

LA TROBE UNIVERSITY  
VICTORIA, AUSTRALIA

PHD THESIS

---

**Understanding and Characterising the Effects  
of Radiation Damage in Protein  
Micro-Crystallography**

---

*Author:*

Hugh MARMAN

Bachelor of Science (Honours)

Diploma of Music Industry (Technical Production)

*A thesis submitted in total fulfilment of the requirements  
for the degree of Doctor of Philosophy*

*in the*

XFEL Group

Department of Mathematical and Physical Sciences

January, 2022

# Declaration of Authorship

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgement in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Signed: Hugh Marman

---

Date: 19/01/2022

---

*“The stubborn critic would say: ‘What is the benefit of these sciences?’ He does not know the virtue that distinguishes mankind from all the animals: it is knowledge, in general, which is pursued solely by man, and which is pursued for the sake of knowledge itself, because its acquisition is truly delightful, and is unlike the pleasures desirable from other pursuits. For the good cannot be brought forth, and evil cannot be avoided, except by knowledge. What benefit then is more vivid? What use is more abundant?”*

Al-Bīrūnī, 973–1048 CE

# *Abstract*

## **Understanding and Characterising the Effects of Radiation Damage in Protein Micro-Crystallography**

by Hugh MARMAN

Macromolecular x-ray crystallography (MX) is the premiere technique for determining the structure of proteins, a critical step in the understanding of biological function, including diseases and their treatment. Radiation damage is a fundamental limiting factor in protein structure solution via MX. The majority of MX experiments are conducted at synchrotrons, where continual technological development means that the conditions to which a protein crystal is routinely exposed change regularly, and the impact of radiation damage must be continually reassessed. Of particular interest in recent years are the growing use of micro- and nano-crystals in serial crystallography techniques, the impact of incredibly high flux-density Gaussian micro-beams, and the potential use of higher beam energies in MX experiments. Here we study the relationship between these factors as they pertain to the progression of radiation damage in MX, with the aim of adapting experimental techniques to take advantage of the properties of micro-crystals, and the technological capabilities of MX beamlines.

Based on fundamental theory of x-ray—matter interaction and the transport of electrons in solid matter, we simulate, with fine spatial- and temporal-resolution, the impact of sample- and beam-parameters on the progression of radiation damage. These simulations predict significant reduction in damage-rate for beam energies above 20 keV, when compared to standard MX energy ( $\approx 12$  keV). Experiments with native lysozyme micro-crystals at a variety of x-ray beam energies provide results that are consistent with a reduction in damage at higher energy, and highlight the technical and procedural challenges of operating a modern MX beamline in the regime required to take advantage of these effects. An in depth meta-analysis of modern radiation damage studies as reported in the literature firmly contextualises the work presented here within the broader body of scientific knowledge. The techniques and results presented here can aid in the planning of new experiments that take advantage of the properties of micro-crystals and modern synchrotron technology to mitigate the effects of radiation damage and improve data quality.

## *Acknowledgements*

I am grateful to my primary supervisors, Assoc. Prof. Brian Abbey and Dr. Connie Darmanin for their consistent attention, advice, and encouragement through the process of researching and writing this thesis. Both provided access to a depth of knowledge and Connie single-handedly presided over my induction into the arcane mysteries and dark arts of the biochemistry laboratory. My co-supervisor Dr. Eugeniu Balaur assisted on many matters technical, both in conversation and by provision of custom-made experimental apparatus. Special thanks is also due to Dr. Hannah Coughlan for her advice and support early in my candidature.

This work was supported by an Australian Government Research Training Program Scholarship, a La Trobe University Postgraduate Research Scholarship, and an additional stipend from the Australian Research Council Centre of Excellence in Advanced Molecular Imaging. The support of the latter also included the opportunity to attend an annual summit with researchers across multiple disciplines and institutions, with benefits social, academic, and professional.

Many thanks to my collaborators, both within the Centre of Excellence and outside it, in particular Elspeth Garman, Joshua Dickerson, Nadia Zatsepin, Martin Fuchs, Sarah Perry, Grant Mills, Harry Quiney, Alex Kozlov, and Alaric Sanders.

The Physics community at La Trobe were a source of community and conversations in the tea-room and corridors, both technical and social, were an important day-to-day feature of the last few years. Many thanks, in particular, to Susannah Holmes, Steve Yianni, Xinxin Peng, Peter Berntsen, Trey Guest, Nick Phillips, Panji Achmani, Narelle Brack, Grant van Riessen, and Marjan Hadian-Jazi.

## *Statement of Contribution*

Parts of the research contributing to this thesis were performed as part of a research group. The purpose of this statement is to render explicit the contributions of the myself and my collaborators.

I am personally responsible for the drafting of the entire manuscript, including the previously-published paper that forms the core of chapter 3. I have also produced all graphics used in the thesis, except those explicitly acknowledged as the work of others.

The whole manuscript has been reviewed by my principle supervisors Brian Abbey and Connie Darmanin, who have made suggestions for edits to content and form. Chapter 4 has also been reviewed by Elspeth Garman, who likewise contributed suggestions for its improvement.

Chapters 1, 2, and 6 represent the introduction, theoretical background and literature review, and conclusion of this thesis. They have been written (and researched) in their entirety by me.

Chapter 3 presents a simulation of radiation damage. I am responsible for the coding of the simulation, its execution, the analysis of results, and the resulting manuscript. Hananah Coughlin provided an SOP for the crystallisation of lysozyme that I used to script a model of a crystalline sample as part of the simulation.

Chapter 4 presents two experiments undertaken at the P11 beamline at PETRA-III and the 17-ID-2 (FMX) beamline at NSLS-II. Dominik Oberthür provided an SOP for the crystallisation of lysozyme that was used to crystallise the samples for both experiments presented in this chapter. I adapted the SOP, performed all crystallisation trials, and produced the samples for the experiments. TEM characterisation of the crystals, including preparation, was performed by Connie Darmanin and myself, with the assistance of Julian Ratcliffe.

