 2, 50% PG (RT)	1.61	0.30

Table 4. 9 – Comparison of the integrated intensity sensitivity, S_I , for lysozyme at 100 K and room temperature using the standard subtracted intensity, I_{sub} . Note that due to the different processing method these values should not be comparable to those listed in Table 4.8.

Sample	S_I (%/kGy)	σ_I (%/kGy)	S_I ratio: RT/100 K
Lys., 36% PG (100 K)	$7.3 \cdot 10^{-4}$	$4.1 \cdot 10^{-4}$	-
Lys., PG-free (RT)	15.3	1.1	21000
Lys., 36% PG (RT)	1.7	0.8	2300

Table 4.10 – Comparison between S_{rg} and S_{mw} for lysozyme at 100 K and at room temperature in the PG-free and 36% w/w PG buffer 2s. The sensitivity ratios are given for the room temperature measurements to the 100 K measurement.

Sample	S_{rg} (%/kGy)	σ_{rg} (%/kGy)	S_{mw} (%/kGy)	σ_{mw} (%/kGy)	S_{rg} ratio: RT/100 K	S_{mw} ratio: RT/100 K
Lys., 36% PG (100 K)	$2.6 \cdot 10^{-4}$	$1.8 \cdot 10^{-5}$	$1.8 \cdot 10^{-3}$	$8.8 \cdot 10^{-4}$	-	-
Lys., PG-free (RT)	16.1	0.7	26.1	0.8	62000	15000
Lys., 36% PG (RT)	0.39	0.08	-0.09	0.04	1500	-50

4.7.3.2.2 Room temperature

Damage rates at room temperature were measured for lysozyme. Figure 4.19 shows scattering profiles vs. dose for lysozyme in the PG-free and 36% w/w PG buffer 2. The sample in the PG-free buffer 2 showed a very fast rise at high and mid q , and a drop at low q . The sample in the 36% w/w PG buffer 2 showed the same uptick at low q that was seen for glucose isomerase in the 36% w/w PG buffer 1. The same metrics were applied as for glucose isomerase, and the results for S_{rg} and S_{mw} are given in Table 4.10. In the 36% w/w PG buffer 2, S_{rg} was within a factor of 2 of that of glucose isomerase, while S_{mw} was 12 times less than that of glucose isomerase. However, in the PG-free buffer 2, S_{rg} was 1200 times greater and S_{mw} 240 times greater than the measured values glucose isomerase in the PG-free buffer 1. Additionally, for glucose isomerase adding 36% w/w PG to buffer 1 increased the damage rate while for lysozyme adding 36% w/w PG to buffer 2 decreased the damage rate. In order to offer another point of comparison to 100 K, the sensitivity was calculated for the standard subtracted intensity, I_{sub} , the S_I value is in Table 4.9.

The initial R_g in the PG-free buffer 2 ranged from 14.0-15.8 Å and the initial molecular weight from 9.5 to 13 kDa. The R_g is in good agreement with the expected 14.3 Å, while the molecular weight is slightly lower than the expected 14.3 kDa. In the 36% w/w PG buffer 2, initial R_g ranged from 14.3-15.2 Å and initial molecular weight from 7.4 to 14.1 kDa, indicating that the addition of the PG had no affect on the size of the lysozyme.

Figure 4.20 shows changes to the protein-free PG-free and 36% w/w PG buffer 2. These are very similar to the damage seen for the PG-free and 36% w/w PG buffer 1. The sensitivities for the buffers are reported in Table 4.2 (protein-free buffer 2 at room temperature). In order to compare

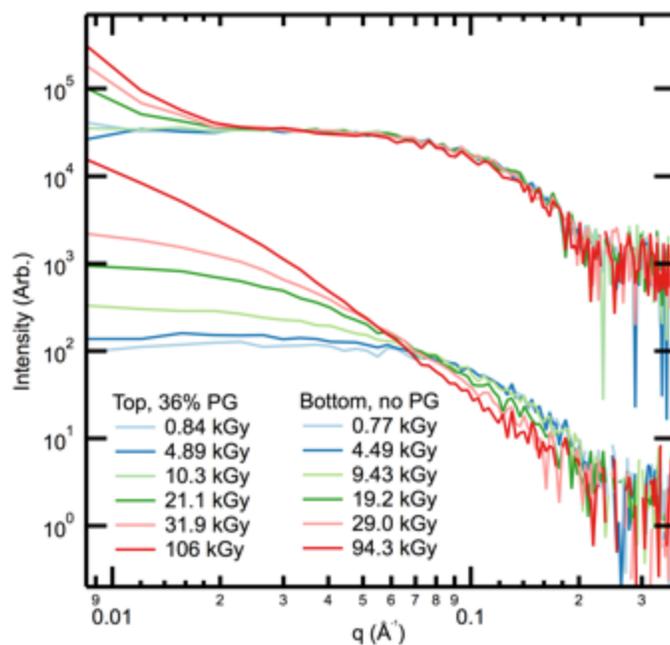


Figure 4.19 – Scattering profiles vs. dose for lysozyme at room temperature in 36% w/w PG buffer 2 (top) and PG-free buffer 2 (bottom). Damage happens much more quickly in the PG-free buffer 2 than in the 36% w/w PG buffer 2. The 36% w/w PG profiles were offset for clarity by multiplication by 250.

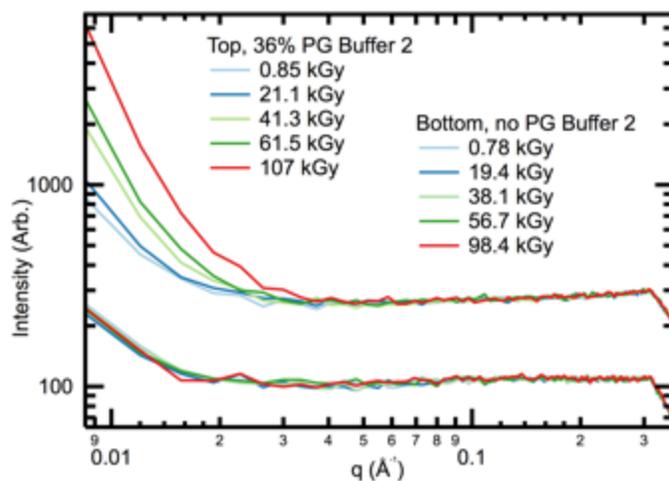


Figure 4.20 – Scattering profiles vs. dose for protein-free 36% w/w PG buffer 2 (top) and PG-free buffer 2 (bottom). There is no apparent change to the PG-free buffer 2, while the 36% w/w PG buffer 2 shows changes similar to that of the 36% w/w PG buffer 1 at room temperature. The 36% w/w PG buffer 2 profiles were offset for clarity by multiplication by 2.

the protein-free buffer 2 sensitivities with protein sensitivities, sensitivities were calculated for $I_{nom,B}(q)$, the calculated S_I values are given in Table 4.8. For the PG-free buffer 2, the protein solution damaged 8200 times faster than the protein-free buffer 2, while for the 36% w/w PG buffer 2 the protein solution damaged 3 times slower than the protein-free buffer 2.

4.7.4 Additional Discussion

4.7.4.1 Comparison between room temperature and 100 K rates

Because the samples are aggregated at 100 K, it was not clear that a comparison of the sensitivities at the two temperatures would yield meaningful results. The comparison was made anyways, but should be treated cautiously. The ratio of S_{rg} for lysozyme in the 36% w/w PG buffer 2 at room temperature to 100 K was 1500. The ratio of S_{rg} from lysozyme in the PG-free buffer 2 at room temperature to 36% w/w PG buffer 2 at 100 K was 62000. The sensitivity ratio for the identical lysozyme samples at the two temperatures was 2.9 times less than the ratio for glucose isomerase. However, due to the extremely large sensitivity of lysozyme at room temperature, the ratio from the PG-free buffer 2 to the 100 K sample was 840 times larger than that for glucose isomerase.

We also compared the molecular weight sensitivities at both temperatures. Again using the sensitivity calculated from the measured initial value, we found that for lysozyme in the 36% w/w PG buffer 2, S_{mw} decreased by a factor of 50 upon cryocooling. Comparing the sample at room temperature in the PG-free buffer 2 to the 100 K sample gives a decrease in sensitivity by a factor of 15000. The sensitivity ratio for protein in 36% w/w PG buffer 2 at different temperatures was much less than the factor of 650 found for glucose isomerase. The ratio of the

sensitivity of the PG-free sample at room temperature to the 100 K sample was much greater than the factor of 65 found for glucose isomerase.

Next we looked at the ratio of S_I calculated for I_{sub} . The ratio of S_I from 36% w/w PG buffer 2 at room temperature and 100 K was 2300. The ratio of S_I from non-PG buffer 2 at room temperature to 36% w/w PG buffer 2 at 100 K was 21000. Similar to S_{rg} , the ratio for identical samples at different temperatures was similar to that of glucose isomerase (2.4 times less), while the ratio from PG-free buffer 2 to the 100 K sample was much larger.

Lysozyme in the PG-free buffer 2 was much more sensitive than glucose isomerase at room temperature, while lysozyme and glucose isomerase at 100 K have similar sensitivities. This led to a much more dramatic decrease in radiation sensitivity upon cooling for lysozyme, from 15000-62000x, than for glucose isomerase, from 33-74x, when comparing lysozyme in the PG-free buffer 2 at room temperature to the 100 K samples. However, when the comparison was made for identical samples at room temperature, the decrease in radiation sensitivity for lysozyme is 50-2300x, while for glucose isomerase it is 650-5500x. The difference may be due to the aggregation of the lysozyme sample at 100 K, certainly that makes this comparison unreliable.

4.7.4.2 Comparison between samples and buffers and to previous results

For glucose isomerase at 100 K, the protein solution had a sensitivity 0.7 times that of the protein-free 36% w/w/ PG buffer 2. However, for lysozyme at 100 K the protein solution had a sensitivity 1.6 times that of the protein-free 36% w/w/ PG buffer 2. In both cases, the standard deviations of the measured sensitivities are large enough we cannot conclusively say if the protein was more, as, or less sensitive than the buffer.

At room temperature, it was observed that the lysozyme in the 36% w/w PG buffer 2 damaged 3 times slower than the protein-free 36% w/w PG buffer 2. This is interesting, as intuitively it seems the protein-free buffer sensitivity should be limiting. One possible explanation is that the lysozyme scavenged some of the PG macroradicals in solution, and that the reaction of those two species did not significantly alter the lysozyme structure. Alternatively, the lysozyme may have undergone damage such as fragmentation that reduced the scatter at low q , while the buffer was aggregating, increasing low q scatter. This is supported by the negative S_{mw} , but would not easily explain the positive S_{rg} , unless the protein is partly unfolding as well. This would reduce the observed rate of increase of the intensity for the protein solution.

For the lysozyme in the PG-free buffer 2 we found $S_{rg} = 16.1 \pm 0.7$ %/kGy and $S_{mw} = 26.1 \pm 0.8$ %/kGy. It has previously been reported that for 4.07 mg/mL lysozyme in a 40 mM sodium acetate pH 4.0 50 mM sodium chloride buffer $S_{rg} = 21.3 \pm 3.7$ %/kGy and $S_{mw} = 37.6 \pm 18.6$ %/kGy (Hopkins & Thorne, 2016). The S_{rg} values are similar, but do not agree within the reported standard deviations. The S_{mw} values agree within the reported standard deviations only because the standard deviation of the previously reported value is quite large. The most likely cause of the discrepancy between the values is the difference in the buffer composition.

4.7.5 Additional Conclusions

The lysozyme data presented in this section supports the conclusions drawn from the glucose isomerase data. We observe a range of critical doses for lysozyme at 100 K, depending on the metric used, from 0.6 to 3.8 MGy, compared with the 0.6-6.7 MGy range for glucose isomerase. Comparisons to the room temperature data are suspect, but for S_{rg} and S_I we see an increase in

the radiation tolerance by a factor of $\sim 1,900$ for the same sample at both temperatures, compared with the $\sim 5,000x$ increase found for glucose isomerase. For S_{mv} the increase in radiation tolerance is a factor of 50, compared with the factor of 650 for glucose isomerase. The increase is much greater when comparing lysozyme in PG-free buffer 2 at room temperature to lysozyme in 36% w/w PG buffer 2 at 100 K, $\sim 15,000$ - $62,000x$ gain in radiation tolerance upon cryocooling, depending on metric. Lysozyme is much more sensitive at room temperature than glucose isomerase, yet the sensitivities at 100 K were similar. This further supports the conclusions drawn in the main paper, as we measured similar 100 K damage rates and gains for both a very sensitive (albeit aggregated) protein and a very insensitive protein.

REFERENCES

- Alexander, P. & Charlesby, A. (1957). *J. Polym. Sci.* **XXIII**, 355–375.
- Audette-Stuart, M., Houée-Levin, C., & Potier, M. (2005). *Radiat. Phys. Chem.* **72**, 301–306.
- Awadelkarim, O., Claesson, O., & Lund, A. (1988). *Zeitschrift Für Naturforsch. A.* **43**, 633–637.
- Baker, L. a. & Rubinstein, J. L. (2010). *Radiation damage in electron cryomicroscopy* Elsevier Masson SAS.
- Barker, S. a, Brimacombe, J. S., & Eades, E. D. M. (1964). *Radiat. Res.* **22**, 357–367.
- Bednarek, J., Plonka, A., Hallbrucker, A., & Mayer, E. (1998a). *Radiat. Phys. Chem.* **53**, 635–638.
- Bednarek, J., Plonka, A., Hallbrucker, A., & Mayer, E. (1998b). *J. Phys. Chem.* **102**, 9091–9094.
- Bednarek, J., Plonka, A., Hallbrucker, A., & Mayer, E. (1999a). 6824–6828.
- Bednarek, J., Plonka, A., Hallbrucker, A., & Mayer, E. (1999b). *Radiat. Phys. Chem.* **55**, 477–481.
- Bednarek, J., Plonka, A., Hallbrucker, A., Mayer, E., & Symons, M. C. R. (1996). *J. Am. Chem. Soc.* **118**, 9387–9390.
- Belloni, J., Delcourt, M. O., Houee-Levin, C., & Mostafavi, M. (2000). *Annu. Reports Sect. “C” (Physical Chem.* **96**, 225–295.
- 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).
- Blanchet, C. E. & Svergun, D. I. (2013). *Annu. Rev. Phys. Chem.* **64**, 37–54.
- 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.
- Burmeister, W. P. (2000). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56**, 328–341.
- Buwalda, S. J., Boere, K. W. M., Dijkstra, P. J., Feijen, J., Vermonden, T., & Hennink, W. E. (2014). *J. Control. Release.* **190**, 254–273.
- Chaudhuri, B. N. (2015). *Protein Sci.* **24**, 267–276.
- Cole, a (1969). *Radiat. Res.* **38**, 7–33.
- 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.
- Dreiss, C. a., Jack, K. S., & Parker, A. P. (2006). *J. Appl. Crystallogr.* **39**, 32–38.
- Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowall, a W., & Schultz, P.

- (1988). *Q. Rev. Biophys.* **21**, 129–228.
- 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.
- Filali-Mouhim, A., Audette, M., St-Louis, M., Thauvette, L., Denoroy, L., Penin, F., Chen, X., Rouleau, N., Le Caer, J.-P., Rossier, J., et al. (1997). *Int. J. Radiat. Biol.* **72**, 63–70.
- Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I., & Craievich, A. F. (2009). *J. Appl. Crystallogr.* **43**, 101–109.
- Fischetti, R., Stepanov, S., Rosenbaum, G., Barrea, R., Black, E., Gore, D., Heurich, R., Kondrashkina, E., Kropf, a J., Wang, S., et al. (2004). *J. Synchrotron Radiat.* **11**, 399–405.
- Fletcher, J. W. & Freeman, G. R. (1967). *Can. J. Chem.* **45**, 635–640.
- Garman, E. F. & Owen, R. L. (2006). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 32–47.
- Garman, E. F. & Weik, M. (2015). *J. Synchrotron Radiat.* **22**, 195–200.
- 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.
- 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.
- Hase, H. & Ishioka, K. (1992). *Int. J. Radiat. Appl. Instrumentation. Part C. Radiat. Phys. Chem.* **39**, 329–331.
- Holton, J. M. (2009). *J. Synchrotron Radiat.* **16**, 133–142.
- Hopkins, J. B., Katz, A. M., Meisburger, S. P., Warkentin, M. A., Thorne, R. E., & Pollack, L. (2015). *J. Appl. Crystallogr.* **48**, 227–237.
- Hopkins, J. B. & Thorne, R. E. (2016). *Submitted*.
- Houée-Levin, C. & Sicard-Roselli, C. (2001). *Studies in Physical and Theoretical Chemistry v87: Radiation Chemistry - Present Status and Future Trends*, C.D. Jonah, & B.S.M. Rao, edited by, pp. 553–584. New York: Elsevier Scientific Pub.
- Howells, M. R., Beetz, T., Chapman, H. N., Cui, C., Holton, J. M., Jacobsen, C. J., Kirz, J., Lima, E., Marchesini, S., Miao, H., et al. (2009). *J. Electron Spectros. Relat. Phenomena.* **170**, 4–12.
- Huang, T. C., Toraya, H., Blanton, T. N., & Wu, Y. (1993). *J. Appl. Crystallogr.* **26**, 180–184.
- Ichikawa, T. & Yoshida, H. (1996). *J. Phys. Chem.* **100**, 9385–9388.
- Jacques, D. & Trewhella, J. (2010). *Protein Sci.* 642–657.
- Jeffries, C. M., Graewert, M. A., Svergun, D. I., & Blanchet, C. E. (2015). *J. Synchrotron Radiat.* **22**, 273–279.
- Kempner, E. S. (1993). *Q. Rev. Biophys.* **26**, 27–48.

- Kempner, E. S. (2011). *J. Polym. Sci. B. Polym. Phys.* **49**, 827–831.
- Keys, K. B., Andreopoulos, F. M., & Peppas, N. A. (1998). 8149–8156.
- Khattab, I. S., Bandarkar, F., Khoubnasabjafari, M., & Jouyban, A. (2012). *Arab. J. Chem.*
- Kirby, N. M. & Cowieson, N. P. (2014). *Curr. Opin. Struct. Biol.* **28**, 41–46.
- Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., & Svergun, D. I. (2003). *J. Appl. Crystallogr.* **36**, 1277–1282.
- Kozak, M. (2005). *J. Appl. Crystallogr.* **38**, 555–558.
- Kroh, J. & Plonka, A. (1977). *Chem. Phys. Lett.* **52**, 371–373.
- Kuwamoto, S., Akiyama, S., & Fujisawa, T. (2004). *J. Synchrotron Radiat.* **11**, 462–468.
- Leapman, R. D. & Sun, S. (1995). *Ultramicroscopy.* **59**, 71–79.
- Lund, A., Nilsson, G., & Samskog, P. (1986). *Radiat. Phys. Chem. Int. J. Radiat. Appl. Instrumentation. Part C.* **27**, 111–121.
- Makarov, I. E., Ershov, B. G., & Pikaev, A. K. (1969). *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **18**, 2024–2029.
- Meents, A., Gutmann, S., Wagner, A., & Schulze-Briese, C. (2010). *Proc. Natl. Acad. Sci. U. S. A.* **107**, 1094–1099.
- Meisburger, S. P., Warkentin, M., Chen, H., Hopkins, J. B., Gillilan, R. E., Pollack, L., & Thorne, R. E. (2013). *Biophys. J.* **104**, 227–236.
- Miller, J. H., Fedoronko, D. a, Hass, B. D., Myint, M., & Kempner, E. S. (1998). *Arch. Biochem. Biophys.* **352**, 281–287.
- Mylonas, E. & Svergun, D. I. (2007). *J. Appl. Crystallogr.* **40**, s245–s249.
- Nave, C. & Hill, M. A. (2005). *J. Synchrotron Radiat.* **12**, 299–303.
- 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.
- Orthaber, D., Bergmann, A., & Glatter, O. (2000). *J. Appl. Crystallogr.* **33**, 218–225.
- Petoukhov, M. V, Konarev, P. V, Kikhney, A. G., & Svergun, D. I. (2007). *J. Appl. Crystallogr.* **40**, s223–s228.
- Petoukhov, M. V & Svergun, D. I. (2013). *Int. J. Biochem. Cell Biol.* **45**, 429–437.
- 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.
- Petrik, N. G. & Kimmel, G. a. (2004). *J. Chem. Phys.* **121**, 3736–3744.
- Pikaev, S. K., Ershov, B. G., & Puntezhis, S. S. (1970). *Radiat. Eff. Inc. Plasma Sci. Plasma Technol.* **5**, 265–268.