The experiment team at PETRA-III consisted of myself, Connie Darmanin, and Grant Mills. All three members of the team contributed to sample preparation and loading, and monitored data collection. Decisions about managing the experiment goals and timeline were made by myself and Connie Darmanin. The silicon chips used for sample mounting were fabricated at the Melbourne Centre for Nano-fabrication. Grant Mills provided some Python scripts to assist in the early stage of data analysis, which I conducted myself, with feedback from Connie Darmanin and Brian Abbey. Some of the analysis was conducted remotely using the DESY MAXWELL computing cluster.

The experiment at NSLS-II was undertaken as part of a block allocation group (BAG), consisting of researchers from the 17-ID-2 FMX Beamline (Martin Fuchs and Wuxian Shi), La Trobe University (myself, Connie Darmanin, Brian Abbey, and Nadia Zatsepin), Oxford University (Joshua Dickerson and Elspeth Garman), and UMass Amherst (Sarthak Saha, Shuo Sui, and Sarah Perry). The experiment was co-conceived by the BAG. I took a leading role in the planning and execution of the experiment, with chief support from Joshua Dickerson, Martin Fuchs,

Brian Abbey, and Elspeth Garman. I was also responsible for all sample preparation. The beamline staff were responsible for the configuration and running of the beamline, and provided facility-related support (e.g. with biochem labs, etc). The Perry lab provided the polymer chips used for sample delivery and adapted them *in situ* to the experimental apparatus. The experiment was performed with the aid of an Eiger 1M-CdTe detector on loan from Dectris.

I conducted the analysis of the NSLS-II experiment, with periodic feedback provided by Connie Darmanin, Brian Abbey, Joshua Dickerson, Elspeth Garman, and Martin Fuchs.

I am solely responsible for the modelling and analysis presented in chapter 5, which includes a meta-analysis of several studies from the literature.

## Publications

**H. Marman**, C. Darmanin, and B. Abbey, 'The Influence of Photoelectron Escape in Radiation Damage Simulations of Protein Micro-Crystallography', *Crystals* (2018).

M. Hadian-Jazi, P. Berntsen, **H. Marman**, B. Abbey, and C. Darmanin, 'Analysis of Multi-Hit Crystals in Serial Synchrotron Crystallography Experiments Using High-Viscosity Injectors', *Crystals* (2021).

D. J. Wells, P. Berntsen, E. Balaur, C. M. Kewish, P. Adams, A. Aquila, J. Binns, S. Boutet, H. Broomhall, C. Caleman, A. Christofferson, C. E. Conn, C. Dahlgqvist, L. Flueckiger, F. Gian-Roque, T. L. Greaves, M. Hejazian, M. Hunter, M. Hadian-Jazi, H. O. Jönsson, S. Kadaoluwa, R. Kirian, A. Kozlov, R. Kurta, **H. Marman**, D. Mendez, A. Morgan, K. Nugent, D. Oberthür, H. Quiney, S. Saha, J. A. Sellberg, R. Sierra, M. Wiedorn, B. Abbey, A. V. Martin, and C. Darmanin, 'Observations of phase changes in monoolein during high viscous injection', *Journal of Synchrotron Radiation*, submitted (2021).

S. Holmes, H. J. Kirkwood, R. Bean, K. Giewekemeyer, **H. Marman**, A. V. Martin, M. Hadian-Jazi, M. O. Wiedorn, D. Oberthür, L. Adriano, N. Al-Qudami, S. Bajt, I. Barák, S. Bari, J. Bielecki, S. Brockhauser, M. A. Coleman, F. Cruz-Mazo, C. Danilevski, K. Dörner, A. M. Gañán-Calvo, R. Graceffa, H. Fangor, M. Heymann, M. Frank, A. Kaukher, Y. Kim, B. Kobe, J. Knoška, T. Laurus, R. Letrun, L. Maia, M. Messerschmidt, M. Metz, T. Michelat, G. Mills, S. Molodtsov, D. C. F. Monteiro, A. Münnich, G. Peña-Murillo, G. Previtali, A. Round, T. Sato, R. Schubert, J. Schulz, M. Shelby, C. Seuring, J. A. Sellberg, M. Sikorski, A. Silenzi, S. Stern, J. Sztuk-Dambietz, J. Szuba, M. Trebbin, P. Vagovic, T. Ve, B. Weinhausen, K. Wrona, P. L. Xavier, C. Xu, O. Yefanov, K. A. Nugent, H. Chapman, A. Mancuso, A. Barty, B. Abbey, and C. Darmanin, 'Megahertz pulse trains enable multi-hit serial crystallography experiments at XFELs', *Nature Communications*, submitted (2021).

# Contents

<b>Declaration of Authorship</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Statement of Contribution</b>	<b>v</b>
<b>Publications</b>	<b>vii</b>
<b>List of Figures</b>	<b>xii</b>
<b>List of Tables</b>	<b>xiv</b>
<b>List of Abbreviations</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Protein Structure . . . . .	1
1.1.1 The Protein Data Bank & UniProt . . . . .	2
1.2 Protein Crystallography . . . . .	3
1.3 Radiation Damage in X-ray Crystallography . . . . .	4
1.4 Summary of This Thesis . . . . .	6
1.4.1 Supporting Material . . . . .	6
1.4.2 Software Referenced in This Thesis . . . . .	7
<b>2 Background &amp; Literature Review</b>	<b>8</b>
2.1 X-rays & Their Interactions with Matter . . . . .	8
2.1.1 Transmission, Scattering, and Absorption . . . . .	8
2.1.2 Interaction Cross Sections . . . . .	10
2.2 Macromolecular Crystallography . . . . .	12
2.2.1 Crystals . . . . .	12
2.2.2 X-ray Diffraction by Crystals . . . . .	16
2.2.3 The Reciprocal Lattice . . . . .	18
2.2.4 Crystallography & Structure Determination . . . . .	21
2.2.5 Protein Crystals . . . . .	24
2.2.6 Synchrotron Light Sources . . . . .	26