- PŁonka, A., Szajdzińska-Piętek, E., & Kroh, J. (1984). *Radiat. Phys. Chem.* **23**, 583–587.
- Pollack, L. (2011). *Biopolymers*. **95**, 543–549.
- Rambo, R. P. & Tainer, J. A. (2013). *Nature*. **496**, 477–481.
- Ravelli, R. B. & McSweeney, S. M. (2000). *Structure*. **8**, 315–328.
- Riederer, H., Hüttermann, J., Boon, P., & Symons, M. C. . (1983). *J. Magn. Reson.* **54**, 54–66.
- Rosiak, J. M. & Ulański, P. (1999). *Radiat. Phys. Chem.* **55**, 139–151.
- Saha, A., Mandal, P. C., & Bhattacharyya, S. N. (1995). *Radiat. Phys. Chem.* **46**, 123–145.
- Shalaev, E., Reddy, R., Kimball, R. ., Weinschenk, M. ., Guinn, M., & Margulis, L. (2003). *Radiat. Phys. Chem.* **66**, 237–245.
- Sharpatyi, V. A. (1963). *Russ. Chem. Rev.* **32**, 321–330.
- Shen, C., Julius, E. F., Tyree, T. J., Moreau, D. W., Atakisi, H., & Thorne, R. E. (2016). *Submitted*.
- Singh, A. & Singh, H. (1982). *Prog. Biophys. Mol. Biol.* **39**, 69–107.
- Skou, S., Gillilan, R. E., & Ando, N. (2014). *Nat. Protoc.* **9**, 1727–1739.
- Stadtman, E. R. & Levine, R. L. (2003). *Amino Acids*. **25**, 207–218.
- Staluszka, J., Plonka, A., Szajdzinska-Pietek, E., Kohl, I., Hallbrucker, A., & Mayer, E. (2003). *Radiat. Phys. Chem.* **67**, 247–250.
- Staluszka, J., Steblecka, M., Szajdzinska-Pietek, E., Kohl, I., Salzmann, C. G., Hallbrucker, A., & Mayer, E. (2008). *J. Phys. Chem. A*. **112**, 8678–8685.
- Svergun, D., Barberato, C., & Koch, M. H. J. (1995). *J. Appl. Crystallogr.* **28**, 768–773.
- Svergun, D. I. & Koch, M. H. J. (2003). *Reports Prog. Phys.* **66**, 1735–1782.
- Terryn, H., Deridder, V., Sicard-Roselli, C., Tilquin, B., & Houée-Levin, C. (2005). *J. Synchrotron Radiat.* **12**, 292–298.
- Vestergaard, B. & Sayers, Z. (2014). *IUCrJ*. **1**, 523–529.
- Wang, B., Kodama, M., Mukataka, S., & Kokufuta, E. (1998). *Polym. Gels Networks*. **6**, 71–81.
- Warkentin, M., Badeau, R., Hopkins, J. B., & Thorne, R. E. (2012). *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 1108–1117.
- Warkentin, M., Badeau, R., Hopkins, J., & Thorne, R. E. (2011). *Acta Crystallogr. D Biol. Crystallogr.* **67**, 792–803.
- Warkentin, M., Berejnov, V., Hussein, N. S., & Thorne, R. E. (2006). *J. Appl. Crystallogr.* **39**, 805–811.
- Warkentin, M., Hopkins, J. B., Badeau, R., Mulichak, A. M., Keefe, L. J., & Thorne, R. E. (2013). *J. Synchrotron Radiat.* **20**, 7–13.
- Warkentin, M., Stanislavskaia, V., Hammes, K., & Thorne, R. E. (2008). *J. Appl. Crystallogr.*

41, 791–797.

Weik, M., Ravelli, R. B., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J., & Sussman, J. L. (2000). *Proc. Natl. Acad. Sci. U. S. A.* **97**, 623–628.

Wenxiu, C., Huaying, B., & Manwei, Z. (1985). *Radiat. Phys. Chem.* **26**, 43–47.

Zakurdaeva, O. a., Nesterov, S. V., & Feldman, V. I. (2005). *High Energy Chem.* **39**, 201–206.

Zakurdaeva, O. a., Nesterov, S. V., & Feldman, V. I. (2010). *Radiat. Phys. Chem.* **79**, 876–879.

Zheng, W., Jewitt, D., & Kaiser, R. I. (2006). *Astrophys. J.* **639**, 534–548.

Zheng, W., Jewitt, D., & Kaiser, R. I. (2007). *Chem. Phys. Lett.* **435**, 289–294.

CHAPTER 5

EFFECT OF COMMON CRYOPROTECTANTS ON CRITICAL WARMING RATES AND ICE FORMATION IN AQUEOUS SOLUTIONS

5.1 Preface

This chapter is somewhat of a departure from the previous work, as it has no direct relation to the cryoSAXS project. The Thorne group has been working to understanding the ice formation behavior in aqueous solutions, as it has practical applications to cryopreservation and protein crystallography (Kriminski et al., 2003; Berejnov et al., 2006; Warkentin et al., 2006, 2008; Warkentin & Thorne, 2009) and provides insight into fundamental physics (Warkentin et al., 2013). This expertise, along with our study of radiation damage in protein crystallography, is what lead us to consider the idea of cryoSAXS. Because of that, this work is indirectly related to the cryoSAXS project.

The rest of this Chapter is the text of a paper which reported on the ‘critical warming rates’ of aqueous solutions (Hopkins et al., 2012). Critical warming rates are the rates necessary to prevent ice formation in aqueous solutions upon warming from a vitrified state to liquid state (100 K to 300 K, roughly). These rates are compared to the cooling rates needed to prevent ice formation in the same solutions, and the differences are understood in terms of the fundamental physics of ice formation and growth as a function of temperature in these solutions.

5.2 Introduction

How can biological samples be cooled, stored and then warmed without damage? Most studies of cryopreservation have focused on the cooling process and its consequences. Recent work suggests that the warming process is of comparable or greater importance in preserving biological function (Seki & Mazur, 2008, 2009; Mazur & Seki, 2011). Significant intracellular ice nucleation and growth can be lethal, and can occur on both cooling and warming (Zhmakin, 2008; Seki & Mazur, 2009). Thus it is important for successful cryopreservation to minimize the formation of intracellular ice.

Pure water can be supercooled to the homogeneous ice nucleation temperature $T_h = 235$ K (-38° C) (Debenedetti, 2003), and can be vitrified by cooling below its glass transition temperature $T_g = 136$ K (-137° C) (Debenedetti, 2003) at rates of order 10^6 K/s (Brüggeller & Mayer, 1980; Kohl et al., 2005). Upon warming, spontaneous ice nucleation and growth occurs above ~ 150 K (-123° C) (Debenedetti, 2003), followed by recrystallization and then melting. Ice volume increases both by nucleation and growth of new ice nuclei and growth of ice formed during cooling. Cryoprotectants suppress ice nucleation and reduce growth rates, and so reduce cooling rates required for sample vitrification (Fuller, 2004) and warming rates to minimize ice formation during warming and sample recovery.

A large number of studies have examined how critical cooling rates (CCRs) required for vitrification (i.e., to keep ice volume fraction formed on cooling below a minimum detectable (small) value) vary with cryoprotectant concentration and type in aqueous solutions and in cells, most at cooling rates below 10 K/s. Experimental study of critical warming rates is more challenging. Ice forms much more readily on warming, so critical warming rates (CWRs) of a given solution are much larger than critical cooling rates. Moreover, ice formation on warming to

room or biological temperature is transient. The slow warming rates characteristic of calorimetric methods used in previous studies have generally been insufficient to measure critical warming rates for moderate concentrations of most cryoprotectants (Boutron & Mehl, 1990). Critical warming rate data have only been reported for a few cryoprotectants (DMSO, levo-2,3-butanediol, propane-1,2-diol, propylene glycol, propylene glycol monomethyl ether, DL-theritol, and 1,4-Di-O-methyl-DL-theritol), and only for critical warming rates below ~ 8.3 K/s (Boutron & Kaufmann, 1978; Boutron, 1990; Boutron & Mehl, 1990; Wowk et al., 1999; Baudot et al., 2000).

Vitrification procedures are becoming increasingly accepted in cryopreservation, especially of single cells like oocytes and embryos (Pereira & Marques, 2008; Gook, 2010). Because cooling is rapid, little water transport across the cell membrane (driven in part by ice formation in the extracellular medium) and thus little intracellular dehydration can occur (Fowler & Toner, 2005). Conventional vitrification techniques employing moderate cooling rates thus require substantial cryoprotectant concentrations to prevent intracellular ice formation. Ultra-rapid vitrification techniques, in which very small volume ($<0.1 \mu\text{l}$) samples are plunge cooled in liquid nitrogen and plunge warmed in aqueous solutions, give much larger cooling and warming rates and thus reduce required cryoprotectant concentrations (Seki & Mazur, 2008, 2009). Recent work on mouse oocytes found that survival increased with warming rate up to the maximum examined rate of ~ 2000 K/s (Mazur & Seki, 2011).

Here we report measurements of the critical warming rates required to prevent ice formation in aqueous solutions of common cryoprotectants. Our data for glycerol, DMSO, ethylene glycol, polyethylene glycol 200 and sucrose cover warming rates between ~ 10 and $\sim 10^4$ K/s, the smallest of these rates being larger than the largest rate measured in previous studies. We find a

strong dependence of critical warming rate on cooling rate, with larger cooling rates giving smaller critical warming rates and allowing the use of smaller cryoprotectant concentrations. We describe the apparatus developed for achieving and measuring these rates, and compare our CWR results with predictions of models for ice nucleation and growth (MacFarlane & Fragoulis, 1986; Toner et al., 1990; Kresin et al., 1991; Karlsson et al., 1993, 1994, 1996; Karlsson, 2001, 2010). We also report measurements of critical cooling rates for DMSO, ethylene glycol and polyethylene glycol 200.

5.3 Materials and Methods

5.3.1 Materials

All cryoprotectants studied were obtained either from Sigma-Aldrich (St. Louis, MO) or Avantor Performance Materials (Phillipsburg, NJ) and were 99+% pure. Each was mixed weight per weight (w/w) with 16.7 M Ω -cm DI water filtered in a Barnstead Nanopure II system (Thermo Scientific, Dubuque, IA). Weighing errors corresponded to less than $\pm 0.2\%$ w/w. In the case of sucrose near its solubility limit, solutions were heated and sonicated to ensure full mixing.

5.3.2 Experimental apparatus for warming

Figure 5.1 shows a schematic illustration of the apparatus developed for measuring critical warming rates. Samples were cooled either by insertion into a cold nitrogen gas stream or by plunging in liquid nitrogen, and were warmed by immersion in hot oil. Ice formation was assessed by a visual assay using a microscope and high-speed camera. A detailed description of each component follows.

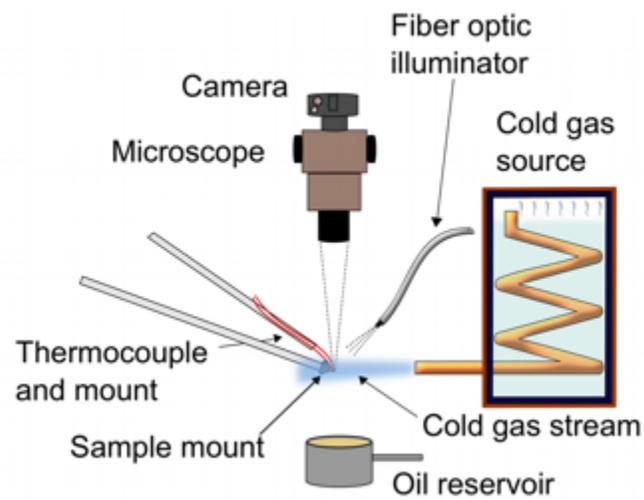


Figure 5.1 – Schematic illustration of the experimental setup used to measure critical warming rates.

Cold gas stream. A $T \sim 100$ K (-173 °C) cold gas stream was produced using a coiled copper tube immersed in liquid nitrogen (LN_2) contained in a sealed storage tank. Nitrogen gas boil-off within the tank entered the coil at the top of the tank, and exited the coil outside the tank. With a flow rate of ~ 0.5 l/s, controlled by environmental heat input to the storage tank, the exit flow was laminar. After filling the tank, the measured gas stream temperature increased at ~ 0.001 K/s as nitrogen boiled off, from ~ 94 K (-179 °C) to a maximum of ~ 103 K (-170 °C), well below water's glass transition temperature. Heaters were incorporated in early versions, but were later deemed unnecessary because precise temperature control was unimportant.

Oil immersion tank. The oil immersion tank consisted of a 60 ml stainless steel cup filled with Light Mineral Oil (Fisher Scientific, Waltham, MA). The oil temperature determined the maximum warming rates, and so the oil was warmed on a hotplate from 295 K (22 °C) up to 450 K (180 °C). Samples were kept cold in the gas stream, and were abruptly warmed by manually lifting the cup up through the cold gas stream until the sample was fully immersed in oil. With a ~ 0.5 m/s vertical cup velocity (determined from high speed imaging) and a cup filled to its brim, the transition time between cold stream cooling and full oil immersion was at most a few ms. This short transition time prevented condensation and ice formation on the sample from the surrounding humid air.

Sample holder. Samples were held either in microfabricated polyimide loops with circular openings of 20 to 300 μm , attached to 0.64 mm diameter stainless steel rods (MicroLoops LD, MiTeGen Ithaca, NY) or in loops of 100 μm diameter polyester thread with nominal loop diameters from ~ 0.3 to 5 mm, glued into 0.64 mm diameter hollow stainless steel rods (MicroTubes, Hampton Research, Aliso Viejo, CA), as shown in Figure 5.2a and 5.2b. The rods were inserted into a magnetic steel base (Reusable Goniometer Base, MiTeGen), which was held

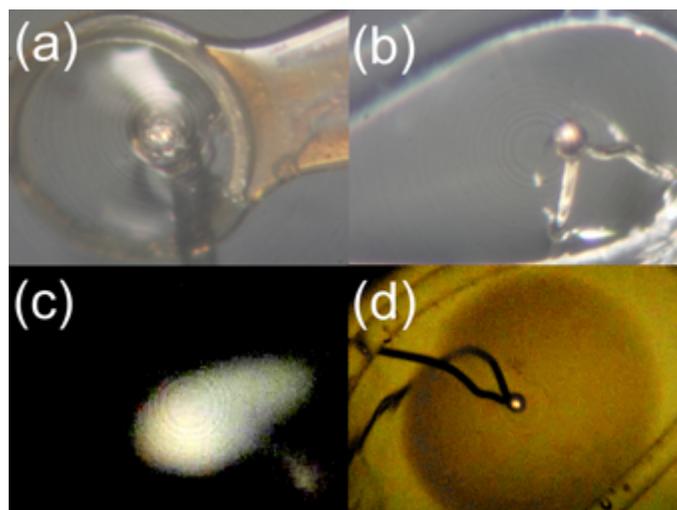


Figure 5.2 – (a) Drop of water-glycerol solution mounted in a 300 μm polyimide loop (MicroLoop LD, MiTeGen) with thermocouple. (b) Drop mounted in a custom loop with thermocouple. (c) Dark field optical ice assay. (d) Bright field optical ice assay.

by a magnet to the end of a 9 mm diameter glass (thermally insulating) rod. This rod was attached to an x-y manual translation stage (Thorlabs, Newton, NJ), allowing fine positioning of the sample in the cold gas stream and in the microscope's field of view.

Thermocouple holder. To measure the temperature of loop-mounted samples, bare-wire type-E thermocouples with standard connectors (OMEGA Engineering, Stamford, CT) were attached to the end of a glass rod. The rod was mounted on an x-y-z manual translation stage (Thorlabs) to facilitate insertion and fine positioning of the thermocouple within the sample. Thermocouple wire diameters ranged from 0.003" (75 μm) to 0.0005" (12.5 μm), with a typical diameter of 0.001" (25 μm). For each thermocouple size, the warming rate of the bare junction when immersed in hot oil was determined. For each sample volume, a thermocouple size was then chosen that would not limit the measured sample warming rate. Thermocouple voltages were measured and digitized using a 6052E DAQ card (National Instruments, Austin, TX), at a sampling rate of 250 kHz.

Sample imaging. Samples were imaged during cooling and warming using a Casio EX-F1 high speed camera (Casio, Dover, NJ), which provided frame rates of 24, 30, 60, 300, 600 and 1200 frames per second. This camera was attached to a Stemi 2000-C microscope (Carl Zeiss MicroImaging, Thornwood, NJ), with primary illumination provided by fiber optics. Still images and high speed videos were acquired in either bright field mode (with diffuse illumination scattered from a white background behind the sample) or dark field mode (looking for scatter from ice with a black background behind the drop), as shown in Figure 5.2c and 5.2d.

5.3.3 Warming rate data collection

Aqueous cryoprotectant solutions with volumes ranging from ~50 nL to 10 μ L were dispensed into appropriate diameter loops. Most samples were cooled in a $T \sim 100$ K nitrogen gas stream, which gave cooling rates from ~5 K/s to ~1200 K/s, depending upon sample volume. Dehydration was negligible in all cases, because water's vapor pressure rapidly decreases as it cools (dropping by a factor of 4 between 20 $^{\circ}$ C and 0 $^{\circ}$ C and another factor of 6 (in the supercooled state) from 0 $^{\circ}$ C to -20 $^{\circ}$ C (Kraust & Greer, 1984)), because the addition of cryoprotectants further reduces water's activity, and because the slowest cooling rates were achieved using the largest samples and containing the largest cryoprotectant concentrations. Samples were warmed by immersion in oil, at rates ranging from ~5 K/s to ~30,000 K/s. Cooling and warming rates are determined by the sample volume and shape, by the cooling or warming medium and its temperature, and by the sample's speed relative to the cooling medium. In practice, sample size produces larger variations in cooling/warming rates than choice of cooling/warming medium, temperature and flow/sample speed (Kriminski et al., 2003). Consequently, cooling and warming rates tend to be strongly coupled at rates above ~100 K/s. Adjusting the other experimental parameters to obtain significant independent variation is difficult except at very low (<10 K/s) rates.

To study the effects of cooling rate on critical warming rates, a method for cooling samples of similar size at rates spanning a few orders of magnitude was required. Smaller cooling rates were obtained by enclosing loop mounted samples in plastic capillary tubes sealed at one end and covered with varying thicknesses of spray foam insulation. For ~0.25 μ L samples, this method gave cooling rates in the nitrogen gas stream as small as 5 K/s. Prior to warming, the capillary could be removed while the sample was held at 100 K (-173 $^{\circ}$ C) in the nitrogen gas stream, to

maximize the warming rate. Warming rates for these samples varied from ~ 200 K/s to ~ 5000 K/s (compared with a maximum rate of $\sim 30,000$ K/s achieved for our smallest samples). The largest cooling rates were achieved by plunge cooling loop mounted samples into liquid nitrogen, following the “hyperquenching” technique described in reference (Warkentin et al., 2006) which involves blowing away the cold gas layer above the liquid nitrogen immediately prior to plunging.

Ice formation was assayed by judging sample transparency/opacity (Berejnov et al., 2006), as shown in Figure 5.2c and 5.2d. X-ray diffraction measurements (Warkentin et al., 2008) suggest that for $\sim \mu\text{l}$ volume drops the assay has a minimum detectible ice volume fraction of ~ 0.01 (1%). Critical cooling rates determined using this assay agree with previous literature values determined by other methods (Warkentin et al., 2008). On cooling, ice formation could be assessed from high resolution still images acquired after cooling was complete, as there was no subsequent sample evolution. But any ice formed on warming disappears when the sample undergoes final melting, and so transient ice was detected by frame-by-frame analysis of high-speed videos.

Warming (cooling) rates were quantified as the average warming (cooling) rate between $T=150$ K (-123 °C) and 240 K (-33 °C), roughly corresponding to the glass transition temperature and the homogeneous nucleation temperature of pure water (Debenedetti, 2003), respectively. Critical warming rates were calculated as the average of the fastest rate that produced detectable ice and the slowest rate that did not. An uncertainty in this value was calculated as half the difference between the two bounding rates. Error and irreproducibility in temperature and rate measurements could result from sample-to-sample variations in the position of the thermocouple within each sample. To quantify any such effect, dual thermocouple probes were constructed to

allow simultaneous temperature measurement at the center and edge of a sample. Cooling and warming rates measured for a range of sample volumes using the two thermocouples differed on average by 1% and at most by 10%.

5.3.4 Critical cooling rates

Critical cooling rates were measured for aqueous solutions of DMSO, ethylene glycol, and PEG 200, at lower concentrations and faster cooling rates than have been previously examined. These measurements were performed using the method described in (Warkentin et al., 2008), by plunging solution- and thermocouple- containing capillaries of different diameters into liquid nitrogen. Using capillaries instead of loops simplified thermocouple positioning and temperature measurements, but limited maximum achievable cooling rates.

5.3.5 Theoretical modeling

Several models based upon classical theories of crystal nucleation and growth have been developed to explain ice formation in aqueous solutions and in cells as a function of temperature (MacFarlane & Fragoulis, 1986; Toner et al., 1990; Kresin et al., 1991; Karlsson et al., 1993, 1994, 1996; Karlsson, 2010). Karlsson (Karlsson, 2010) recently compared these models, and showed that the model of Karlsson, Cravalho and Toner (Karlsson et al., 1993, 1994, 1996) (hereafter referred to as the KCT model) most closely matches experimental results for melting point depression, nucleation point depression and time to nucleation in NaCl (0-6 mol/kg) and glycerol (40% w/w) solutions (Karlsson, 2010). The KCT model has also been applied to describe ice formation upon warming (Karlsson, 2001).

The KCT model was developed to describe cells, and includes a model for transport of water across the cell membrane and intracellular dehydration in response to extracellular ice formation. To analyze our data for cell-free aqueous solutions, we use the ice nucleation and growth components of the KCT model, which build upon earlier analyses of nucleation (Toner et al., 1990) and growth (Christian, 1975; MacFarlane & Fragoulis, 1986). We summarize below these model components, and then use glycerol-specific parameters to model glycerol critical cooling and warming rate data. Unless otherwise noted, all temperatures are in K and, to facilitate connection to experiment, all concentrations are in units of weight per weight ($0 < c < 1$).

The total ice fraction formed during warming is obtained by summing contributions from ice nucleation and growth of those nuclei and from growth of pre-existing nuclei formed during cooling. The homogeneous ice nucleation rate, J , is given by (Toner et al., 1990; Karlsson, 2010)

$$J(c, T) = \Omega(c, T) \exp[-\kappa(c, T)\tau(c, T)] , \quad (5.1)$$

where c is the solute (in our case, cryoprotectant) concentration, T is the temperature, $\Omega(c, T)$ and $\kappa(c, T)$ are the kinetic and thermodynamic homogeneous nucleation rate coefficients (Toner et al., 1990) as defined in (Karlsson, 2010), and $\tau(c, T)$ is a dimensionless temperature (Karlsson et al., 1994).

Following (Karlsson et al., 1994), ice growth is described using a non-isothermal diffusion-limited growth model (MacFarlane & Fragoulis, 1986). The radius at a time t of a crystal nucleated at time t' is (MacFarlane & Fragoulis, 1986)

$$r(t; t') = \left[\int_{T(t')}^{T(t)} \alpha(T)^2 \bar{D}(c, T) \left(\frac{dT}{q} \right) \right]^{1/2} , \quad (5.2)$$

where $\alpha(T)$ is a dimensionless crystal growth parameter (of order unity), $\bar{D}(c, T)$ is an effective diffusion coefficient as defined in Eq. (5.9) below, and q is the cooling or warming rate. Assuming spherical ice growth and that growing ice fronts do not run into each other (no impingement), the extended volume fraction of ice formed on cooling is given by

$$\frac{V_{ice,c}(c, q_c)}{V_0} = \left(\frac{4\pi}{3} \right) \int_{T_i}^{T_f} J(c, T) \left[\int_T^{T_f} \alpha(T')^2 \bar{D}(c, T') \left(\frac{dT'}{q_c} \right) \right]^{3/2} \left(\frac{dT}{q_c} \right), \quad (5.3)$$

where V_0 is the sample volume, $q_c < 0$ is the cooling rate, T_i is the initial temperature and T_f is the (lower) final temperature. This calculated volume fraction overestimates the actual ice fraction at larger fractions when the probability of ice fronts intersecting is much larger.

Similarly, the extended ice volume fraction for warming is

$$\begin{aligned} \frac{V_{ice,w}(c, q_c, q_w)}{V_0} = & \left(\frac{4\pi}{3} \right) \int_{T_i}^{T_f} J(c, T) \left[\int_T^{T_f} \alpha(T')^2 \bar{D}(c, T') \left(\frac{dT'}{q_c} \right) \right]^{3/2} + \left[\int_{T_f}^{T_i} \alpha(T')^2 \bar{D}(c, T') \left(\frac{dT'}{q_w} \right) \right]^{3/2} \left(\frac{dT}{q_c} \right), \quad (5.4) \\ & + \left(\frac{4\pi}{3} \right) \int_{T_f}^{T_i} J(c, T) \left[\int_T^{T_i} \alpha(T')^2 \bar{D}(c, T') \left(\frac{dT'}{q_w} \right) \right]^{3/2} \left(\frac{dT}{q_w} \right) \end{aligned}$$

where $q_w > 0$ is the warming rate. The first term — an integral from warm T_i to cold T_f — gives the ice fraction due to crystal nucleation and growth on cooling and then subsequent additional growth of those crystals on warming. The last term — an integral from cold T_f to warm T_i — gives the ice fraction due to ice nucleation on warming and the subsequent growth of those nuclei (Karlsson, 2001). The actual (experimental) volume ice fraction χ , which includes the effects of (soft and hard) impingement of ice fronts during growth, is conventionally estimated (Karlsson et al., 1994) as

$$\chi = 1 - \exp[-V_{ice,w} / V_0], \quad (5.5)$$

χ is smaller than $V_{ice,w} / V_0$, but approaches $V_{ice,w} / V_0$ when $V_{ice,w} / V_0$ is small.

The model requires as input a function for the concentration- and temperature-dependent solution viscosity, the concentration-dependent homogeneous nucleation temperature and melting temperature, and the concentration- and temperature-dependent diffusion coefficient. Karlsson (Karlsson, 2010) provides a parameterization of the viscosity for DMSO and glycerol and for the melting and nucleation temperatures for a general aqueous solution. For glycerol, the viscosity used in (Karlsson, 2010) differs from measurements given in (Trejo González et al., 2011), and the melting temperature in (Karlsson, 2010) disagrees with measured values, particularly at high concentration (Bohon & Conway, 1972). We have used the concentration and temperature dependent viscosity model from (Trejo González et al., 2011) and a fit to $T_m(c)$ from (Bohon & Conway, 1972) given by

$$T_m(c) = -(188.25 \text{ K})c^3 + (64.633 \text{ K})c^2 - (30.21 \text{ K})c + 273.15 \text{ K}, \quad (5.6)$$

In (Karlsson et al., 1994), the diffusion constant is evaluated using the Stokes-Einstein equation,

$$D_{SE}(c, T) = \frac{k_B T}{6\pi a_0 \eta(c, T)}, \quad (5.7)$$

where $\eta(c, T)$ is the solution viscosity, k_B is the Boltzmann constant, and a_0 is the hydrodynamic radius of water. Recent experimental evidence suggests that nanoconfined water, protein hydration water, bulk water, and bulk water with solutes violate the Stokes-Einstein relation at low temperatures (Chen et al., 2006; Banerjee et al., 2009; Mallamace et al., 2010). The effects of solutes on the diffusion constant and non- Stokes-Einstein behavior have not yet been systematically evaluated. We thus calculate the diffusion constant of the solution as

$$D(c, T) = \left(\frac{\eta_w(T)}{\eta(c, T)} \right) D_w(c, T), \quad (5.8)$$

where η is the solution viscosity (Trejo González et al., 2011), η_w is the viscosity of water (Trejo González et al., 2011), and $D_w(c,T)$ is the diffusion constant of water as measured in (Chen et al., 2006).

Finally, for diffusion-limited growth, water near ice crystals is depleted relative to the bulk solution. Following (Karlsson et al., 1994), we average the diffusion coefficient of Eq. (5.8) between the edge of the ice crystal and a point in the bulk solution to find the effective diffusion coefficient,

$$\bar{D}(c,T) = \frac{1}{c_{w\infty,m} - c_{wliq,m}} \int_{c_{wliq,m}}^{c_{w\infty,m}} D(1 - c_w, T) dc_{w,m}, \quad (5.9)$$

where $c_{w\infty}$ is the water concentration far from a crystal, $c_{wliq} < c_{w\infty}$ is the liquidus water concentration, $c_w = 1 - c$, and the “m” subscripts denote concentrations in mol/L (Karlsson et al., 1994); the conversion from w/w to mol/L is given by

$$c_m = \frac{1000\rho_2}{M_1(1/c - (1 - \rho_2/\rho_1))}, \quad (5.10)$$

where $c = m_1/(m_1 + m_2)$ is the concentration in w/w, m_1 and ρ_1 (m_2 and ρ_2) are the mass and density of the first (second) component of the solution in g/cm³, and M_1 is the molar mass of the first component in mol/g. The liquidus concentration is the water concentration at the ice-liquid boundary, where the solution melting point is equal to the temperature. The equilibrium water-glycerol phase diagram has a eutectic point at 66.7 % w/w glycerol and 226.5 K (-46.6° C), but at higher glycerol concentrations the observed crystallization is of glycerol, not water. Thus, following (Bohon & Conway, 1972), the water melting curve is simply extrapolated to higher concentration and lower temperatures, and c_{wliq} is found by solving equation (5.6) for $c(T_m)$.

The dimensionless crystal growth parameter α is defined by the transcendental equation (MacFarlane & Fragoulis, 1986)

$$\alpha(T)^3 = 2c_{ws}(T) \exp\left[-\alpha(T)^2 / 4\right] \int_{\alpha(T)}^{\infty} x^{-2} \exp\left[-x^2 / 4\right] dx \quad (5.11)$$

with

$$c_{ws}(T) = \frac{c_{w\infty,m} - c_{wliq,m}(T)}{c_{w\beta,m} - c_{wliq,m}(T)} \quad (5.12)$$

where c_{ws} is the nondimensional supersaturation and $c_{w\beta}$ is the water concentration inside the ice crystal (assumed to be 55.6 M), and the “m” subscript again denotes concentrations in mol/L (Karlsson et al., 1994). Equation (5.11) can be solved numerically.

To evaluate the predictions of this model for critical cooling rates of aqueous glycerol solutions, Mathematica was used to numerically solve equation (5.3) and (5.5), using $T_i = T_m(c)$ and $T_f = T_g(c)$, temperatures corresponding to the melting transition and glass transition temperature of the solution. The assumed critical ice volume fraction was 0.01 (1%) roughly matched the minimum detectable ice fractions in our experiments.

The ice fraction formed on warming depends on the cooling history through the first term in Eq. (5.4). Critical warming rates were calculated using the same cooling rates as were used in preparing the samples for each CWR measurement (Table 5.1) at all glycerol concentrations except 40% w/w where the predicted CCR is higher than the experimental cooling rate.

To obtain a lower bound on the predicted ice fraction and thus a lower bound on the critical warming rate, we neglected ice nucleated on cooling and evaluated only the second term of equation (5.4), which gives the ice that forms on warming a perfectly vitrified sample. CWRs

for glycerol concentrations from 1% to 55% w/w. were evaluated numerically in Mathematica, using the same temperature limits as the CCR calculations.

CCR and CWR calculations were repeated for DMSO using functional forms for the viscosity as a function of temperature and concentration and for the melting/liquidus temperature as a function of concentration given in (Karlsson, 2010). The calculated CCRs and CWRs (not shown) were orders of magnitude greater than those calculated for glycerol, likely due to less accurate input parameters. In (Karlsson, 2010) functional forms for the viscosity and melting temperature of aqueous solutions of DMSO are provided. However, no comparison with experimental data is provided for the viscosity, and the melting temperature expression is based on data extending only to ~ 253 K (-20 °C). The viscosity and melting temperature for glycerol in (Karlsson, 2010) does not agree with measurements (Bohon & Conway, 1972; Trejo González et al., 2011) and so there may be similar inaccuracies in the DMSO data.

5.4 Results

5.4.1 Critical warming and cooling rate measurements

Figure 5.3 shows semi-log plots of critical warming rates, measured after cooling to $T=100$ K in the cold gas stream, as well as new and published critical cooling rates as a function of cryoprotectant concentration for glycerol (Boutron & Kaufmann, 1979b; Sutton, 1991; Wowk et al., 1999; Berejnov et al., 2006; Warkentin et al., 2008), DMSO (Sutton, 1991; Baudot et al., 2000; Berejnov et al., 2006), ethylene glycol (Boutron & Kaufmann, 1979b; Wowk et al., 1999; Baudot & Odagescu, 2004; Berejnov et al., 2006), PEG 200 (Berejnov et al., 2006) and sucrose (Berejnov et al., 2006). Table 5.1 gives the measured critical warming rates and Table 5.2 gives the measured critical cooling rates. Over the concentration range measured, both critical

Figure 5.3 – (a)-(e): Measured critical warming rates (red circles); previously predicted critical warming rates (orange squares) for glycerol (Boutron & Mehl, 1990), DMSO (Boutron & Kaufmann, 1978; Boutron & Mehl, 1990; Baudot et al., 2000), and ethylene glycol (Boutron & Mehl, 1990; Wowk et al., 1999; Baudot & Odagescu, 2004); previously measured CWRs (pink circles) for DMSO (Boutron & Kaufmann, 1978; Boutron & Mehl, 1990; Baudot et al., 2000); previously measured CCRs (blue diamonds) for glycerol (Boutron & Kaufmann, 1979b; Sutton, 1991; Wowk et al., 1999; Berejnov et al., 2006; Warkentin et al., 2008), DMSO (Sutton, 1991; Baudot et al., 2000; Berejnov et al., 2006), ethylene glycol (Boutron & Kaufmann, 1979b; Wowk et al., 1999; Baudot & Odagescu, 2004; Berejnov et al., 2006), PEG 200 (Berejnov et al., 2006), and sucrose (Berejnov et al., 2006) solutions; and the CCR of water (teal diamond) (Brüggeller & Mayer, 1980; Kohl et al., 2005). The red and blue (upper and lower) lines are exponential fits to the critical warming and cooling rate data, respectively. (f): Measured critical warming rates (circles), the CCR of water (diamond) (Brüggeller & Mayer, 1980; Kohl et al., 2005), and exponential fits (solid lines) to the CWRs for the aqueous solutions of (a)-(e).

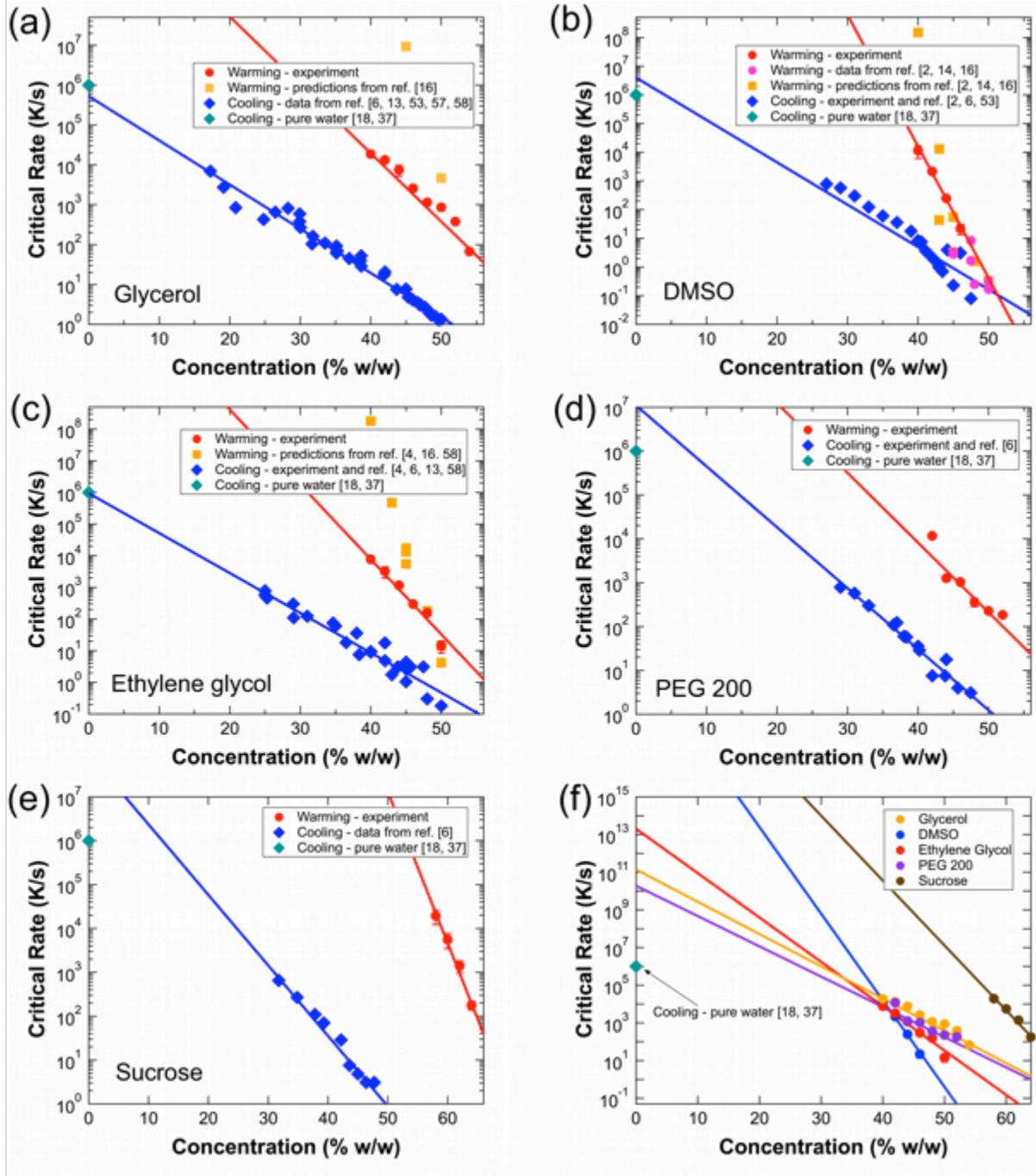


Table 5.1 – Measured critical warming rate, CWR; the uncertainty in the measurement, as discussed in the Methods section; the cooling rate, CR, at which the critical warming rate was measured; and previously measured critical cooling rate, CCR, if known. As discussed in Section 5.2.3, warming rates are varied by varying the drop size, which also leads to variations cooling rate. The conversion from concentration c in w/w to v/v is given by $c_w = \frac{1}{\rho_s / c - (\rho_s - 1)}$ where ρ_s is the solute density in g/cm^3 . Drop volumes were roughly 5, 1, 0.2, 0.01 and 0.005 μl for CWRs of 20, 300, 1200, 7000 and 19,000 K/s, respectively.