2.2.7	X-ray Free Electron Lasers . . . . .	32
2.2.8	Serial Crystallography . . . . .	33
2.3	Radiation Damage in Synchrotron MX . . . . .	34
2.3.1	Dose . . . . .	35
2.3.2	Radiation Damage Metrics . . . . .	36
2.3.3	Tolerable Dose . . . . .	37
2.3.4	Empirical Models of Spot Fading . . . . .	39
2.3.5	Mitigation of Radiation Damage . . . . .	40
2.3.6	The Role of Photoelectrons in Radiation Damage . . . . .	41
2.4	Summary . . . . .	45
<b>3</b>	<b>A Spatially- and Temporally-Resolved Model of Radiation Damage in Macromolecular Crystallography</b>	<b>47</b>
3.1	Introduction . . . . .	47
3.2	Simulation Design . . . . .	49
3.2.1	Top-Level Structure . . . . .	50
3.2.2	Monte Carlo Electron Trajectory Model: CASINO . . . . .	50
3.2.3	Space-Resolved Dose Mapping . . . . .	51
3.2.4	Time-Resolved Spot Fading . . . . .	55
3.3	Simulation Execution . . . . .	56
3.3.1	Simulating a Crystal Sample: Chemical Composition & Weighted Cross Sections . . . . .	57
3.3.2	CASINO Execution . . . . .	58
3.3.3	Space-Resolved Dose Map Execution . . . . .	59
3.3.4	Time-Resolved Spot Fading Execution . . . . .	59
3.4	Paper . . . . .	59
3.5	Summary of Results & Discussion . . . . .	81
3.5.1	Effect of Beam/Crystal Size . . . . .	81
3.5.2	Effect of Beam Energy . . . . .	81
3.5.3	Diffraction Efficiency at Higher Beam Energies . . . . .	82
3.5.4	Evolution of the Diffraction Volume . . . . .	82
3.6	Additional Discussion . . . . .	82
3.7	Summary . . . . .	84
<b>4</b>	<b>Radiation Damage Experiments Using Protein Micro-Crystals</b>	<b>86</b>
4.1	Introduction . . . . .	86
4.2	Materials and Methods . . . . .	90
4.2.1	Sample Preparation & Characterisation . . . . .	90
4.2.2	X-ray Data Collection at PETRA-III . . . . .	96
4.2.3	X-ray Data Collection at NSLS-II . . . . .	100
4.3	Results . . . . .	106
4.3.1	PETRA-III Results . . . . .	106
4.3.2	NSLS-II Results . . . . .	107
4.4	Discussion . . . . .	110

4.4.1	Context . . . . .	111
4.4.2	Experimental Design . . . . .	112
4.4.3	PETRA-III Results: Overview . . . . .	114
4.4.4	NSLS-II Results: Overview . . . . .	116
4.4.5	NSLS-II Results: Discrepancy Between Experimental Intensities and Computational Predictions . . . . .	117
4.4.6	NSLS-II Results: The Influence of Beam Energy on Bragg Spot Fading . . . . .	121
4.4.7	Considerations For Future High Energy Diffraction Experiments . . . . .	123
4.5	Summary . . . . .	125
<b>5</b>	<b>Comparison of Spot Fading Models</b>	<b>127</b>
5.1	Introduction . . . . .	127
5.2	Background . . . . .	129
5.2.1	Modelling the Impact of Radiation Damage on Diffracted Intensity . . . . .	129
5.2.2	MX at Different X-Ray Energies . . . . .	132
5.3	Methods . . . . .	133
5.3.1	A Predictive Version of the Atakisi Model . . . . .	134
5.3.2	Collection & Analysis of Primary Data . . . . .	135
5.3.3	Re-analysis of Previous Studies . . . . .	136
5.4	Results . . . . .	136
5.4.1	Shell Half Doses & Model A . . . . .	136
5.4.2	Micro-Crystallography at Cryogenic Temperature and Standard MX X-Ray Energy . . . . .	137
5.4.3	Micro-Crystallography at Cryogenic Temperature and High X-Ray Energy . . . . .	141
5.4.4	Re-analysis of Previous Studies . . . . .	144
5.5	Discussion . . . . .	146
5.5.1	Setting $\kappa$ and $\alpha$ in Model A . . . . .	146
5.5.2	Comparison of Model A & Model H in Reference to Our Experimental Data Sets . . . . .	147
5.5.3	Comparison of Model A & Model H in Reference to Prior Spot Fading & Shell Half Dose Studies . . . . .	149
5.5.4	Model H & Integrated Half Dose Values . . . . .	157
5.6	Summary . . . . .	159
<b>6</b>	<b>Thesis Summary &amp; Further Work</b>	<b>162</b>
6.1	Summary of Work & Results of This Thesis . . . . .	163
6.1.1	Simulations Predict Significant Damage Reduction to Micro-crystals Due to Photoelectron Escape . . . . .	163
6.1.2	Early Experiments Highlight Technical Challenges in Micro-crystallography at Higher Beam Energies . . . . .	164
6.1.3	Meta-Analysis Demonstrates Capabilities & Limitations of Predictive Models of Radiation Damage . . . . .	166
6.2	Suggestions for Further Work . . . . .	166
6.2.1	Extended Experimental Study of Radiation Damage at High Energy . . . . .	167