Cryoprotectant	Concentration (% w/w)	CWR (K/s)	CWR Uncertainty (K/s)	CR (K/s)	CCR (K/s) if known
Glycerol $\rho_s = 1.261\text{g} / \text{cm}^3$	54	67.1	3	38.8	
	52	379	46	66.1	
	50	866	162	168	0.7-1.4 (Boutron et al., 1979; Sutton, 1991; Boutron, 1993; Wowk et al., 1999)
	48	1150	50	95.1	2.3 (Sutton, 1991)
	46	2590	490	500	4.0-4.8 (Sutton, 1991; Berejnov et al., 2006)
	44	7520	2310	753	7.5 (Berejnov et al., 2006)
	42	13300	700	856	20.7 (Warkentin et al., 2008)
	40	18700	400	1050	35.8
DMSO $\rho_s = 1.1004\text{g} / \text{cm}^3$	46	22.1	8.8	9	3.1 (Berejnov et al., 2006)
	44	248	45	31.9	0.7-4.0 (Sutton, 1991; Berejnov et al., 2006)
	42	2160	400	309	1.8 (Sutton, 1991)
	40	11400	5300	738	7.5-8.3 (Baudot et

					al., 2000; Berejnov et al., 2006)
Ethylene Glycol $\rho_s = 1.1132g / cm^3$	50	14.3	5.9	8	0.08-18 (Wowk et al., 1999; Baudot & Odagescu, 2004)
	48	156	34	42.6	0.3-3.1 (Baudot & Odagescu, 2004; Berejnov et al., 2006)
	46	299	9	141	3.1 (Berejnov et al., 2006)
	44	1160	140	178	3.1 (Berejnov et al., 2006)
	42	3230	1270	295	4.8-17.7 (Berejnov et al., 2006)
	40	7690	6	567	9.0-9.5 (Wowk et al., 1999; Baudot & Odagescu, 2004)
PEG 200 $\rho_s = 1.1239g / cm^3$	52	185	16	69.4	
	50	227	11	60.9	
	48	353	79	71.7	3.1 (Berejnov et al., 2006)
	46	1040	140	189	4.0 (Berejnov et al., 2006)
	44	1270	200	259	7.5-17.8 (Berejnov et al., 2006)
	42	11700	1800	231	7.5 (Berejnov et al., 2006)
Sucrose $\rho_s = 1.587g / cm^3$	64	178	40	37.1	
	62	1400	430	107	
	60	5560	2090	564	
	58	19400	7090	812	

Table 5.2 – Critical cooling rates (CCR) for DMSO, ethylene glycol and PEG 200, measured as described in Section 5.2.4.

Cryoprotectant	Concentration (% w/w)	CCR (K/s)
DMSO	39	17.7
	37	35.8
	35	60
	33	123
	32	300
	29	574
	27	784
Ethylene Glycol	42	17.7
	38	35.8
	35	60
	31	123
	29	300
	25	784
PEG 200	44	17.7
	40	35.8
	38	60
	37	123
	33	300
	31	574
	29	784

warming rates and critical cooling rates roughly fall on straight lines, which suggests an exponential dependence on concentration. Figure 5.3 shows that the critical warming rates are uniformly larger – by factors of 10 to 1000 – than the critical cooling rates at the same concentration. The slopes of the critical warming rate data are larger than those of the critical cooling rate data, as summarized in Table 5.3. Figure 5.3b shows that previous critical warming rate data for large concentrations of DMSO (Boutron & Kaufmann, 1978; Boutron & Mehl, 1990; Baudot et al., 2000) are consistent with the present measurements at lower concentrations and higher warming rates.

The effect of cooling rate on the critical warming rate was measured for a 44% w/w ethylene glycol solution. The critical cooling rate of this solution, $\sim 3\text{-}5$ K/s (Boutron & Kaufmann, 1979b; Wowk et al., 1999; Baudot & Odagescu, 2004; Berejnov et al., 2006), was just below the smallest cooling rate we could achieve with our setup. This allowed the effect of cooling rate on critical warming rate to be measured for cooling rates from ~ 1 to 10^2 times the CCR. Figure 5.4 shows the results. At cooling rates from ~ 2 to 20 times the critical cooling rate, the critical warming rate was larger than 5000 K/s (the fastest rate measurable with the apparatus used for this set of measurements). However, at a cooling rate of ~ 1000 K/s, ~ 200 times the critical cooling rate, the critical warming rate dropped to ~ 350 K/s.

5.4.2 Model calculations

Figure 5.5 shows critical warming rates predicted assuming the experimental cooling rates (and thus including effects of ice formed during cooling), lower bound critical warming rates predicted assuming warming from an ice-free (fully vitrified) initial state, predicted critical cooling rates, critical warming rates predicted by Eq. (15) of Ref. (Boutron & Mehl, 1990), and

Table 5.3 – Coefficients α_{CCR} , α_{CWR} , $q_{0,CCR}$ and $q_{0,CWR}$ for exponential fits of the form $q_{crit}(c) = q_0 \exp(-\alpha c)$ to experimental data for the critical cooling warming rates, respectively, for aqueous solutions of five cryoprotectants. Note that c is in w/w, and takes values from 0 to 1.

Cryoprotectant	α_{CCR}	α_{CWR}	$\alpha_{CWR} / \alpha_{CCR}$	$q_{0,CCR}$ (K/s)	$q_{0,CWR}$ (K/s)
Glycerol	26	39	1.5	$5.3 \cdot 10^5$	$1.4 \cdot 10^{11}$
DMSO	34	110	3.2	$4.0 \cdot 10^6$	$3.6 \cdot 10^{22}$
Ethylene Glycol	29	54	1.9	$9.4 \cdot 10^5$	$2.3 \cdot 10^{13}$
PEG 200	32	37	1.2	$1.1 \cdot 10^7$	$1.9 \cdot 10^{10}$
Sucrose	37	79	2.1	$9.5 \cdot 10^7$	$2.2 \cdot 10^{24}$

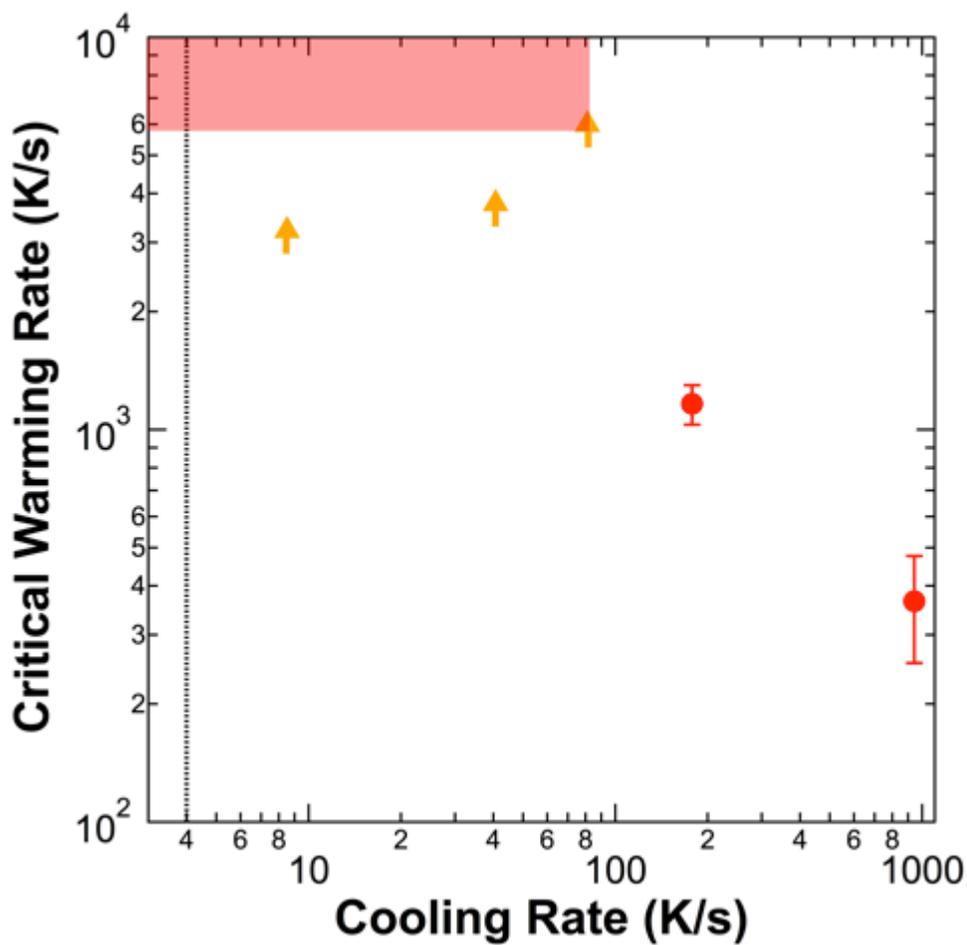


Figure 5.4 – Measured critical warming rate versus sample cooling rate for 44% w/w ethylene glycol solutions. Red circles indicate measured critical warming rates. Orange arrows indicate measured lower limits for the critical warming rate when the rate was too fast to measure; the actual warming rates must lie in or above the red shaded region. The dashed line indicates the measured critical cooling rate (Boutron & Kaufmann, 1979b; Wowk et al., 1999; Baudot & Odagescu, 2004; Berejnov et al., 2006).

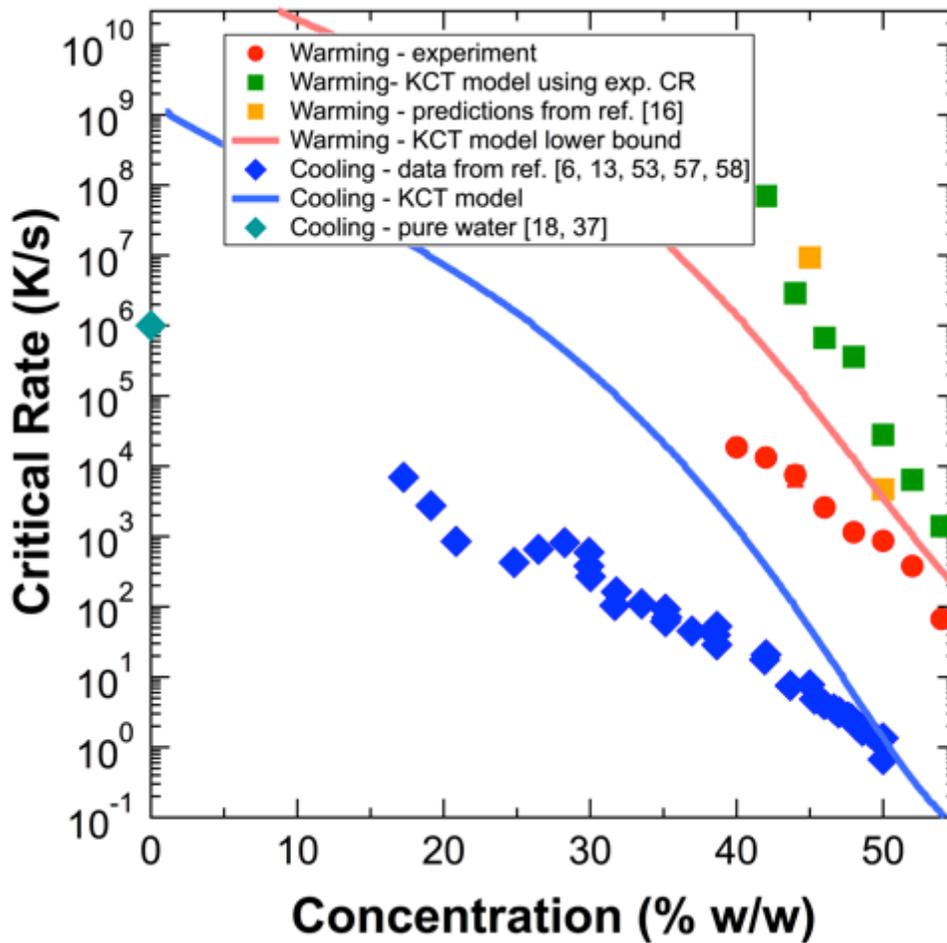


Figure 5.5 – Measured critical warming (red circles) and cooling rates (blue diamonds) (Boutron & Kaufmann, 1979b; Sutton, 1991; Wowk et al., 1999; Berejnov et al., 2006; Warkentin et al., 2008) , predicted critical warming rates from previous data (yellow squares) (Boutron & Mehl, 1990), our modeled critical warming rates using the experimental cooling rate at each data point (green squares), our modeled lower bound on the critical warming rates (assuming warming from an ice-free state) (red line), and our modeled critical cooling rates (blue line). CCR of water (diamond) (Brüggeller & Mayer, 1980; Kohl et al., 2005) shown for reference.

our experimental data versus concentration for glycerol. At all concentrations, the model's predicted lower bound on the critical warming rate is larger than the critical cooling rate, consistent with experiment. This predicted lower bound is in rough agreement with measured CWRs at larger concentrations, but is orders of magnitude larger than observed rates at smaller concentrations. Critical warming rates from 42-54% w/w glycerol predicted using the experimental cooling rates are uniformly higher than the lower bound prediction by a factor of ~ 10 , and larger than measured rates by a factor of $\sim 10-100$. Predicted critical cooling rates are in rough agreement with experiment at large concentrations, are orders of magnitude larger than experiment at small concentrations, and do not reproduce the observed exponential concentration dependence. The discrepancies between predicted and measured rates are relatively insensitive to the assumed "critical" ice fraction: increasing (decreasing) the fraction from 0.01 to 0.1 (0.001) changes the critical warming and cooling rates by less than a factor of four over the examined concentration range. Consequently, the discrepancies are relatively insensitive to uncertainty/variation in the minimum detectable ice fractions in our experiments. From a previous simulation of the KCT model (Karlsson, 2001), a critical warming rate of ~ 20 K/s can be deduced for an initially ice-free (i.e., fully vitrified) solution containing 7.9 M (63% w/w) glycerol and 0.6 M NaCl, too large given our measured CWR of 67.7 K/s at 54% w/w glycerol. Our calculated lower bound is 0.06 K/s, and our experimental data extrapolates to ~ 1 K/s. Slow rates at this concentration are expected, as water crystallization in glycerol concentrations greater than 66.7% w/w is so slow as to be unobservable on laboratory timescales (Bohon & Conway, 1972).

As a check on the model and our calculations in their most likely regime of applicability (i.e., at high concentrations), nucleation rates and isothermal ice growth rates were calculated as a

function of temperature for a 50% glycerol solution. For the diffusion limited growth model used in the KCT model, the linear growth rate depends on the size r of the particle (MacFarlane & Fragoulis, 1986)

$$\frac{dr}{dt} = \frac{\alpha(T)^2 D(c,T)}{2r}. \quad (5.13)$$

To compare model predictions with the experimental data in (Hey & MacFarlane, 1998), an initial radius of $0.1 \mu\text{m}$ – an estimate for the smallest detectable nucleus based on images in (Hey & MacFarlane, 1998) - was used. As shown in Figure 5.6, predicted initial growth rates are a factor of four below experimentally measured growth rates (Hey & MacFarlane, 1998). The temperature of the peak nucleation rate is roughly 10 K below the temperature with the smallest time to transition (183 K (-90° C)) on the measured TTT curve (Hey & MacFarlane, 1996). The calculations thus provide a reasonable but not completely accurate description of the experimental nucleation and growth rates relevant at high glycerol concentrations.

5.5 Discussion

5.5.1 Concentration dependence

For all cryoprotectants studied here, the available data for both the critical cooling and warming rates show an approximately exponential variation with cryoprotectant concentration (w/w) c i.e.

$$q_{crit}(c) = q_0 \exp(-\alpha c) \quad (5.14)$$

where q_{crit} is the critical rate, q_0 is the critical rate at zero concentration and α is a scale factor.

Table 5.3 gives parameters obtained by fitting Eq. (5.14) to data for several cryoprotectants. For all cryoprotectants, α is larger, by factors ranging from 1.2 to 3.2, for warming than for cooling.

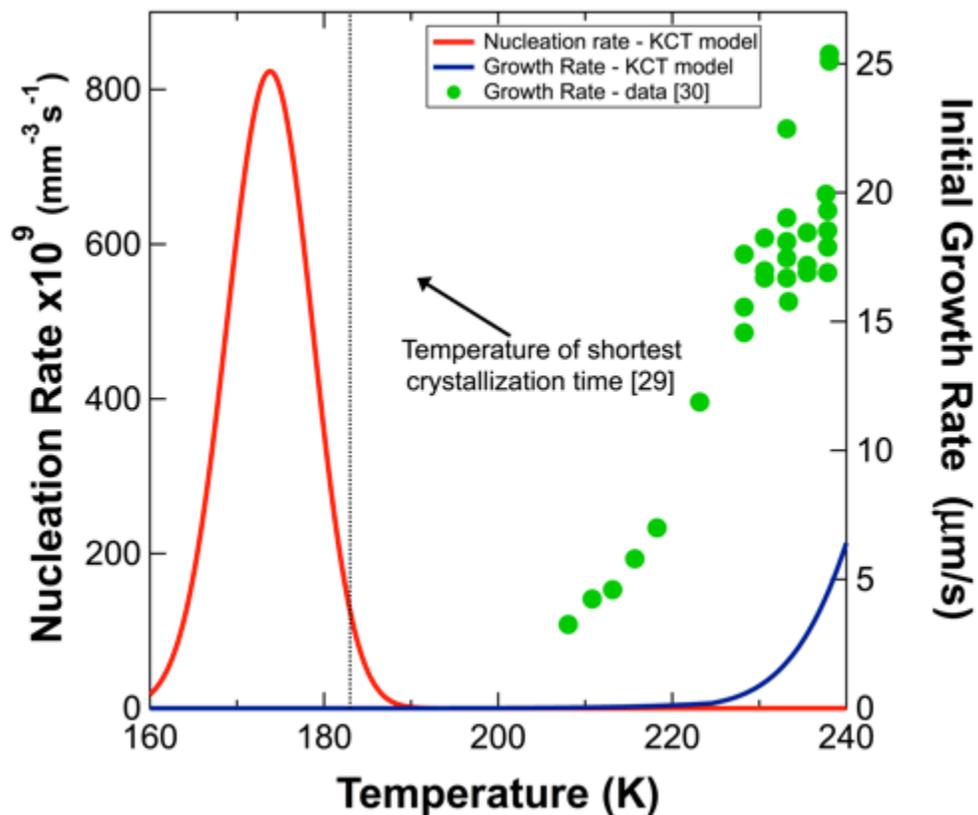


Figure 5.6. Calculated nucleation rate (red, left axis) and growth rate (blue, right axis) for a 50% w/w glycerol solution, using the KCT model. Experimentally measured growth rates (green circles) and the temperature of least time on a TTT curve (dotted vertical line) are shown for comparison (Hey & MacFarlane, 1996, 1998).

Previous exponential fits to critical cooling rate data for glycerol (Warkentin et al., 2008) extrapolated at $c = 0$ to the measured critical cooling rate of pure water, as one would expect if the concentration dependence of the critical cooling rate was, in fact, exponential. As shown in Figure 5.3f, extrapolations of exponential fits to the present CWR data do not converge to the same $c = 0$ value — the (presently unknown and possibly unmeasurable) critical warming rate of pure water.

5.5.2 Variation of critical warming rate with cooling rate

The critical warming rate strongly depends upon the cooling rate, even when cooling rates are much larger than the critical cooling rate. As shown in Figure 5.4, increasing the cooling rate from ~ 20 to ~ 200 times the critical cooling rate decreases the critical warming rate by a factor of at least 15. This strong dependence is likely due to the rapid growth during warming of ice that nucleated during cooling.

The measured critical cooling rate is defined as the minimum rate required so that no ice is detectable, and so depends on the ice detection sensitivity. Calculations suggest that reducing the ice volume fraction from 2.3% to $10^{-4}\%$ requires cooling rates be increased by roughly two orders of magnitude (Kresin et al., 1991). Consequently, even cooling rates much larger than the measured critical cooling rates should yield ‘vitreous’ samples with significant ice nuclei concentrations.

The strong cooling rate dependence of the critical warming rates in Figure 5.4 suggests that, for the cryoprotectant concentration and cooling rate range shown, growth during warming of nuclei formed on cooling produces much more ice than growth of nuclei formed on warming. With sufficiently large cooling rates, this should no longer be the case, and the CWR should asymptote

to a minimum, cooling-rate-independent value (which of course will still depend upon the minimum detectable ice fraction.) These CWR values might then show simpler asymptotic variation to $c=0$ than is seen in Figure 5.3. Achieving the required cooling rates, especially at smaller concentrations, may be difficult or impossible.

5.5.3 Excess of CWR over CCR

For the concentration ranges shown in Figure 5.3, critical warming rates for all cryoprotectants examined are larger, by a factor of 10 to 1000, than the critical cooling rates for the same concentration. This difference can be explained by differences in ice crystal nucleation and growth. As shown in Figure 5.6, on cooling from the liquid phase the probability per unit time of ice nucleation grows, reaching a maximum at some temperature below the homogeneous nucleation temperature. The final ice fraction within a sample then depends upon the number of nuclei formed and how much they grow before the temperature drops below the vitrification temperature. Growth rates decrease with decreasing temperature, and are small at temperatures comparable to and below that corresponding to the peak nucleation rate (e.g., as experimentally observed for glycerol and DMSO (Hey & MacFarlane, 1996, 1998)). Consequently, on cooling only modest ice growth following nucleation occurs, and so relatively modest cooling rates are required to achieve small ice fractions.

On warming above the devitrification temperature, existing ice crystals can resume their growth, new crystals can nucleate and grow, larger crystals can grow at the expense of smaller ones (recrystallization), and growth can continue all the way up to the melting temperature. More nuclei combined with much larger growth rates at temperatures above that of the peak nucleation rate lead to much larger ice fractions for a given rate q_w than on cooling. Critical warming rates

to achieve a given ice fraction are thus much larger than the critical cooling rates of the same solution.

5.5.4 Predictions of critical warming rates based upon previous data

Differential scanning calorimetry has been used to measure the devitrification temperature T_d and the melting temperature T_m of many aqueous solutions at small warming rates ($\sim 0.04 - 5$ K/s). As the warming rate q_w increases, $T_m - T_d$ decreases, and over the range of warming rates explored $T_m - T_d$ varies roughly linearly with $\log(1/q_w)$ (Boutron & Kaufmann, 1978, 1979a; Boutron et al., 1982, 1986, 1979, 1980; Mehl & Boutron, 1987; Boutron, 1990, 1993; Boutron & Mehl, 1990; Mehl, 1993; Baudot et al., 1996, 2000, 2002; Wowk et al., 1999). Ice formation on warming at a given rate is reduced if the difference between T_d and T_m is small. By extrapolating the measured warming rate dependence of $T_m - T_d$ to an arbitrary threshold $T_m - T_d = 5$ K, the critical warming rate was estimated.

A reformulation of this method used a critical ice fraction rather than a $T_m - T_d$ threshold to obtain improved, somewhat lower estimates of the CWR (Boutron & Mehl, 1990). From measurements of the devitrification temperature at two different warming rates q_{w_1} and q_{w_2} , the critical warming rate $q_{w,crit}$ was calculated as

$$\log q_{w,crit} = \frac{T_{d_1} T_{d_2}}{T_{d_1} - T_{d_2}} \left[\frac{1}{A_{crit} T_m} \log \left(\frac{q_{w_2}}{q_{w_1}} \right) + \frac{1}{T_{d_2}} \log q_{w_1} - \frac{1}{T_{d_1}} \log q_{w_2} \right]. \quad (5.15)$$

Here A_{crit} corresponds to the ratio T_d / T_m at the desired critical volume fraction of ice, and has usually been chosen to be 0.95. Predictions based upon data collected at warming rates between 0.04 and 5 K/s are available for glycerol (Boutron & Mehl, 1990), DMSO (Boutron &

Kaufmann, 1978; Boutron & Mehl, 1990; Baudot et al., 2000) and ethylene glycol (Boutron & Mehl, 1990; Wowk et al., 1999; Baudot & Odagescu, 2004), and are shown in Figure 5.3a, 5.3b and 5.3c. These calculated rates are close to the measured ones at large concentrations, i.e., closer to the range explored in previous critical warming rate measurements. But at the smaller concentrations examined here, the predicted critical warming rates rapidly diverge above our measured values.

5.5.5 Predictions of critical warming and cooling rates based upon models of ice nucleation and growth

The predicted critical warming rates (using measured cooling rates) and the predicted critical cooling rates differ from the experimental results as shown in Figure 5.5, and are orders of magnitude larger than experiment at lower cryoprotectant concentrations. Since the predicted critical cooling rates are much larger than observed, the pre-existing ice fraction at the start of warming following cooling at a given rate should be overestimated, which should lead to an over-estimate (within the model) of the critical warming rate. Even a predicted lower bound on critical warming rates (assuming warming from an initially ice-free state) is far larger than experiment. As discussed in Section 5.3.2 differences between model predictions and experiment are relatively insensitive to differences between calculated and experimentally estimated ice fractions.

These very large differences between model predictions for CCRs and CWRs and experiment at lower concentrations could be due to flaws in the input parameters and/or in the model itself. Available data for water-glycerol solution viscosities only extend to ~240 K (-33° C)(Trejo González et al., 2011), and so extrapolations to lower temperatures may be in error. The non-

Stokes-Einstein fragile to strong transition observed in the diffusion constant and assumed in our calculations is not well characterized in the presence of solutes (Mallamace et al., 2010), and may not exist at large concentrations (Nakanishi et al., 2012). However, simulations using the Stokes-Einstein diffusion relation using otherwise unchanged parameters resulted in a decrease in the predicted CCRs and an increase in the predicted CWRs by at most a factor of ~ 5 . Critical cooling rates for water-glycerol mixtures predicted for the vitrification regime (where water transport across the cell membrane becomes unimportant) in Ref. (Karlsson et al., 1994), using a different viscosity model, are also orders of magnitude larger than experiment. Reasonable quantitative agreement between simulations using our original viscosity model and CCR data for glycerol concentrations below 40% w/w can be obtained by increasing the viscosity at all temperatures by a factor of 10^4 , but we have no physical basis for this increase.

Classical nucleation theories typically yield reasonable estimates for the thermodynamic barriers to nucleation, but predicted nucleation rates are often off by many orders of magnitude due to challenges in correctly modeling nucleation kinetics. The nucleation model used here (Toner et al., 1990) has only been tested at small (~ 2 K/s) cooling rates, where ice forms at relatively high temperatures for which experiment-based estimates of model parameters are more accurate. The model has not been verified at large cooling rates, where large supercooling may occur before appreciable nucleation and growth, nor has it been verified during warming from the vitrified state. In addition, the model calculates steady state rates, which may not be appropriate for large cooling/warming rates and small samples. On abruptly cooling (or warming) to a given temperature the nucleation rate does not instantaneously reach the steady-state value, as time is required to establish the steady-state distribution of sub- and supercritical nuclei sizes (Kelton et al., 1983). Transient nucleation rates are smaller than steady-state nucleation rates (Kelton et al.,

1983), and so could result in smaller predicted critical cooling and warming rates in the low-concentration, high rate regime, closer to experiment. While a numerical calculation of the lag time of the transient nucleation rate in an aqueous solution is beyond the scope of this paper, we note that for some systems measured lag times are as large as 10^4 s (Gutzow et al., 1997) and that calculated lag times range from at least 10^{-7} (Demetriou et al., 2002) to 10^7 s (Demo & Kozisek, 1996). A lag time of ~ 1 ms or greater would have an effect on our measured critical rates at low concentration.

The diffusion limited growth model (MacFarlane & Fragoulis, 1986; Karlsson et al., 1994) we use has not been verified at the large cooling and warming rates studied here, or at lower temperatures. In this model, the growth rate varies as $t^{1/2}$, but measured isothermal growth rates in 50% glycerol solutions were time independent at temperatures lower than ~ 233 K (-40° C) (Hey & MacFarlane, 1998). At large supersaturations and large cooling and warming rates, transfer of latent heat from the phase boundary becomes a bottleneck and the assumption of local equilibrium may break down. Dendritic growth may then dominate (Hey & MacFarlane, 1998). Dendritic growth has been observed to be the primary mode of ice growth in solutions of ~ 20 - 60% glycerol at temperatures less than ~ 200 K (-73° C) (Rapatz & Luyet, 1966).

5.5.6 Application to biological cryopreservation

The primary objective of the work presented here is to quantify fundamental aspects of ice formation in aqueous cryoprotectant solutions, especially in the largely unexplored regime of large warming rates. These results for two-component systems provide a basis for understanding the behavior of more complex systems.

What complications may arise in applying these results to cellular systems? Unlike at small cooling and warming rates, at high rates characteristic of vitrification procedures the effects of water transport across cell membranes and intracellular dehydration driven by extracellular ice nucleation can be neglected (Karlsson et al., 1994; Karlsson, 2001; Zhao et al., 2006).

Nucleation and growth behavior may be different in the cellular environment. Some experiments on cellular solutions have been interpreted as providing evidence that ice forms by heterogeneous nucleation (Franks et al., 1983), while others suggest that cryoprotectants suppress heterogeneous nucleation so that homogeneous nucleation dominates (Rall et al., 1983). These studies used small cooling rates (< 0.3 K/s), where there is ample time for heterogeneous nucleation before temperatures at which rapid homogeneous nucleation occurs are reached. At the much larger cooling and warming rates studied here, homogeneous nucleation should dominate, consistent with our high speed video observation that ice always first appeared in the bulk both on cooling and on warming.

Cellular components could also affect ice crystal growth and recrystallization rates. Our preliminary attempts to measure critical cooling and warming rates in aqueous solutions containing up to 50% w/w lysozyme suggest that soluble proteins are very poor cryoprotectants, as has been found in studies of ice formation in hydrated protein powders (Sartor et al., 1995).

Consequently, at the large cooling and warming rates studied here, critical warming and cooling rates for cryoprotectant-containing cellular solutions may be similar to those obtained here for cell-free solutions.

5.6 Conclusions

The present work provides a more quantitative basis for rational selection of cryoprotectants for cryopreservation. We have measured critical warming rates for several cryoprotectants at rates up to three orders of magnitude larger than have been previously reported. For all cryoprotectants studied, measured critical warming rates are orders of magnitude larger than the critical cooling rates. For comparable rates, warming yields much larger maximum ice fractions than cooling. Consequently, an increase in warming rate has a much larger effect on maximum ice fraction in freeze-thaw cycles than a comparable increase in cooling rate. This is consistent with recent observations that warming rate is more important to biological survival than cooling rate (Seki & Mazur, 2008, 2009; Mazur & Seki, 2011).

Critical warming rates are also strongly dependent on cooling rates, and can be decreased by an order of magnitude or more by increasing cooling rates. Since maximum warming rates are limited by maximum tolerable sample temperatures, this suggests that samples should always be cooled at the maximum rate possible, and that maximizing both cooling and warming rates are necessary to minimize ice formation and required cryoprotectant concentrations.

REFERENCES

- Banerjee, D., Bhat, S. N., Bhat, S. V., & Leporini, D. (2009). *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11448–11453.
- Baudot, A., Alger, L., & Boutron, P. (2000). *Cryobiology*. **40**, 151–158.
- Baudot, A., Cacela, C., Duarte, M. L., & Fausto, R. (2002). *Cryobiology*. **44**, 150–160.
- Baudot, A. & Odagescu, V. (2004). *Cryobiology*. **48**, 283–294.
- Baudot, A., Peyridieu, J. F., Boutron, P., Mazuer, J., & Odin, J. (1996). *Cryobiology*. **33**, 363–375.
- Berejnov, V., Husseini, N. S., Alsaied, O. A., & Thorne, R. E. (2006). *J. Appl. Crystallogr.* **39**, 244–251.
- Bohon, R. L. & Conway, W. T. (1972). *Thermochim. Acta*. **4**, 321–341.
- Boutron, P. (1990). *Cryobiology*. **27**, 55–69.
- Boutron, P. (1993). *Cryobiology*. **30**, 86–97.
- Boutron, P., Delage, D., & Roustit, B. (1980). *J. Chim. Phys.* **77**, 567–570.
- Boutron, P., Delage, D., Roustit, B., & Körber, C. (1982). *Cryobiology*. **19**, 550–564.
- Boutron, P. & Kaufmann, A. (1978). *Cryobiology*. **15**, 93–108.
- Boutron, P. & Kaufmann, A. (1979a). *Cryobiology*. **16**, 557–568.
- Boutron, P. & Kaufmann, A. (1979b). *Cryobiology*. **16**, 83–89.
- Boutron, P., Kaufmann, A., & Dang, N. Van (1979). *Cryobiology*. **16**, 372–389.
- Boutron, P. & Mehl, P. (1990). *Cryobiology*. **27**, 359–377.
- Boutron, P., Mehl, P., Kaufmann, A., & Angibaud, P. (1986). *Cryobiology*. **23**, 453–469.
- Brüggeller, P. & Mayer, E. (1980). *Nature*. **288**, 569–571.
- Chen, S.-H., Mallamace, F., Mou, C.-Y., Broccio, M., Corsaro, C., Faraone, A., & Liu, L. (2006). *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12974–12978.
- Christian, J. W. (1975). *The theory of transformations in metals and alloys*.

- Debenedetti, P. G. (2003). *J. Phys. Condens. Matter*. **15**, R1669–R1726.
- Demetriou, M. D., Ghoniem, N. M., & Lavine, A. S. (2002). *J. Chem Phys*. **117**, 10739.
- Demo, P. & Kozisek, Z. (1996). *Thermochem. Acta*. **280/281**, 101–126.
- Fowler, A. & Toner, M. (2005). *Ann. N. Y. Acad. Sci.* **1066**, 119–135.
- Franks, F., Mathias, S. F., Galfre, P., Webster, S. D., & Brown, D. (1983). *Cryobiology*. **20**, 298–309.
- Fuller, B. J. (2004). *CryoLetters*. **25**, 375–388.
- Gook, D. A. (2010). *Reprod. Biomed. Online*.
- Gutzow, I., Schmelzer, J., & Dobрева, A. (1997). *J. Non-Cryst. Solids*. **219**, 1–16.
- Hey, J. M. & MacFarlane, D. R. (1996). *Cryobiology*. **33**, 205–216.
- Hey, J. M. & MacFarlane, D. R. (1998). *Cryobiology*. **37**, 119–130.
- Hopkins, J. B., Badeau, R., Warkentin, M., & Thorne, R. E. (2012). *Cryobiology*. **65**, 169–178.
- Karlsson, J. O. M. (2001). *Cryobiology*. **42**, 154–169.
- Karlsson, J. O. M. (2010). *Cryobiology*. **60**, 43–51.
- Karlsson, J. O. M., Cravalho, E. G., Borel Rinkes, I. H. M., Tompkins, R. G., Yarmush, M. L., & Toner, M. (1993). *Biophys. J.* **65**, 2524–2536.
- Karlsson, J. O. M., Cravalho, E. G., & Toner, M. (1994). *J. Appl. Phys.* **75**, 4442–4455.
- Karlsson, J. O. M., Eroglu, A., Toth, T. L., Cravalho, E. G., & Toner, M. (1996). *Hum Reprod.* **11**, 1296–1305.
- Kelton, K. F., Greer, a. L., & Thompson, C. V. (1983). *J. Chem. Phys.* **79**, 6261.
- Kohl, I., Bachmann, L., Hallbrucker, A., Mayer, E., & Loerting, T. (2005). *Phys. Chem. Chem. Phys.* **7**, 3210–3220.
- Kraust, G. F. & Greer, S. C. (1984). *J. Phys. Chem.* **88**, 4781–4785.
- Kresin, M., Körber, C., & Körber, C. (1991). *J. Chem. Phys.* **95**, 5249–5255.
- Kriminski, S., Kazmierczak, M., & Thorne, R. E. (2003). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **59**, 697–708.