6.2.2	Reducing the Photoionisation Cross Section of the Solvent . . . . .	168
6.2.3	Adapting Current Models for Room Temperature Crystallography . . . . .	168
<b>Appendix A Example RADDPOSE-3D Input File</b>		<b>170</b>
<b>Appendix B Spot Fading Meta-Analysis: Additional Plots</b>		<b>172</b>
<b>Appendix C Image Permissions</b>		<b>178</b>
C.1	Images Used Under the Auspice of a Creative Commons License . . . . .	178
C.2	Images Used by Permission of the Owner . . . . .	179
<b>Bibliography</b>		<b>180</b>

# List of Figures

2.1	X-ray-matter interactions . . . . .	9
2.2	Selected elemental x-ray interaction cross sections . . . . .	11
2.3	Crystal structure . . . . .	13
2.4	The unit cell & crystal lattice . . . . .	15
2.5	Relationship between the diffraction of a molecule and that of a molecular crystal .	17
2.6	Crystal planes & Miller indices . . . . .	18
2.7	Bragg's law. . . . .	19
2.8	The Ewald sphere . . . . .	20
2.9	Structure and crystallisation of lysozyme: A model protein . . . . .	25
2.10	PDB structure statistics . . . . .	29
2.11	Diffraction patterns of 1 <sup>st</sup> and 3 <sup>rd</sup> generation synchrotrons . . . . .	30
2.12	MX beamline optics. . . . .	31
2.13	Serial femtosecond crystallography. . . . .	35
2.14	Radiation damage & loss of diffracted intensity . . . . .	37
2.15	Electron inelastic mean free path & penetration depth. . . . .	43
3.1	Radiation damage simulation: Flow diagram . . . . .	50
3.2	Model of the energy deposition due to a photoionisation event . . . . .	53
4.1	Detector quantum efficiency & diffraction efficiency . . . . .	90
4.2	Lysozyme crystals: Optical microscopy . . . . .	93
4.3	Size distribution in crystallisation trials . . . . .	94
4.4	Lysozyme crystals: Transmission electron microscopy . . . . .	95
4.5	Silicon sample chips . . . . .	97
4.6	PETRA-III sample mounting . . . . .	98
4.7	Polymer sample chips . . . . .	101
4.8	NSLS-II sample loading . . . . .	102
4.9	NSLS-II sample mounting . . . . .	103
4.10	PETRA-III results: Relative intensity as a function of time . . . . .	106
4.11	PETRA-III results: Bragg diffraction . . . . .	107
4.12	NSLS-II results: Bragg diffraction . . . . .	108
4.13	NSLS-II results: Relative intensity as a function of dose . . . . .	109
5.1	Blake-Phillips model of crystal damage. . . . .	130

5.2	Meta-analysis of shell half dose. . . . .	137
5.3	Spot fading by resolution shell, 12.66 keV. . . . .	139
5.4	Loss of integrated intensity, 12.66 keV. . . . .	140
5.5	Spot fading by resolution shell, 26.00 keV. . . . .	142
5.6	Loss of integrated intensity, 26.00 keV. . . . .	143
5.7	Two models of spot-fading compared to previously published data. . . . .	145
5.8	The Howells-Holton model & integrated intensity. . . . .	158
B.1	Model A & Model H meta-analysis: Bourenkov, FAE. . . . .	173
B.2	Model A & Model H meta-analysis: Bourenkov, FtsH. . . . .	174
B.3	Model A & Model H meta-analysis: Bourenkov, p19-siRNA-1. . . . .	175
B.4	Model A & Model H meta-analysis: Warkentin. . . . .	176
B.5	Model A & Model H meta-analysis: Liebschner. . . . .	177

# List of Tables

1.1	Software versions . . . . .	7
2.1	Bravais lattices . . . . .	14
3.1	Simulated sample parameters . . . . .	57
4.1	Lysozyme crystallisation trial . . . . .	91
4.2	PETRA-III beamline parameters . . . . .	99
4.3	NSLS-II beamline parameters . . . . .	103
4.4	NSLS-II half dose . . . . .	110
5.1	Resolution shells . . . . .	136
5.2	Spot fading meta-analysis samples and resolutions . . . . .	144
5.3	Integrated half dose values & the Howells-Holton model of spot fading . . . . .	159

# List of Abbreviations

<b>ANL</b>	Argonne National Laboratory
<b>BCDI</b>	Bragg Coherent Diffractive Imaging
<b>BM</b>	Bending Magnet
<b>BNL</b>	Brookhaven National Laboratory
<b>CCDC</b>	Cambridge Crystallography Data Centre
<b>CSDA</b>	Continuous Slowing Down Approximation
$D_{1/2}$	Half Dose
<b>DE</b>	Diffraction Efficiency
<b>DESY</b>	Deutsches Elektronen Synchrotron
<b>DWD</b>	Diffraction Weighted Dose
<b>EG</b>	Ethylene Glycol
<b>EM</b>	Electron Microscopy
<b>EMBL-EBI</b>	European Bioinformatics Institute
<b>FEL</b>	Free Electron Laser
<b>FWHM</b>	Full Width at Half Maximum
<b>HEWL</b>	Hen Egg-white Lysozyme
<b>ID</b>	Insertion Device
<b>IMFP</b>	Inelastic Mean Free Path
<b>LCLS</b>	Linac Coherent Light Source
<b>LINAC</b>	Linear Accelerator
<b>MBA</b>	Multi-Bend Achromat
<b>MC</b>	Monte Carlo
<b>MCN</b>	Melbourne Centre for Nanofabrication
<b>MX</b>	Macromolecular X-ray Crystallography
<b>NMR</b>	Nuclear Magnetic Resonance
<b>NSLS-II</b>	National Synchrotron Light Source
<b>PAD</b>	Pixel Array Detector
<b>PDB</b>	Protein Data Bank
<b>PE</b>	Photoelectron
<b>PEG</b>	Polyethylene Glycol
<b>PIR</b>	Protein Information Resource
<b>QE</b>	Quantum Efficiency
<b>RCSB</b>	Research Collaboratory for Structural Bioinformatics
<b>RT</b>	Room Temperature
<b>SASE</b>	Self Amplified Stimulated Emission