- MacFarlane, D. R. & Fragoulis, M. (1986). *Phys. Chem. Glas.* **27**, 228–234.
- Mallamace, F., Branca, C., Corsaro, C., Leone, N., Spooren, J., Stanley, H. E., & Chen, S.-H. (2010). *J. Phys. Chem. B.* **114**, 1870–1878.
- Mazur, P. & Seki, S. (2011). *Cryobiology.* **62**, 1–7.
- Mehl, P. & Boutron, P. (1987). *CryoLetters.* **8**, 64–73.
- Mehl, P. M. (1993). *Thermochim. Acta.* **226**, 325–332.
- Nakanishi, M., Griffin, P., Mamontov, E., & Sokolov, A. P. (2012). *J. Chem. Phys.* **136**, 124512.
- Pereira, R. M. & Marques, C. C. (2008). *Cell Tissue Bank.* **9**, 267–277.
- Rall, W. F., Mazur, P., & McGrath, J. J. (1983). *Biophys. J.* **41**, 1–12.
- Rapatz, G. & Luyet, B. (1966). *Biodynamica.* **10**, 69–80.
- Sartor, G., Hallbrucker, A., & Mayer, E. (1995). *Biophys. J.* **69**, 2679–2694.
- Seki, S. & Mazur, P. (2008). *Biol. Reprod.* **79**, 727–737.
- Seki, S. & Mazur, P. (2009). *Cryobiology.* **59**, 75–82.
- Sutton, R. L. (1991). *J. Chem. Soc. Faraday Trans.* **87**, 101.
- Toner, M., Cravalho, E. G., & Karel, M. (1990). *J. Appl. Phys.* **67**, 1582–1593.
- Trejo González, J. A., Longinotti, M. P., & Corti, H. R. (2011). *J. Chem. Eng. Data.* **56**, 1397–1406.
- Warkentin, M., Berejnov, V., Hussein, N. S., & Thorne, R. E. (2006). *J. Appl. Crystallogr.* **39**, 805–811.
- Warkentin, M., Sethna, J. P., & Thorne, R. E. (2013). *Phys. Rev. Lett.* **110**, 015703.
- Warkentin, M., Stanislavskaja, V., Hammes, K., & Thorne, R. E. (2008). *J. Appl. Crystallogr.* **41**, 791–797.
- Warkentin, M. & Thorne, R. E. (2009). *J. Appl. Crystallogr.* **42**, 944–952.
- Wowk, B., Darwin, M., Harris, S., Russell, S., & Rasch, C. (1999). *Cryobiology.* **39**, 215–227.
- Zhao, G., Luo, D., & Gau, D. (2006). *AiChE J.* **52**, 2596–2606.

Zhmakin, A. I. (2008). *Phys.-Usp.* **51**, 231–252.

CHAPTER 6

CURRENT AND FUTURE WORK ON CRYOSAXS

6.1 Current problems and possible solutions

6.1.1 Problems

Currently, given a macromolecule in a buffer that allows for high quality cryocooling, data of quality similar to or better than that collected at room temperature can be obtained from as little as 1-5 μL of sample. However, several issues make this challenging, and prevent widespread adoption of the technique:

- 1) Proper buffer conditions need to be identified for the macromolecule. The cryoprotectant concentrations needed for vitrification do not always leave the protein structure unchanged at room temperature. More troubling, proteins can change during cooling. For example lysozyme in 36% w/w propylene glycol is monodisperse at room temperature but aggregated after cooling to 100 K, as shown in Figure 6.1 (note that this is for lysozyme in a different buffer than Figure 4.17, where, at least a CHESS, cryocooling has been successful).
- 2) The scattering profile at low temperature is not always easily interpretable. An example is xylanase in 45% w/w PEG, which has lower scattering from the protein solution than the buffer at high q (something that at room temperature would be considered nonsense), Figure 6.2. This could be due to changes in scattering from the bulk solvent and/or hydration shell.
- 3) The current procedure for data collection in the fixed path length silicon sample holders is time intensive: at least 5 minutes per sample is spent not collecting data. This is due to the volumetric precision needed when loading the sample, and the cleaning needed to reuse the sample holder for both buffer and sample. These requirements are explored in detail in Chapter

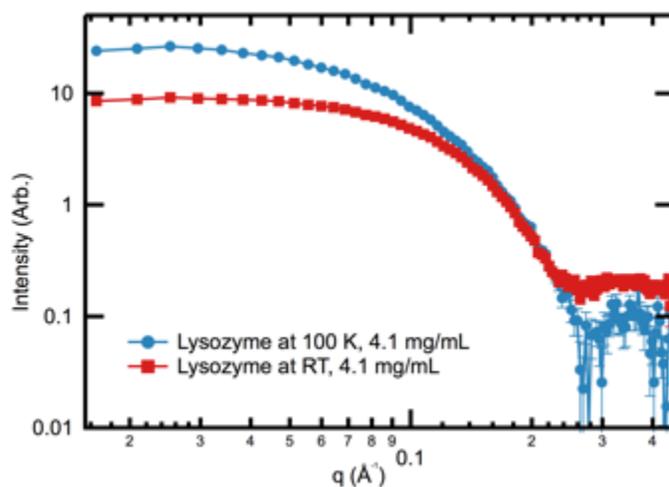


Figure 6.1 – Lysozyme at 4.1 mg/mL in a 36% w/w propylene glycol buffer at room temperature and 100 K. The profiles have been scaled to match at $q = 0.2 \text{ \AA}^{-1}$. The change in shape of the profiles indicates changes to the protein upon cooling. The larger low q at 100 K indicates a higher molecular mass (based on the extrapolation to zero scattering intensity). Evaluation of the $p(r)$ function using DATGNOM (Petoukhov & Konarev, 2007) shows a change upon cooling in D_{max} from 41 to 72 \AA and in $p(r)$ calculated R_g from 14.3 to 21.1 \AA . The increased size and mass at 100K indicate aggregation of the lysozyme during the cooling process.

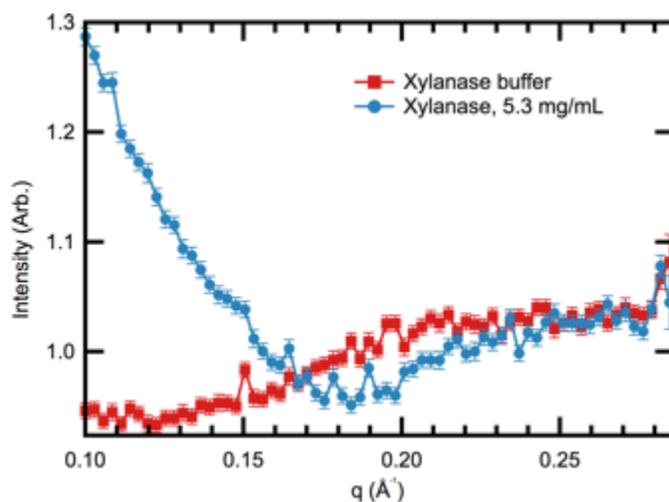


Figure 6.2 – 100 K scattering profiles for xylanase (5.3 mg/mL) and xylanase buffer, both with 45% w/w PEG 200. The protein curve (blue/circles) crosses below the buffer curve (red/squares) at high q . At room temperature this would be non physical, as scattering is not negative, and would indicate a mismatch with either the buffer or protein measurement. This could come from contamination of the sample holder, changes in the beam conditions, or mismatch between the buffer the protein is in and the buffer measure for background subtraction. While these are possibilities at 100 K, we have repeatedly observed this type of negative scattering and believe it may be associated with changes in the hydration shell upon cooling.

2. One of the possible advantages to cryoSAXS is that it could increase the rate of high throughput SAXS. Use of windowless holders could eliminate need for cleaning the sample holder, which at room temperature takes substantially more time than the actual data collection. Currently, data collection takes much longer than standard high throughput SAXS setups, so this improvement has not been achieved.

4) The current fixed path length sample holders and cooling method result in macroscopically visible fractures in the sample. These can lead to anisotropic scattering in the scattering profile. This anisotropic scatter can sometimes be masked out, but two identical measurements of any scattering profile are required to determine if the scattering has been affected. More details are available in Chapter 2.

5) Use of the original, windowless, sample holders was never completely reliable. In recent experiments with these sample holders, measured scattering profiles were mismatched by more than 1%, outside the experimental error. The original paper never discussed reproducibility (Meisburger et al., 2013), but personal experience indicates it is not ideal.

6) Preventing ice formation requires large concentrations of cryoprotectant, raising questions about if/how the protein is affected by these additives. Currently that question is answered by comparison to scattering profiles in standard buffers without cryoprotectants, but this is not a feasible procedure for many of the targets of cryoSAXS: macromolecules that are difficult or costly to express in the quantities typically needed for room temperature SAXS. In the long term, cryoSAXS must be divorced from room temperature SAXS to be an effective technique.

6.1.2 In situ deposition and cooling as a possible solution

Development of an in situ deposition and cooling method may be a solution to some of the

problems. Samples are currently cooled by placing them into a cold nitrogen gas stream. This approach is not optimized for maximum cooling rates. We have been working on a new system to deposit room temperature solution into a pre-cooled sample holder. This should significantly increase cooling rates (similar to splat cooling liquid onto cold copper). Faster cooling means less time for changes to the protein solution during cooling, and allows use of lower cryoprotectant concentrations, helping mitigate issues 1 and 6. It is likely that the fractures come from differential contraction between the sample holder and sample. If the sample holder is already cold, there is no differential contraction, so this approach may also ameliorate issue 4. This development is also one approach to time resolved cryoSAXS. It would allow mixing of samples in a standard microfluidic mixer, followed by ejection of the sample from the mixer and cooling in the cold sample holder. This is further discussed Section 6.2.

There are three main design challenges for in situ deposition and cooling. The first is how to mount a sample holder in the system and cool it. The sample holder must be cold (100 K or less), while the air very near the sample holder needs to be warm (273 K or greater), to prevent pre-cooling of the sample that would lower cooling rates. Further, the same sample holder must be used for both sample and buffer, and it must be returned to the same position with ~ 10 microns. This means that our sample holder must be able to be warmed, cleaned, and dried before a new sample is deposited. Additionally, all of this must be accomplished on reasonable timescales, so that experiments can be carried out in the limited available beam time.

To cool the sample holder, we use a copper cold finger in a liquid nitrogen dewar. The dewar is capped just below the top of the cold finger to prevent cold gas from swirling up around the sample holder. The cold finger is threaded at the top, allow a small piece of copper with a sample holder mounted to it to be screwed in. A thermal grease, such as Apiezon N Grease, helps

maintain good thermal contact, with measured sample temperatures ranging from ~85-95 K. To clean the sample holder, the dewar is drained and the cold finger is rapidly heated to room temperature (via an external heater). The sample holder mount is unscrewed, and the sample holder cleaned. The mount is then screwed back into the cold finger, and the cold finger cooled to liquid nitrogen temperatures.

The second requirement is that we must deposit a precise volume of liquid into the sample holder. In the current design liquid is loaded into 1/8" tubing connected to a 25-gauge needle, which is held a few inches above the sample holder. A fast solenoid is opened to place a 30 psi overpressure on the liquid. This ejects a stream of liquid from the needle. Simultaneous to the solenoid opening, a liquid shutter is activated. This shutter starts out blocking the liquid jet, moves to let it pass through to the sample holder, and then blocks it again. This allows deposition of a precise volume that depends on the speed of the shutter. When the solenoid is closed, the overpressure vents and deposition stops. Sometimes the sample doesn't freeze in the sample holder, instead splashing out. However, adjustment of the deposition pressure seems to help alleviate this issue. One entire cycle, putting in the sample holder, cooling down, depositing sample, warming up, and cleaning the holder, takes about 30-45 minutes. While it is not fast, several experiments can reasonable be carried out during a single beam time.

The third major design consideration is that to prevent deposition of water vapor or other gasses on the surface of the sample holder, it must be in a high purity nitrogen, or vacuum, environment. The SAXS signal from ice is extremely strong, capable of overwhelming the protein signal. Growth of ice also changes the scatter as a function of time, preventing longer exposure times. The presence of ice on the sample holder may act as heterogeneous nucleation sites for ice in the sample, so for a given cooling rate a much higher cryoprotectant concentration

would be needed to prevent ice formation in the sample.

Our current design has the sample holder inside by a small (~3 in x 3 in x 1 in) aluminum box, which is flushed with ultra high purity nitrogen (99.9993+% pure). This box has polystyrene x-ray windows, and small hole in the top to allow the sample to enter. The shutter blocks this hole except during sample deposition, and the box is maintained at a modest overpressure for minimal contamination. One side of the box can be removed, so that the sample holder mount can be put on and taken off the cold finger. Figure 6.3 is a CAD image showing the major features. There is a 10 W (maximum) resistive heater in the box to maintain the interior nitrogen temperature, preventing pre-cooling or freezing of the sample prior to deposition.

This design is unable to prevent ice formation on the sample holder. Over the course of several minutes, there is a visible change in the scattering profile, Figure 6.4, which we attribute to deposition of impurity gases on the sample holder. There was no visually observable ice formation on the sample holder and the box had no significant leaks. When the aluminum box was removed and the sample holder was put in a warm nitrogen stream fed by liquid nitrogen boil off (a 300 K nitrogen stream from an Oxford Cryosystems 700 Series Cryostream), scattering from the sample holder remained mostly constant for more than 30 minutes. The observed variation in the scattering in this setup was non systematic, suggesting it came from either frost near the edges of the stream growing and occasionally getting blown off or swirling frost and cold air in the beam path (given the dramatic changes in setup this required, we were unable to test deposition in this configuration while at CHESS). This suggests the problem is likely one of insufficient purity of our nitrogen flush in the chamber. As the cold finger and sample holder are a cryopump, even small amounts of contaminants could lead to significant deposition.

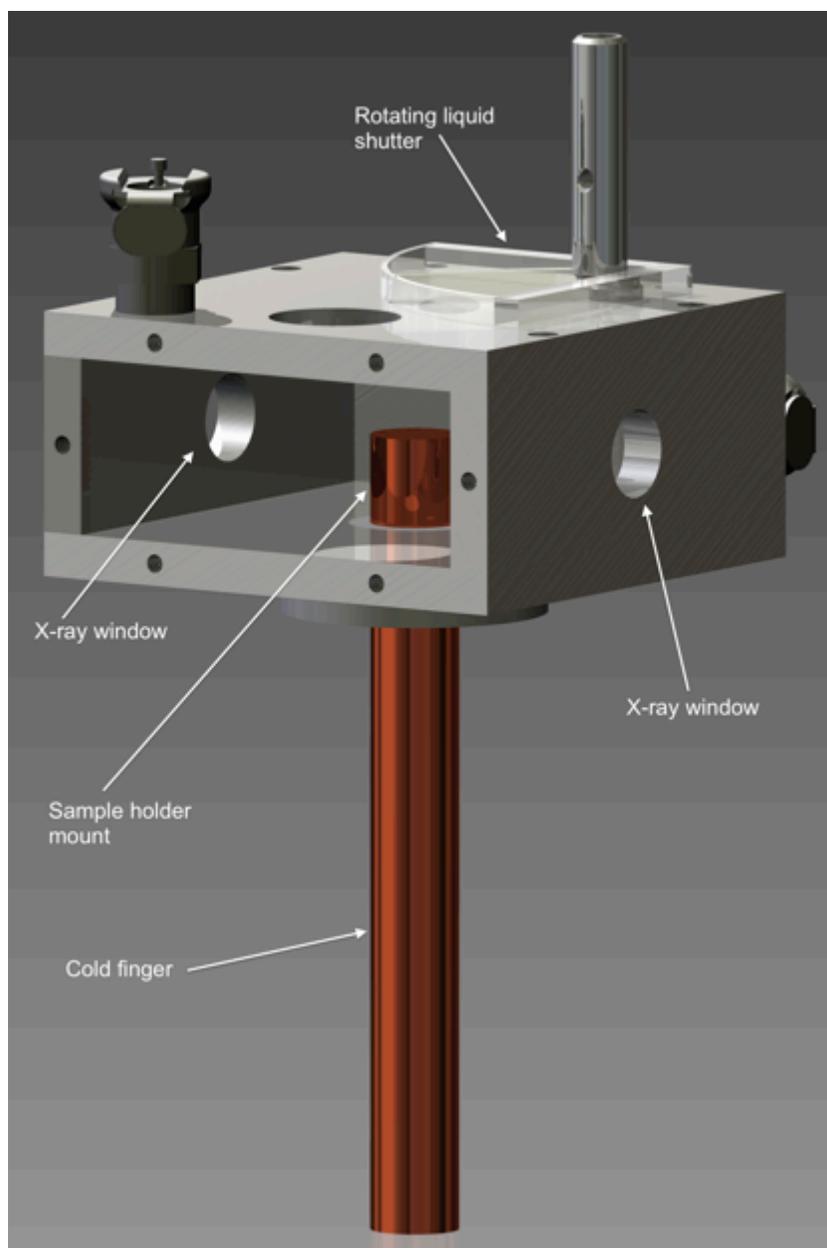


Figure 6.3 – CAD drawing of the current in-situ deposition and cooling apparatus. One panel of the dry nitrogen box has been removed to show the interior, particularly the sample holder mount, which screws into the cold finger. Not shown are interior heaters and thermocouple, the liquid nitrogen dewar the cold finger sits in, the motor that turns the rotating shutter, the needle that deposits the liquid, and the mounting bracket that connects this box to sample motors and allows the sample holder to be positioned in the X-ray beam. Image from CAD created by Andrea Katz, used with permission.

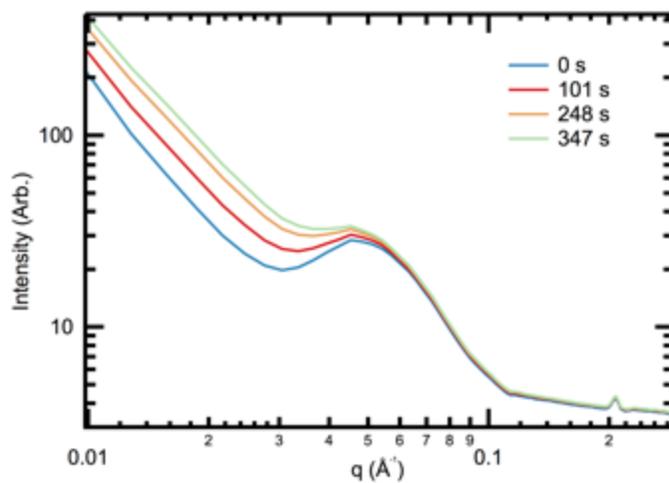


Figure 6.4 – Systematic increase in low- q scattering indicative of ice formation on the sample holder. The peaks at $q \sim 0.05 \text{ \AA}^{-1}$ and $q \sim 0.2 \text{ \AA}^{-1}$ are artifacts from the way the semi-transparent beamstop was set up, and subtract out when background subtraction is done.

One possible source of contamination is water vapor or other gases adsorbed on the gas supply line and fittings, or on the aluminum box. By choosing materials that minimize adsorption and or allow removal of water vapor and other contaminants (e.g. by baking out the aluminum) this can be (mostly) eliminated. Alternatively, sample deposition may generate a small amount of spray, or simply evaporation from the stream into the dry environment, that introduces moisture into the chamber. An additional cold finger with larger surface area (such as a cold finger with molecular sieve on it) could act as a second, stronger, cryopump to reduce or eliminate deposition on the sample holder.

6.1.3 Improvements in the sample holders

Samples frozen in the windowless sample holders used for the original work did not show fracturing. It is possible that a clever choice of geometry, materials, and size of a windowed sample holder could eliminate fractures upon cooling, and thus solve issue 4. Making such a sample holder requires development focused around simulation of cooling conditions and stresses, and iterative prototyping of a variety of different sample holder materials and geometries. If the new sample holders are more uniform than the current ones this could also solve issue 3. More details are given in Sections 2.4.5 and 2.6.1. Elimination of fracturing would also eliminate the need for comparison of nominally identical measurements, reducing sample consumption.

This redesign is already in progress by Masters students in the Thorne group. ANSYS simulations are being used to determine proper choices of geometry and materials. These students are also investigating the effect of cooling rates on fracturing. Possible designs are currently under consideration, and then prototypes will be fabricated and tested at the beam line.