<b>SEM</b>	Scanning Electron Microscopy
<b>SFX</b>	Serial Femtosecond X-ray Crystallography
<b>SIB</b>	Swiss Institute of Bioinformatics
<b>SNR</b>	Signal-to-Noise Ratio
<b>SOP</b>	Standard Operating Protocol
<b>SSX</b>	Serial Synchrotron X-ray Crystallography
<b>TEM</b>	Transmission Electron Microscopy
<b>UniProt</b>	Universal Protein Resource
<b>XFEL</b>	X-ray Free Electron Laser
<b>XRD</b>	X-ray Diffraction

#### Chemical Abbreviations

---

<b>CdTe</b>	Cadmium-Telluride
<b>HCl</b>	Hydrogen Chloride (Hydrochloric Acid)
<b>NaCl</b>	Sodium Chloride
<b>NaOH</b>	Sodium Hydroxide

# Chapter 1

## Introduction

Proteins are complex biological macromolecules that are responsible for myriad biological processes. Proteins are divided into several groups or families, based on their conformational and ancestral similarity, and their function. *Enzymes* catalyse biochemical reactions, such as breaking molecules into smaller components as part of metabolism or separating DNA helices for transcription; *hormones* act as inter-cellular messengers; *collagen*, *keratin*, and *elastin* provide structure to cells; *antibodies* bind to foreign particles like bacteria and viruses; and other classes of proteins are involved in nutrient transport, organ function, fluid- and pH-balance within the body, and more. Drugs must interact with proteins in order to inhibit or supplement their function, and to gain entrance to cells. Understanding the roles played by proteins in the body is a Herculean undertaking that has been the subject of significant research effort for well over a century and has important ramifications for maintaining a healthy society. It is the aim of *structural biology* to understand protein function via its strong correlation to composition and form [1–3]. Macromolecular x-ray crystallography is the primary technique employed in structure determination. One of the fundamental challenges faced by protein crystallographers is radiation damage. As the standard protocols of crystallography experiments evolve, a continuous reassessment of the impact of radiation damage is essential.

### 1.1 Protein Structure

Proteins are sequences of amino-acid residues that fold in complex ways to adopt specific three-dimensional structures when in the appropriate environment (dictated by temperature, pH, electrodynamic properties, etc). Amino acids share a common backbone structure, with each amino acid having a unique side chain attached to this backbone. Amino acids connect to each other along the backbone through peptide bonds. This process results in the elimination of a water

molecule, leaving each amino acid slightly reduced, hence the term 'residue'. Protein structure is defined in a four-tier hierarchy:

- *Primary structure* is the specific sequence of amino-acid residues that are combined to form a protein;
- *Secondary structure* describes the folding of sections of the amino-acid chain into sub structures such as  $\alpha$ -helices and  $\beta$ -sheets. These assemblies are the result of backbone-backbone interactions;
- *Tertiary structure* describes the full three-dimensional shape of an individual protein molecule, or *monomer*. The additional folding leading to this structure is caused by weak electromagnetic interactions between the side chains of the protein;
- *Quaternary structure* describes the assembly of multiple monomers (of the same or different type) into larger protein complexes.

Once folded, a protein (or protein complex) will have a particular size and shape, with particular residues exposed to the surface of the protein. Part of protein will form a *binding pocket/active site*, the location of which can be inferred from its shape. The residues of active site will elucidate the binding mechanism and define the protein's biological function. This is the structure-function correlation: a detailed understanding of the three-dimensional structure of a protein can reveal what it does based on the molecule(s) to which it will bind and how the residues of the active site will then interact with the bound molecule(s).

### 1.1.1 The Protein Data Bank & UniProt

The *Protein Data Bank* (PDB) [4] is a database maintained by the *Research Collaboratory for Structural Bioinformatics* (RCSB), where three-dimensional structures of proteins, nucleic acids, and complex assemblies are deposited by researchers. Its archives contain over 1 TB of data that is publicly accessible. In 2021, the PDB celebrated 50 years of open access to structure information. As of October 2021, there were 182 949 structures documented on the PDB, of which 159 817 were isolated proteins (as opposed to proteins with bound ligands, or protein complexes) [5].

The *Universal Protein Resource* (UniProt) [6], maintained in collaboration between the *European Bioinformatics Institute* (EMBL-EBI), the *Swiss Institute of Bioinformatics* (SIB) and the *Protein Information Resource* (PIR), is a similar project to the PDB, but its database contains protein *sequence* data, rather than structure data. As of October 2021, there were 565 254 sequences on Swiss-Prot, UniProt's repository of curator-reviewed sequences; on TrEMBL, which contains yet-to-be-reviewed sequences with limited automatic annotations, there were 219 174 961 entries! The

disparity between the number of known sequences and potential sequences on UniProt, and the number of structures on the PDB highlights the extent of the task still facing structural biologists.

## 1.2 Protein Crystallography

A detailed understanding of protein function typically requires its structure to be known to atomic resolution. Such resolution (on the order of  $1 \text{ \AA}$ , or  $1 \times 10^{-10} \text{ m}$ ) is four orders of magnitude smaller than the wavelengths of visible-light. Protein structures are therefore imaged using techniques like nuclear magnetic resonance (NMR), electron microscopy (EM), and x-ray diffraction (XRD). X-ray diffraction is currently the most successful technique, responsible for 87.6% of the structures in the PDB [5].

X-rays interact weakly with matter, particularly with the light elements such as carbon, oxygen, and nitrogen that are the majority constituents of proteins. Of those x-ray photons that do interact with a sample, many will undergo inelastic interactions, including being absorbed, that typically result in damage rather than contributing to x-ray diffraction. In order to acquire sufficient diffraction data to obtain a protein structure, it is necessary to combine the diffraction signal from many hundreds of thousands of proteins, in order to spread the damage between them. This is accomplished via macromolecular x-ray crystallography (MX): a protein crystal is produced by tailoring the protein's environment in such a way as to cause individual molecules to arrange themselves in a regular, periodically repeating fashion (a crystal); this crystal then produces a diffraction signal that is many hundreds of thousands of times stronger than that of a single protein.

There are several potential bottlenecks in the process of structure determination via MX, the first of which is obtaining a sufficiently large ( $\sim 10 \mu\text{m} - 100 \mu\text{m}$ ), well-ordered crystal of the protein target [7, 8]. There is no fundamental theoretical method for precisely predicting protein crystallisation; the specific conditions required to crystallise a protein must be determined, in part, by trial-and-error [2]. This process can be time consuming and expensive. Knowledge of the crystallisation conditions of one protein may or may not be helpful in determining the required conditions for another. The larger a crystal can be grown, the closer it comes to approximating the ideal mathematical abstraction of an infinite crystal, upon which the methods of interpreting its structure are based. The addition of more individual protein molecules makes a larger crystal a more powerful amplifier. Departures from crystalline order also reduce the diffraction quality of a crystal. Atomic and small-molecule crystals are characterised by high internal order and rigidity, however proteins are large molecules and protein crystals contain solvent channels that account for significant fractions of their volume ( $\sim 30\text{--}60\%$ ); this allows for deviations

in molecule-alignment, both translational and rotational. Such deviations can be exaggerated by the addition of thermal energy to the system, causing small-scale molecular motion. Slight variations in folding can result in molecules of slightly different size and shape, termed *conformational disorder*. Protein crystals also suffer from *mosaicity*, like atomic crystals, where small regions of the crystal that internally exhibit a high degree of order, are misaligned relative to each other, reducing the order of the crystal as a whole. So even once crystallisation conditions are discovered, there is no guarantee that a large crystal will result, or that it will be sufficiently well ordered to be useful for MX experiments.

X-ray crystallography has been used in structure determination for over 60 years and as proteins that readily crystallise have had their structures solved, the remaining targets are increasingly those that do not form large, well-ordered crystals. Technological advances at synchrotron x-ray sources have helped to offset this issue, and the high flux densities now available permit the use of smaller, less perfect crystals for structure determination. Crystals that would have been unsuitable for MX a decade or two ago can now be used in experiments, leveraging serial crystallography techniques [9, 10]. These techniques present new challenges, and/or a different perspective on challenges faced in more traditional crystallography, and understanding and refining the theory and practice of micro-crystallography are areas of active research.

### 1.3 Radiation Damage in X-ray Crystallography

One area where small protein crystals ( $\sim 1 \mu\text{m}$ ) may behave quite differently to their larger counterparts is in terms of the damage that they undergo as a result of the interaction between x-rays and matter [11, 12]. Damage is an inevitable consequence of interaction with x-rays. In an MX experiment, the requirement is to obtain sufficient diffraction information before the accumulation of damage reaches a threshold that precludes the further collection of useful data. Radiation damage in MX can be viewed as an accumulation of acquired disorder, which, in tandem with any inherent disorder the crystal originally possessed, serves to reduce the diffraction quality of the crystal. Although it can't be avoided outright, damage can be mitigated. For example, the use of cryogenic cooling became routine in MX in the 1990s; by cooling the sample to 100 K, radical species that spread damage are immobilised, and the protein molecules are fixed in place, hence they do not drift out of alignment as quickly, even when the weak bonds that originally fixed their crystalline positions are broken. Use of cryogenic cooling increased the radiation tolerance of protein crystals by a factor of 70 and also improved diffraction quality through the immobilisation of the protein molecules in the crystal lattice [13]. Within two decades, cryogenic crystallography had almost entirely replaced room temperature MX [14].

When cryogenic cooling was introduced, it rendered the typical crystal sample of the time effectively immortal under the illumination conditions available [15]. In 1990, Henderson [16] calculated that a protein crystal could survive an entire day's illumination by the most powerful synchrotron source available at that time. However, synchrotrons have evolved significantly in the last three decades. For example, the flux density available at BioMAX [17], the first MX beamline at the MAX-IV, 4<sup>th</sup> generation synchrotron, is six orders of magnitude higher than what was available when Henderson performed his calculations. This increase in flux density is so great that radiation damage is now a significant issue even for small molecule crystallography [18]. As noted above, the availability of high flux density has also made possible the use of smaller crystals with reduced internal order. The diffraction from these targets is typically poor compared to larger crystals, even before any damage has occurred. Therefore, the accrual of damage can very rapidly degrade them to the point that they are no longer usable.

It is possible, however, that smaller crystals may have an advantage over larger ones, when it comes to radiation damage. The primary driver of damage in MX is *photoelectric absorption*. The damage caused by this effect is spread over a small volume (within several microns) surrounding the interaction site. This distribution of damage has no noticeable impact, for example, when illuminating a 100  $\mu\text{m}$  crystal, but if the crystal is of similar dimension to the primary photoelectron range (ideally below 5  $\mu\text{m}$ ), there is a good chance that damage could be reduced relative to the level of x-ray illumination it receives, by depositing energy into the solvent surrounding the crystal [11]. Given the increasing need to use small crystals in MX, there is a strong motivation to understand and take advantage of this effect.