6.1.4 In-vacuum cryoSAXS

In-vacuum cryoSAXS is another option for in situ deposition and cooling. A good vacuum will prevent ice formation on the sample holder, however depositing the sample becomes significantly more challenging. An in-vacuum system could also improve the performance of the windowless sample holders. In my experience, the windowless holders perform best when we have minimized the background scattering. This may be due to experimental errors in the calculation of the sample transmission, which would cause error in the subtraction of the instrument background. Minimizing the instrument background would thus minimize any error for this source. In combination with the other improvements in reproducibility already achieved, the reduction in background scattering of a sample in vacuum may improve, or solve, issue 5. Sublimation of the sample is not a concern, as it is quite slow at 100 K (Andreas, 2007).

If the sample will be cooled in-situ (in-vacuo), evaporative cooling of the sample must be minimized, and the vacuum must be maintained while liquid is deposited. Alternatively, transferring a sample into a vacuum chamber requires keeping it cold and frost-free during transfer. As sub-visual amounts of ice on the sample holder can cause unacceptable changes in the scattering, this may prove to be quite difficult. The deposition or transfer of the sample and the requirements for successful SAXS place stringent demands on the cryostat used. An appropriate system would have to be found, purchased, and likely modified, or designed and built from scratch, either of which would be costly (in time, money, or both).

Rather than in-vacuum cooling, it would be simpler to cool the sample prior to placing it in a vacuum. This would be done in a dry nitrogen environment that also includes the sample vacuum chamber. The sample is then transferred to the vacuum chamber (which is flushed with nitrogen),

the chamber is sealed, and then pumped down. However, as in 6.1.2, the cryostat and cold sample will be sensitivity to contamination in the dry nitrogen environment. Thus, the ice formation problems that plagues the in situ deposition and cooling setup above will also have to be solved for this approach.

6.2 Other questions

In addition to the above problems and possible solutions there are other areas of research that are either uniquely enabled by cryoSAXS or essential for developing and promoting cryoSAXS.

6.2.1 Time resolved cryoSAXS

Time resolved SAXS is an exciting technique capable of revealing structural changes in macromolecules on timescales from microseconds to days (Kirby & Cowieson, 2014). It provides a solution-based label-free method to study, for example, protein and RNA folding, protein allostery and structural dynamics, filament assembly, and complex formation (Mertens & Svergun, 2010; Amenitsch & Marmiroli, 2011; Kathuria et al., 2011; Pollack, 2011; Kirby & Cowieson, 2014; Chaudhuri, 2015). Successful experiments often use extreme amounts of sample, up to 50-100 mg per time point, though standard systems available for users at beamlines can use as little as 2-3 mg per time point (Graceffa et al., 2013). For comparison, a standard room temperature SAXS measurement uses ~30 µg of sample, while a typical cryoSAXS measurement uses ~0.8 µg.

One of the long-term goals of cryoSAXS is to enable routine time resolved measurements with minimal sample volume. Generally, we will use a microfluidic mixer to initiate a reaction, and then freeze cryoSAXS samples at varying times after the mixing. The current plan is to do this

using the in situ deposition system described in Section 6.1.2. Depositing from a microfluidic mixer into the cold sample holder would allow initiation of reactions in a standard way, followed by cooling and measurement at 100 K. Variation of the time of flight to the sample holder would allow measurements of different time points, while the minimum time point is set by the time of flight plus the cooling time.

In addition to needing the basic in situ system to work, there are additional challenges. A sample could either be deposited all at once, or built up slowly out of small drops. Either method may lead to slower cooling rates near the top of the sample. This would lead to a difference in the time since mixing across the sample holder, decreasing the time resolution. Depending on the velocity of the sample leaving the mixing device, large heights might be necessary for longer time points, requiring highly accurate alignment for deposition into the sample holder. It may also be hard to maintain a coherent stream or drop over a large distance. Finally, while the sample itself would be low volume, the shuttering method leads to waste. If the reaction is reversible, the volume blocked by the shutter could be recovered returned to the initial state.

Other methods may prove fruitful for time resolved cryoSAXS. One possibility is a return to windowless sample holders. Again using a conventional microfluidic mixer, sample could be ejected into a liquid cryogen. The sample could then be fished out, and mounted in the beam. This is not trivial. Ejecting drops into cryogen can be difficult, particularly for small drops, which have a tendency to float on the surface of the cryogen, yielding slower cooling times. Great care must also be taken to remove any cold gas layer above the cryogen. For small samples, most or all of the cooling can happen in the cold gas layer, resulting in much decreased cooling rates compared to those in the cryogen (Warkentin et al., 2006). Once the plunging is completed, small drops must be identified, mounted, and then placed in a cold gas stream or on a

cryostage or cold finger without any frost accumulation.

The most ambitious, and potentially rewarding, approach to time resolved cryoSAXS is to freeze the entire microfluidic mixer. This would allow an entire time sequence to be captured from a single cooling event, simply by measuring at different distances along mixer as in room temperature continuous flow time resolved SAXS. This would likely give the optimal sample use of any of the methods mentioned so far. There are two recent reports in the literature on cryocooling microfluidic mixers (Jahn et al., 2013; Mejia et al., 2014). Both approaches claim to achieve cooling rates greater than 10^4 K/s, sufficient for ~ 10 ms minimum time resolution. Either approach would need significant modification for use with cryoSAXS. One significant drawback from both methods is that sample thicknesses of ~ 15 -50 microns are required to achieve the fast cooling rates. This is 20 times shorter than the current path length used for cryoSAXS, and would require at least 20 times longer data acquisition times to compensate. However, if it could be made to work, cooling the mixer would likely be significantly more productive in the long run.

6.2.2 Hydration shells and modeling scattering at 100 K

One unresolved question for cryoSAXS is how the hydration shell of a protein changes with the addition of cryoprotectant and with cryocooling. The ability to accurately predict scattering curves is essential for comparison to crystal structures, determination of solution constituents for time resolved SAXS or other multicomponent solutions, and many other areas. Figure 2.11 shows an example of how changes in hydration shell can alter the measured scattering profiles. Understanding this behavior would provide a solution to problem 2, as it would allow disentanglement of changes in solvent scattering vs. changes in protein structure. It is also

interesting fundamental science in its own right.

This project would measure scattering from a set of proteins with different cryoprotectant and cryocooling conditions. SAXS measurements in conjunction with other methods for investigating hydration layer structure such as crystallography, small angle neutron scattering, or molecular dynamics would illuminate how the cryoprotectant affects the hydration shell at room temperature (Svergun et al., 1998; Merzel & Smith, 2002; Ball, 2008; Stanley et al., 2008; Ortore et al., 2011; Harada et al., 2012; Kim & Gabel, 2015). Using the bulk solution densities at 100 K (values currently being measured by undergraduates in the Thorne group), best guesses would be generated for hydration layers at 100 K using a variety of assumptions (no change in composition, density increase equal to the bulk solution, etc.). Predicted scattering profiles would then be generated and compared the data. The 100 K data would also be used to generate a best-fit set of parameters for the hydration shell, under the assumption that the protein structure is not changing upon cooling. These predictions would allow comparison with the model parameters, and lead to a consensus on the state of the hydration shell.

6.2.3 High profile target

Demonstration of the utility of cryoSAXS with a high profile biological target would spark interest in the technique in the larger community and lead to more excitement and adoption. What makes a good target? Something that is highly radiation sensitive, available only in small quantities, or is unstable as prepared at room temperature (or all three). For better or worse, I am not a biologist, so I currently have less insight into this area than many. However, there is expertise in this area available on campus. Consultation with biology groups that use SAXS as a technique, such as the Crane group, may yield an interesting target. Consultation with

MacCHESS, particularly Richard Gillilan, and through him the MacCHESS SAXS user community, would also be valuable.

6.2.4 Screening for proper buffer conditions

For certain proteins, the shape/scattering upon cooling to 100 K varies significantly depending on the choice of cryoprotectant. For example, xylanase in 45% w/w PEG 200 has a measured radius of gyration (R_g) at room temperature and 100 K of 16.9 Å and 17.9 Å respectively. In contrast, xylanase in 36% w/w PG has a measured R_g at room temperature and 100 K of 16.8 Å and 22.4 Å respectively. The 6% increase for Xylanase in PEG 200 is comparable to the increase observed in the R_g s upon cryocooling for glucose isomerase, 2% in PG (Hopkins et al., 2015) and 3% in PEG 200 (Meisburger et al., 2013), and for lysozyme, 9% in PEG 200, from 13.8 Å to 14.6 Å (previously unreported, and measured without a perfect background match at high q).

The observed increasing change in R_g with decreasing protein size can be understood by noting that the contribution of the hydration layer to the scattering depends on the size of the protein, assuming a fixed thickness hydration layer. The 33% increase in R_g for xylanase in PG upon cryocooling is much larger than the other observed increases mentioned above. It is not easily explainable by changes in the hydration shell, as we do not expect significant differences between the hydration shells in PEG 200 and PG. This is supported by the glucose isomerase measurements, which show almost no change between cryoprotectants. This indicates that the protein is either elongating (such as full or partial unfolding) or aggregating on cryocooling. Examination of a Porod plot of the room temperature and 100 K scattering profiles shows no indication of flexibility or disorder, such as would be expected in protein unfolding (Rambo & Tainer, 2011). Thus, the protein is most likely aggregating upon cryocooling. An upturn in a

Guinier plot at low q is characteristic of aggregation, and normally would provide verification. Our data is not repeatable at low q (see Section 2.4.4), so low q assays including Guinier plots cannot reliably be used. A similar (presumed) aggregation is seen for lysozyme in PG, as mentioned in Section 6.1.1 and Figure 6.1.

Cryocooling that does not significantly affect the protein structure is a necessary starting point for any cryoSAXS experiment. We currently have little intuition for what influences the protein during cooling. We have done some screening for appropriate buffers using cryoSAXS, and found that, in addition to choice of cryoprotectant, buffer type (e.g. HEPES, Tris, Phosphate), buffer and salt concentration, and pH can alter the 100 K scattering profile. It is well known that cooling can change the pH of a solution, and that some buffers change less with temperature than others. This has recently been exploited to create a temperature independent pH buffer, a concept that may be useful for cryoSAXS going forward (Sieracki et al., 2008).

Key to enabling use of cryoSAXS with a wide variety of molecules is development of an off-line method for determining if a buffer yields successful cryocooling. Beam time can be limited to as little as 3-4 days every 4 months, and so it is necessary to verify ahead of time that data can be successfully collected at 100 K. A successful off-line assay must either be carried out at 100 K, or requires that the sample state after cryocooling and warming correlate with the sample state at 100 K. It also must yield information that is relevant to the structures measured by cryoSAXS. Size and molecular mass are two good target quantities.

In order to develop this method, I would measure a set of proteins in a variety of buffers using cryoSAXS. Data from techniques like SEC, MALS, and DLS, and UV-Vis spectroscopy, carried out on samples that were cooled and warmed would be compared to the SAXS data to determine if the techniques can detect changes from cryocooling. This will pave the way to understanding

the behavior of samples and buffers upon cooling. Changes upon cooling may include changes in pH, salt solubility, protein solubility, or hydration as well as the possibility of cold denaturation. Much of the work for this project can be done not at the beamline, which will make progress faster.

6.3 Summary of future work

In closing, there are two necessary but divergent components to the future work mentioned above. The first is continued development of the cutting edge of the technique, by pushing to faster cooling (Section 6.1.2), time resolved cryoSAXS (Section 6.2.1), and investigation of low temperature protein hydration behavior (Section 6.2.2). The second is the improvement and regularization of the technique for non-expert users, by fabricating new sample holders (Section 6.1.3), moving to in-vacuum cryoSAXS (Section 6.1.4), demonstrating of cryoSAXS with a high profile target (Section 6.2.3), and investigation of appropriate buffer conditions for cryocooling (Section 6.2.4). Both of these paths represent necessary directions of research and development, and will require ingenuity and perseverance. However, these paths are not necessarily equally appealing for a particular researcher, or even equally appropriate for pursuit by a research group vs. a user facility. The right projects should be paired with the right people and venue for best success.

REFERENCES

- Amenitsch, H. & Marmiroli, B. (2011). *Rend. Lincei*. **22**, 93–107.
- Andreas, E. L. (2007). *Icarus*. **186**, 24–30.
- Ball, P. (2008). *Chem. Rev.* **108**, 74–108.
- Chaudhuri, B. N. (2015). *Protein Sci.* **24**, 267–276.
- 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.
- Harada, R., Sugita, Y., & Feig, M. (2012). *J. Am. Chem. Soc.* **134**, 4842–4849.
- Hopkins, J. B., Katz, A. M., Meisburger, S. P., Warkentin, M. A., Thorne, R. E., & Pollack, L. (2015). *J. Appl. Crystallogr.* **48**, 227–237.
- Jahn, A., Lucas, F., Wepf, R., & Dittrich, P. (2013). *Langmuir*. **29**, 1717–1723.
- Kathuria, S. V., Guo, L., Graceffa, R., Barrea, R., Nobrega, R. P., Matthews, C. R., Irving, T. C., & Bilsel, O. (2011). *Biopolymers*. **95**, 550–558.
- Kim, H. S. & Gabel, F. (2015). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 57–66.
- Kirby, N. M. & Cowieson, N. P. (2014). *Curr. Opin. Struct. Biol.* **28**, 41–46.
- Meisburger, S. P., Warkentin, M., Chen, H., Hopkins, J. B., Gillilan, R. E., Pollack, L., & Thorne, R. E. (2013). *Biophys. J.* **104**, 227–236.
- Mejia, Y. X., Feindt, H., Zhang, D., Steltenkamp, S., & Burg, T. P. (2014). *Lab Chip*. **14**, 3281–3284.
- Mertens, H. D. T. & Svergun, D. I. (2010). *J. Struct. Biol.* **172**, 128–141.
- Merzel, F. & Smith, J. C. (2002). *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5378–5383.
- Ortore, M. G., Mariani, P., Carsughi, F., Cinelli, S., Onori, G., Teixeira, J., & Spinozzi, F. (2011). *J. Chem. Phys.* **135**, 245103.
- Petoukhov, M. & Konarev, P. (2007). *J. Appl. Crystallogr.* **40**, 223–228.
- Pollack, L. (2011). *Biopolymers*. **95**, 543–549.
- Rambo, R. P. & Tainer, J. a (2011). *Biopolymers*. **95**, 559–571.

Sieracki, N. a, Hwang, H. J., Lee, M. K., Garner, D. K., & Lu, Y. (2008). *Chem. Commun.* 823–825.

Stanley, C., Krueger, S., Parsegian, V. A., & Rau, D. C. (2008). *Biophys. J.* **94**, 2777–2789.

Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S., & Zaccai, G. (1998). *Proc. Natl. Acad. Sci. U. S. A.* **95**, 2267–2272.

Warkentin, M., Berejnov, V., Hussein, N. S., & Thorne, R. E. (2006). *J. Appl. Crystallogr.* **39**, 805–811.

APPENDIX A

PROTOCOL FOR SAMPLE HOLDER FABRICATION

This appendix describes in more detail the protocol for producing the sample holders described in Chapter 2. Section 2.2.2 gives an overview of the procedure and describes the materials used. This supplies more details specific to the cleanroom in the Cornell Nanofabrication Facility (the CNF), the Thorne lab, and the tools available circa 2015.

Applying the nitride

1. MOS clean the wafers. Clean in the base tank for 10 minutes, then run through the automatic rinser. Clean in the acid tank for 10 minutes and run through the automatic rinser again. The acid bath's rinse tank is faster than the base bath's and can be used for both rinses.
2. Dry the wafers with compressed nitrogen, since they are too small to fit in the spin dryer.
3. Use the nitride furnace to deposit a layer of low stress nitride (LSN) about 100 nm thick (about 30 minutes of deposition time). Since there is not a properly sized set of tongs to remove the top of the three inch furnace boat, let the boat sit for a while after it is unloaded until it is cool enough to remove the top with gloved hands (note: currently the three inch furnace boat has no top so this is no longer an issue).
4. Use the Filmetrics Film Measurement System to measure the thickness of the nitride film on each wafer. This will be important when etching the nitride later on.

Photolithography

1. Clean a spinner with acetone and IPA.
2. Find a chuck smaller than the wafer and place this chuck in the spinner.
3. On the spinner computer, select the recipe 500 500 90s. This will spin at 500 rpm, will ramp the speed from 0 to 500 rpm in one second, and continue to spin for 90 seconds.
4. Use the centering tool to center the wafer on the chuck, then use the spinner computer to hold vacuum. Check the centering using the spinner computer. Recenter if necessary.
5. Start the wafer spinning, then spray with acetone followed by IPA to clean the wafer, letting each spin dry. Afterwards, use the compressed air gun in the hood to spray across the wafer while it is still spinning. This helps to remove dust particles.
6. Apply the primer P20. An amount covering about 2/3 of the wafer's surface seems good. Try to avoid bubbles. Start the spin process as quickly as possible after the application.
7. Apply NLOF 2020, a negative photoresist. An amount covering a little more than 2/3 of the wafer's surface or a little more seems to be a good amount. Try to avoid bubbles. Start the spin process as quickly as possible after the application.
8. Examine the wafer in the hood to make sure coverage is complete and reasonably uniform.
9. Bake the wafer at 115 °C for 60 seconds.
10. Exposure is done with the Suss MA6 contact aligner. Change the alignment gap to 20 μm and the exposure time to 5.5 s @ 12 W/cm^2 (*adjust time for changing exposure*)

intensities). Using topside alignment mode, line the alignment marks on the mask up with the flat of the wafer. Do an alignment check to make sure that everything looks okay, and expose the wafer.

11. Do a post-exposure bake at 115 °C for 60 seconds.
12. Develop by hand using AZ-726 MIF developer. Put a wafer into a slot on one end of a three inch wafer boat, orienting it so that the resist side will be down if the wafer boat is set on that end. Pour the developer into a beaker, then set the wafer boat on its end in the developer. Agitate slightly for two minutes, thirty seconds. Remove the wafer boat from the developer and rinse with DI water from the hood's spray hose. Dry the wafer with the hood's nitrogen gun. Pour the developer into the water developer waste bottle, and take both the wafer boat and the beaker to the dirty dish rack.
13. De-scum the wafer using the YES asher. Readjust the pins on the substrate holder so that they will fit a 3 inch wafer, then select Recipe 2. Check the settings to make sure that they agree with the recipe list on the side of the tool! Recipe 2 should have an RF power of 50 W, a temperature setpoint of 30 C, 70 sccm O₂, and should be a 60 second process.

Nitride Etching and Clean-up

1. Before etching the nitride, use a cleanroom swab to paint NLOF 2020 around the outside of the wafer, and to protect any places where there is a scratch or pinhole in the nitride. If there are any features that don't reach to the edge of the wafer, use the NLOF to connect them to the edge. This prevents erosion at the corners during the etch.
2. Bake at 95 °C for 60 seconds, so that the resist around the edge of the wafer is no longer

tacky.

3. Using the Oxford 81 etcher, do a 10 minute oxygen clean – 50 sccm O₂, 60 mTorr chamber pressure, 150 W RF power -- on the empty chamber.
4. Load wafers into the Oxford, then use the CF₄ Etch recipe – 30 sccm CF₄, 40 mTorr, 150 W RF power -- to etch away the nitride. For a nitride thickness of about 100 nm, a 2 minute etch should be sufficient. Check the most recent recipe calibration for the tool to select the best time.
5. Stop the chamber from venting after the CF₄ Etch, and instead do a 2 minute oxygen clean on the wafers.
6. Strip the photoresist after removing the wafers from the Oxford. Put the wafers in each of the hot strip baths for ten minutes, then use the automatic rinse tank to clean them. Hand-dry the wafers with the hood's nitrogen gun.
7. MOS clean the wafers again. Clean in the base tank for 10 minutes and run through the automatic rinser, then clean in the acid tank for 10 minutes and run through the automatic rinser again. The acid bath's rinse tank is faster than the base bath's and can be used for both rinses.
8. Dry the wafers with compressed nitrogen.