To maximise the potential for reducing damage by increasing the likelihood of inelastic interactions occurring outside of the crystal, higher energy x-rays should be used [12, 19], as the length scale of damage distribution is proportional to energy. MX experiments are typically performed within a narrow range of energies around 12.4 keV, with some variation required for phasing techniques that rely on the absorption edges of certain elements [20]. This has been the standard for many decades, and is firmly entrenched in the design and operation of MX beamlines. Some beamlines are incapable of operation at energies much higher than this, and those that are capable of deploying higher energy beams typically do not have well established protocols for doing so. Technology renders it possible, and the opportunity for damage reduction may render it desirable, but a period of trial and error is required to render it routine, and therefore, usable by the general experimenter.

## 1.4 Summary of This Thesis

The overarching goal of this thesis is to investigate radiation damage, especially with reference to micro-crystals and modern synchrotron apparatus involving intense, Gaussian micro-beams, through simulations, experiments, and meta-analysis.

Chapter 2 provides an overview of the theoretical background of macromolecular crystallography and the causes and effects of radiation damage. This material covers a range from fundamental theory to the current state of scientific knowledge as reported in the literature. It is intended to provide the general reader with sufficient context to comprehend the discussion presented in the following chapters.

Chapter 3 documents the scripting and execution of a discrete spatially- and temporally-resolved simulation of radiation damage that accounts for the finite spread of photoelectrons, and the energy they deposit, through the sample material. The core of this chapter is a paper originally published in the journal *Crystals* [21], it is presented here with additional material explaining the simulation in greater detail, and extending the original discussion of results to contextualise it within the thesis as a whole.

Chapter 4 presents two experiments designed to test the utility of using higher than normal x-ray energies in macromolecular crystallography with micro-crystals. Here we explore the technical challenges and test the application of the theory of photoelectron escape under real world conditions.

Chapter 5 is a comparison of two published models of radiation-induced Bragg spot fading. We have examined the predictions of these models with reference to data gathered in the experiments described in Chapter 4, and a meta-analysis of spot fading results from the literature.

Chapter 6 concludes this thesis with a summary of our activities and findings, and suggestions for further work.

### 1.4.1 Supporting Material

The MATLAB and Python scripts produced during the course of the research presented in chapters 3, 4, and 5 have been uploaded to an online repository, hosted by the Open Science Framework. They can be accessed with the following link: <https://osf.io/jv9he/>.

## 1.4.2 Software Referenced in This Thesis

In addition to the code produced specifically for this research, several software applications were used extensively, some of which are in continual, active development. Typically, the versions used were the most recent available at the time the research was conducted; the notable exception being RADDPOSE v2, which was used in addition to the then most recent version of RADDPOSE-3D, as discussed in Section 4.9 of the paper ‘The Influence of Photoelectron Escape in Radiation Damage Simulations of Protein Micro-Crystallography’ (see Section 3.4). Table 1.1 outlines which software versions were used throughout the research presented in this thesis.

Table 1.1: Versions of software used and referenced in this thesis.

Application	Version
RADDPOSE-3D	v1.2.427 used in Chapter 3
	v4.0 used in our analysis in Chapters 4 and 5
	v1.2.427 also relevant to discussion in Chapter 5
CASINO	v2.42 used in Chapter 3
Albula	v3.2.0 used in Chapter 4
CrystFEL	v0.8.0 used in Chapters 4 and 5

## Chapter 2

# Background & Literature Review

In this chapter, we present a theoretical and empirical basis for the concepts explored in subsequent chapters. Section 2.1 covers the interactions of x-rays and matter; section 2.2 deals with the theoretical basis of macromolecular crystallography (MX) and provides a brief review of synchrotron radiation and modern synchrotron light sources; and section 2.3 focuses on the causes, nature, and effects of radiation damage to crystals during MX experiments. Particular attention is given to loss of intensity in Bragg reflections and the role of photoelectrons in damage.

### 2.1 X-rays & Their Interactions with Matter

X-rays make up the band of electromagnetic radiation between ultraviolet and gamma-rays. As electromagnetic radiation is a continuous spectrum, placing hard limits on the energies (or equivalently, wavelengths or frequencies) which define a particular band is problematic, but x-rays may typically be taken as having energies on the order of 100 eV to 100 keV, with the x-ray band further split into soft x-rays (up to several keV) and hard x-rays ( $> 5$  keV or 10 keV). In MX, the desire to image proteins at as close to atomic-resolution as possible necessitates a probe with a wavelength of approximately 1 Å or shorter, which corresponds to roughly 12.4 keV (or higher).

#### 2.1.1 Transmission, Scattering, and Absorption

Hard x-rays interact only weakly with matter, particularly when dealing with materials composed primarily of lighter elements, as is the case with biological samples. Roughly 73% of 12.4 keV x-ray photons will pass through a 1 mm thick soft-tissue sample (ICRU-44) [22], and at the length-scales common in MX, transmission is even higher; over 99% of 12.4 keV x-ray photons are transmitted through 10 μm of the same sample. X-ray photons that do not pass through a material can interact in a variety of ways, depending on their energy; for the energies

and samples used in MX, the relevant interactions are elastic scattering, also known as *Rayleigh scattering*; inelastic scattering, or *Compton scattering*; and photoionisation, also known as photoelectric absorption. These interactions are summarised in fig. 2.1, and detailed in the following sections.