KOH Etching Procedure

1. Set the sonicator inside a secondary container and place both inside the fumehood. Set up the water bath, also in a secondary container, on a table next to the hood such that the desired water level in both devices is at the same height. If desired, carry out the optional

glassware cleaning step procedure below.

2. Fill water bath and sonicator. Use the pump line from the water bath to fill two siphons and get them started.
3. Turn on the sonics and heat the water bath to 60 °C. Allow the system to run for a couple of hours so that the water has time to outgas.
4. Re-start the siphons, filling with 60 °C water from the pump line.
5. Add a siphon line from the water bath to one or more additional reservoirs to help with evaporative loss. The extra reservoir(s) can be room temperature since we are not drawing much water from them.
6. Add 350-500 mL of 45% w/w KOH solution to the beaker. Add 500 nL per L of Triton X-100 ultra-grade surfactant. Suspend the beaker in the sonicator with our wire basket. Let it heat up to 60 °C.
7. Put a wafer into the cut-up three inch wafer boat so that it will be right-side up when the boat is turned on its side and placed in the KOH. Make sure that the sonics are on when you add the wafer boat and wafer to the KOH. Cover with a watch glass and tape loosely to ensure that the glass does not walk off the top of the beaker.
8. Cover everything with aluminum foil to help prevent evaporative losses
9. Etch for 16 hours.

KOH Post-Etch Procedure

1. Heat a beaker of nanopure water to 60 °C on a hotplate. Heat a second beaker of

nanopure water to 50 °C on a hotplate.

2. Remove the wafer and wafer boat from the KOH beaker and immediately place in the 60 °C beaker of nanopure water.
3. After two minutes, transfer the wafer and wafer boat to the 50 °C beaker. Heat another beaker of nanopure water to 40°C. Note: the first two rinses should be neutralized with the KOH. The rest of the rinses may be drain disposed.
4. After two minutes, transfer the wafer and wafer boat to the 40 °C beaker. Repeat this transfer ever 2 minutes, lowering the temperature of the new beakers by 10 °C each time, until room temperature is reached.
5. Continue transferring the wafer between fresh room temperature rinses until a total of 10 rinses are reached (4 heated rinses and 6 room temperature rinses).
6. Remove the wafer from the wafer boat, and gently blow dry with compressed air. If a faster/gentler dry is desired, a quick dip into pure IPA and gentle agitation will remove most of the liquid.
7. Place the wafer on a low-fiber composite wipe to finish drying.

KOH Cleanup Procedure

1. Stop the sonicator. Remove the KOH beaker from the sonicator and set it in the back of the hood, covered and out of the way, to deal with later.
2. Remove all siphons. Drain the reservoirs and sonicator. Remove the sonicator from the hood.

3. Place the KOH, the first two wafer rinses, and rinse the KOH beaker with nanopure and put that rinse water in a *large* beaker (2+ liter). Place the large beaker on a stirplate, and place a stirbar in the beaker.
4. To neutralize the KOH: Dilute it to two or three times its original volume. Add pH color indicator. Start the stirbar stirring, and set up a thermocouple to monitor the temperature. This reaction is very exothermic!
5. Wearing appropriate PPE, add undiluted HCl carefully. I recommend no more than 100 mL at a time. If it hisses/pops, dilute the KOH mixture more. If KCl salt precipitates out of solution, dilute the KOH mixture more. A total HCL volume roughly equal to the initial volume of KOH will be needed to neutralize
6. Continue to add HCl until the solution turns green. When the color starts to change towards green, add HCL dropwise to avoid any overshooting. Verify that the pH is between 5-9 (7 preferred) with litmus paper. Dropwise add KOH to return the pH to that range if the neutralization overshoots to a pH less than 5.
7. Pour the neutralized solution down the sink, followed by 20 times that volume of water (~3 minutes of running the tap at full volume).
8. Wash glassware and wipe down the inside of hood.
9. Take the wafer to the cleanroom for analysis.

Glassware Cleaning Procedure (Optional)

If you are worried about clean glassware, or want to ensure that the etch has the fewest contaminants possible, you can carry out an initial clean of the glassware prior to etching. This

procedure assumes that the glassware has already been washed.

1. Fill up three identical beakers with nanopure water (these beakers are the KOH etch beaker, and the two rinses beakers needed for the post-etch rinse). Place the first beaker in the wire basket, and place it in the sonicator.
2. Sonicate the beaker with maximum sonics. Remove the beaker after 2 minutes, dispose of the water, and place another beaker in the sonicator. Repeat with the third beaker.
3. Fill the first beaker with methanol. Sonicate for 2 minutes.
4. Remove the beaker and pour the methanol into the second beaker. Place this beaker in the sonicator. Repeat step 3 and 4 for this beaker, pouring in the third beaker.
5. Dispose of the methanol in the appropriate waste container.
6. Repeat steps 3-5 with acetone instead of methanol.
7. Repeat steps 3-5 with isopropanol instead of methanol.
8. Repeat steps 3-4 with nanopure water instead of methanol.
9. Dispose of the water, turn the beakers upside down on a low fiber wipe, and let drip dry.

In order to generate an accurate mask for the sample holders, a simple python script was written. This script outputs a text file that, when it's input into an L-Edit script file and run, generates the appropriate sample holders. More detail is available in the code comments. The full program is reproduced below.

```
#####
#This script contains a set of functions that will generate a L-Edit script file that,
#when run, generates a single parallelogram sample holder or a full line of them. There
#are two functions: "par" crates the coordinates for a single parallelogram sample holder
#while "row" creates a row of parallelograms, all connected. There are five input parameters.
#t1 is the thickness of the x-ray windows, h2 is the height of the inner parallogram (the
#pathlength of the x-rays within the sample holder), t2 is the thickness of the supporting
#side walls, t3 is the size of the inscribed square between the slanting sidewalls, and
#angle is the angle in the bottom left corner of the parallelogram. In the "row" program
#there is an additional parameter, offset, which sets the distance between adjacent sample
#holders. Units of the lengths (t1, t2, t3, h2, offset) are in mm while the angle is in degrees.

import scipy

#This program generates an L-Edit file for a single parallelogram. Inputs, in order, are t1, t2, t3,
#h2, angle. Default parameters are set for the standard sample holders.
def par(t1=0.03, t2=.056, t3=1., h2=1., angle=70.52):
    angler=scipy.radians(angle) #Convert angle to radians
    a=t2/scipy.sin(angler)+t1/scipy.tan(angler) #calculate 'a', the horizontal offset of the inner parallelogram
    x2=h2/scipy.tan(angler) #calculate 'x', the horizontal offset of the upper left vertex
    w2=t3+x2

    #make a list of the vertex positions for the inner parallelogram
    v2=[(a,t1),(x2+a,h2+t1),(w2+x2+a,h2+t1),(w2+a,t1)]

    h1=h2+2*t1
    b=t2/scipy.sin(angler)-t1/scipy.tan(angler)
    w1=w2+a+b
    x1=h1/scipy.tan(angler)

    #make a list of the vertex positions for the outer parallelogram
    v1=[(0,0),(x1,h1),(w1+x1,h1),(w1,0)]
    print 'Outer vertices:'
    for i in v1:
        print i
    print '\n'
    print 'Inner vertices:'
    for i in v2:
        print i

    f=open('./par_'+str(t1)+'.txt','w')
    f.write('layer 2')

    f.write('polygon -! %f%f%f%f%f%f%f%f\n'%(v1[0][0],v1[0][1],v1[1][0],v1[1][1],v1[2][0],v1[2][1],v1[3][0],v1[3][1]))
    f.write('polygon -! %f%f%f%f%f%f%f%f'%(v2[0][0],v2[0][1],v2[1][0],v2[1][1],v2[2][0],v2[2][1],v2[3][0],v2[3][1]))
    f.close()

    return

#This program generates a full length set of parallelograms. Inputs, in order, are t1, t2, t3,
#h2, offset, angle. Default parameters are set for the standard sample holders.
def row(t1=0.03, t2=.056, t3=1., h2=1., offset=5, angle=70.52):
    angler=scipy.radians(angle) #Convert angle to radians
    a=t2/scipy.sin(angler)+t1/scipy.tan(angler) #calculate 'a', the horizontal offset of the inner parallelogram
    x2=h2/scipy.tan(angler) #calculate 'x', the horizontal offset of the upper left vertex
    w2=t3+x2

    #make a list of the vertex positions for the left wall
    v1=[(0,0),(x2,h2),(x2+a,h2),(a,0)]

    h1=h2+2*t1 #calculate h1
    b=t2/scipy.sin(angler)-t1/scipy.tan(angler)
    w1=w2+a+b #calculate w1
    x1=h1/scipy.tan(angler) #calculate x1

    #make a list of the vertex positions for the right wall
    v2=[(w2+a,0),(w2+x2+a,h2),(w2+x2+2*a,h2),(w2+2*a,0)]
```

```

v1=scipy.array(v1)+[-(w2+x2)/2-a,-h2/2]
v2=scipy.array(v2)+[-(w2+x2)/2-a,-h2/2]
v1=v1.tolist()
v2=v2.tolist()
print 'left vertices:'
for i in v1:
    print i
print '\n'
print 'right vertices:'
for i in v2:
    print i

#calcualte box dimensions and position:
bwidth=t1
bheight=90
blcent=-t1/2-h2/2
brcent=t1/2+h2/2

print 'Box dimensions:'
print 'Width: %f' %(bwidth)
print 'Height: %f' %(bheight)
print 'Left center: %f' %(blcent)
print 'Right center: %f' %(brcent)

f=open('./row_'+str(h2)+'.txt','w')
f.write('layer 2')
for i in range(17):
    f.write('polygon -! %f%f%f%f%f%f%f%f\n' %(v1[0][0]-40.+i*offset,v1[0][1],v1[1][0]-
40.+i*offset,v1[1][1],v1[2][0]-40.+i*offset,v1[2][1],v1[3][0]-40.+i*offset,v1[3][1]))
    f.write('polygon -! %f%f%f%f%f%f%f%f\n' %(v2[0][0]-40.+i*offset,v2[0][1],v2[1][0]-
40.+i*offset,v2[1][1],v2[2][0]-40.+i*offset,v2[2][1],v2[3][0]-40.+i*offset,v2[3][1]))
    f.write('box -! %f%f%f%f%f\n' %(0+bheight/2, blcent+bwidth/2,0-bheight/2,blcent-bwidth/2))
    f.write('box -! %f%f%f%f%f\n' %(0+bheight/2, brcent+bwidth/2, 0-bheight/2, brcent-bwidth/2))

f.close()

return

```

APPENDIX B

SAXS DATA ANALYSIS PROGRAM

Overview

The goal of this Appendix is to give a brief overview of key parts and features of the custom code base established for processing SAXS data, particular radiation damage data. The main code is located on the cryoSAXS iMac at `~/Documents/cryosaxs/scripts/saxs_analysis.py`, and consists of over 7,000 lines of code (though not all of that is useful, and some of it represents similar, mostly redundant functions for dealing with data from different beamlines). This code is capable of doing a number of things, I will list some of the main features: reading in images from Pilatus detectors, and their associated header files; radially averaging these images, masking them, and normalizing them to convert to standard scattering profiles; doing normalization from semi-transparent beamstops; taking a set of images and automatically processing them to generate a set of scattering profiles, calculating various quantities from those scattering profiles (including radius of gyration, molecular weight, and maximum dimension), calculating the absorbed dose for each image, and plotting the various calculated quantities as a function of absorbed dose; calculating beam profiles and FWHM from the scan of a knife edge through the beam. There are also functions in the program dedicated to further processing of radiation damage data. There is a function that applies the dose correction to the data and functions that fit the data in various ways (most particularly the linear fits described in Chapter 3). There is also a function that calculates the correct dose based on a beam profile, and figures out the difference between that dose and the FWHM dose.

The file also contains a number of basic processing functions such as loading, scaling, adding, subtracting, averaging, binning, and calculating Guinier fits of scattering profiles, all of which are useful for producing publication quality data. It can also run GNOM on a scattering profile completely automatically for a specified range of D_{max} values. It has a program that runs a fixed number of iterations of DAMMIF on a given input file, with a multithreaded implementation which can specify the number of iterations, cores used for calculation, and various other variables. There are also other smaller utilities, such as the ability to automatically calculate the scale factor and merge SAXS and WAXS data.

The `saxs_analysis.py` file contains a large number of functions. It is designed as a library of SAXS data processing functions, which should be imported into the environment you want to run them in. Some of these functions I designed to be run interactively, sometimes asking for input, or displaying plots. Most of them I wrote to be run as parts of scripts, and take all of their input as passed variables.

Overview of basic processing

I want to provide a guide to the basic processing: reading in images and integrating them into scattering profiles. A reasonable amount of effort has gone into optimizing the speed of these functions, since the integration was initially one of the bottlenecks for processing radiation damage data. The function that loads images is called `get_img_old` (there's a new version, that should have been more general, but I never got it to work quite right). It needs the file name, data directory, and detector dimensions as inputs. It uses the `struct.unpack` command to load in the image. It returns two items, a scipy array indexed as `pic[ypixel, xpixel]`, where the value obtained at each pixel is simply the intensity. It also returns the image name. A note: this function reads in

signed integers, which means that it reads in bad Pilatus detector pixels as -1. The rest of the processing automatically ignores negative values, so there is no need to mask out bad pixels on the detector.

The next step in data collection is typically to integrate the image. This happens in two phases. The first thing that happens is an integration map is generated using the *int_map* function. This function takes inputs of the detector dimensions in pixels, the beam position (in pixels), and whatever mask you are going to use for the image. Optionally, it has parameters that allow it to apply an elliptical correction to the detector image, in case the detector was not oriented square to the incident beam. This function basically does two things. First, it generates a map of the distance of every valid pixel from the beam center, or, if such a map has been generated and saved before, it loads that map (which significantly saves on computation time). Then it generates a map of which pixels need to be ignored, due either to having negative values (bad pixels), or due to being masked out by the user. It also carries out a number of calculations that will be needed for integration of an individual image (such as the minimum and maximum distances from the beam center). For a given set of parameters (detector size, beam center, and mask) *int_map* needs to be run only once. The output contains the map of all pixel distances to the beam center, an array containing true/false values representing if a pixel should be included in the integration, a flattened version of the pixel map, an array with the number of pixels in each bin, an array with $1/\text{npix}$ for each bin ($1/\text{previous array}$), the minimum distance in the map and the maximum distance in the map.

Once the integration map has been calculated, integration can be carried out on an image. This is done using the *integrate_pic* function. This takes as input the image itself (the first output from *get_img_old*), all of the output from the *int_map* function (as a list), the pixel size, sample to

detector distance, and wavelength for calibrating the q values. Because most of the calculation has been already done by the *int_map* function this function is quite fast, which is good for integrating large numbers of images that have the same parameters. Basically the function uses an obscure scipy histogramming function, *scipy.bincount*, where the bins are the distance from the detector. Given the map of distances, it calculates a histogram of the number of pixels with a given distance, using the values of those pixels as weights. There's a similar calculation for the error in the intensity, which calculates the standard error (rather than the standard deviation). The function then calculates the q values associated with the intensity, and returns 3 arrays, q , I , and the error in I .

Overview of radiation damage processing

There are a number of different functions that do radiation damage processing. All of these are wrapped up into the *rad_dam* function, which is designed to run automatically on a large data set. Note that there are multiple *rad_dam* functions, which extended names like *rad_dam_biocat* which are just variants set up to handle the slightly different formats of data from different beamlines. The basic *rad_dam* function is designed to run on data from G1 at CHESS. The program is designed to run on a set of images that are from the same sample that is being damaged. That is, this generates data assuming all images are from the same sample, and it is being sequentially dosed every image.

The function takes a huge number of inputs. The most basic ones are the file prefix, the starting and ending run number (can be the same), a mask, and all of the necessary calibration data. It also needs data location, header location, and save location, detector size, what counter to normalize by, what counter to use for calculating the flux for a given image, other parameters for

calculation dose such as the sample path length. It then has a set of optional parameters. These let you set it up to use a semitransparent beamstop for normalization, apply an elliptical correction to the detector, skip points at the beginning and end of the data, pick how many items get automatically plotted, and a few other things. The best way to learn to use this program is twofold. Going through the code is good. Also, the script (in the same directory) called *2015_mar06_chess_rd_sub.py* has a detailed setup and commenting of how to run this function on actual experimental data.

The program also includes the possibility of doing minimal post-processing of the scattering profiles. This can be done if an appropriate background is not available (for example, looking at damage in a buffer, rather than a protein solution), or simply for speed.

The basic outline of how the program works follows. 1) Load in background file for subtraction. 2) Get a list of the files to be processed. 3) Load each file, generate the corresponding subtracted scattering profile. 4) For each scattering profile, process it to get r_g , molecular weight, maximum dimension, and other parameters (using a variety of methods). 5) Calculate the dose for each frame. 6) Save the data, and automatically generate plots showing the change in scattering profile with dose, change in the various calculated parameters with dose.

Once this data is generated, additional processing may be desired. One possibility is a diffusion correction (calculated in mathematica) may wish to be applied. This can be done with the *dose_cor* function. Curves may need to be fit, the *rd_lin_fit* family of functions will do this. For an example of the linear fitting, see the *2015_mar06_chess_rt_rd_linfit.py* file in the scripts folder.

Final notes

This is not intended to thoroughly document the program. Instead, it is intended to provide some insight into the primary functionality of the program, and how to do some of the more common operations. The code is commented (sometimes extensively, sometimes minimally), and should be mostly readable. The best way to learn it is by reading through the code and trying it out, and modifying it as need be. Hopefully this provides a starting point, so that in the future this code base does not need to be completely rewritten to handle SAXS processing. The scripts folder contains a number of examples of most of the functions in use, so looking through the processing for particular data sets there is a good place to start learning this toolset. While I certainly hope that the code is mostly bug free, I cannot guarantee that. It has gone through extensive testing, and outputs have been compared to those of other conventional programs (particularly for the generation of scattering profiles). It is also rather fragile code, it has very little error checking and handling (more in the more automated functions, such as the *rad_dam* set). This means that whenever the input is changed a bit (data from a different beamline, header in a new format, etc) some modification is usually needed to make the program work well with the data.

The use of BioXTAS RAW, Scatter, or the ATSAS programs are recommended for simple processing, graphical interfaces, etc. I also want to note that I've written a set of functions that allow you to import settings and masks from RAW (which is what we typically use at the G1 beamline for online processing). These are the *load_raw_...* functions. This is particularly useful if you want to use the RAW click and drag masking functionality rather than the coordinate based masking functionality included in this set of programs.

Finally, if you want to display images, a few options exist. *plot_pic* and *plot_tiff* are functions which take either *pic* arrays (from the *get_img_old* function) or tiffs appropriately (mins and

maxes can be modified). These are convenient because you can use them as a basis and then use the other python plotting functions to add to the images, such as easily overlaying masks, showing resolution rings, etc. For independent viewers, ImageJ works, as does Albula, which is provided by Dectris specifically for viewing Pilatus images.