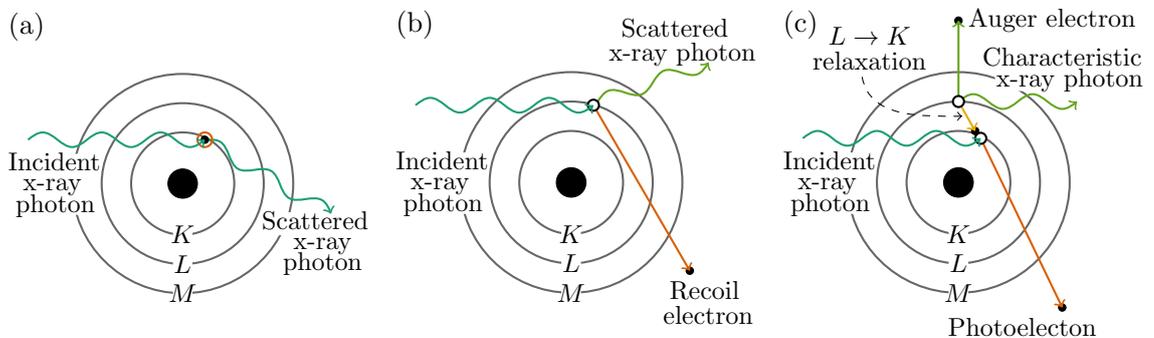


Figure 2.1: **X-ray-matter interactions relevant in the 10 keV to 30 keV hard x-ray energy regime.** (a) An incident x-ray photon is *elastically scattered*, i.e. the scattered photon has energy equal to the incident photon. (b) An incident x-ray photon is *inelastically scattered*, producing a lower energy photon and a low-energy *recoil electron*. (c) An incident x-ray photon causes *photoionisation*, resulting in the ejection (from a core-shell) of a high-energy *photoelectron* (PE). Following photoionisation, an electron relaxation will result in the emission of a (lower energy) characteristic x-ray photon or an Auger electron.

### Elastic Scattering

In the classical description of elastic scattering, the electric field of the incident x-ray beam causes atomic electrons to vibrate parallel to the electric-field of the beam, giving rise to *electromagnetic dipole radiation*. The vibrating electron then acts as a source of radiation at the same energy as the incident radiation. As no energy is lost to the material, this process does not contribute to radiation damage. It is the interference of elastically scattered radiation at the detector plane that is observed in MX and used to determine the position of the scattering centres (the electron distribution of the material). For a  $10\ \mu\text{m}$  soft-tissue sample illuminated by 12.4 keV x-rays, 0.02% of the incident x-ray photons will result in elastic scattering, which accounts for roughly 6% of the interaction between the beam and the sample; the remaining 94% of the interactions result in radiation damage.

### Inelastic Scattering

Also known as *Compton scattering* after its discoverer Arthur Holly Compton [23], inelastic scattering involves, as the name suggests, a transfer of energy from the photon to an atomic electron.

The energy,  $E'$  of the scattered photon is given by

$$E' = \frac{E}{1 + (E/m_e c^2)(1 - \cos(\theta))}, \quad (2.1)$$

where  $E$  is the incident photon energy,  $m_e$  is the electron mass,  $c$  is the speed of light in a vacuum, and  $\theta$  is the angle at which the photon is scattered. The difference in energy,  $\Delta E = E - E'$ , is deposited in the material and contributes to radiation damage. For an incident photon with energy  $E = 12.4 \text{ keV}$ , the maximum energy deposited in the material (if the scattered photon's direction is  $180^\circ$  to the incident photon) is slightly less than  $600 \text{ eV}$ , and the average energy deposited by a Compton scattering event is roughly half this. For a  $30 \text{ keV}$  incident photon, the maximum deposited energy is  $3.2 \text{ keV}$ , and the average energy is  $1.6 \text{ keV}$ .

### Photoelectric Absorption

Photoelectric absorption is the most likely outcome of an x-ray–matter interaction at the energies typically used for MX experiments. Here, an incident x-ray photon displaces a core electron (typically a  $K$ -shell electron as they have the highest interaction cross-section), ionising the atom involved; the resulting *photoelectron* (PE) is emitted from the atom with energy equal to the difference between the incident photon energy and the binding energy of the shell from which it was ejected ( $E_{PE} = E_\Phi - E_K$ ). The binding energies of  $K$ -shell electrons for light elements like hydrogen, carbon, nitrogen, and oxygen—the primary elements of proteins—are all under  $1 \text{ keV}$ , so PEs produced in MX experiments typically have energies very close to the incident beam energy.

Photoionisation leaves the atom involved in an energised state; it will subsequently go through an energy relaxation when a higher-shell electron collapses to fill the vacancy left in the core shell, resulting in emission of either a fluorescent x-ray photon or Auger electron. The mode, and product, of this relaxation, though important in some cases, is not relevant here; instead we are concerned with the behaviour of the PE subsequent to its emission from the ionised atom, a topic which is considered in section 2.3.6.

#### 2.1.2 Interaction Cross Sections

The likelihood of a particular interaction is given by its interaction cross-section and can vary significantly with the energy of the incident x-ray beam, and with the sample. Figure 2.2 shows the cross sections of carbon, sulphur, and gold, for x-rays between  $1 \text{ keV}$  and  $30 \text{ keV}$ . Carbon is a primary component of proteins, and sulphur is a heavier atom reasonably common in proteins via inclusion of the amino acids methionine and cysteine. Gold is included as an example of

a heavier element, with which x-rays interact more strongly (relative to lighter elements). The orders-of-magnitude difference in total cross section between carbon and the heavier elements typical of small-molecule and atomic crystals highlights the relative transparency of proteins (and protein crystals) to hard x-rays. Also of note is that, in general, the cross section for photoionisation is typically two to three orders of magnitude greater than the cross section for elastic scattering, though in the case of carbon, this difference shrinks to around a factor of 2 at 30 keV. In the case of carbon, the cross section for Compton scattering exceeds that of photoionisation above roughly 22 keV, causing it to contribute more significantly to radiation damage. The energy deposited in the sample *per interaction* remains roughly 10 times greater for photoionisation so, even at this energy, its contribution to damage remains slightly higher than that of Compton scattering.

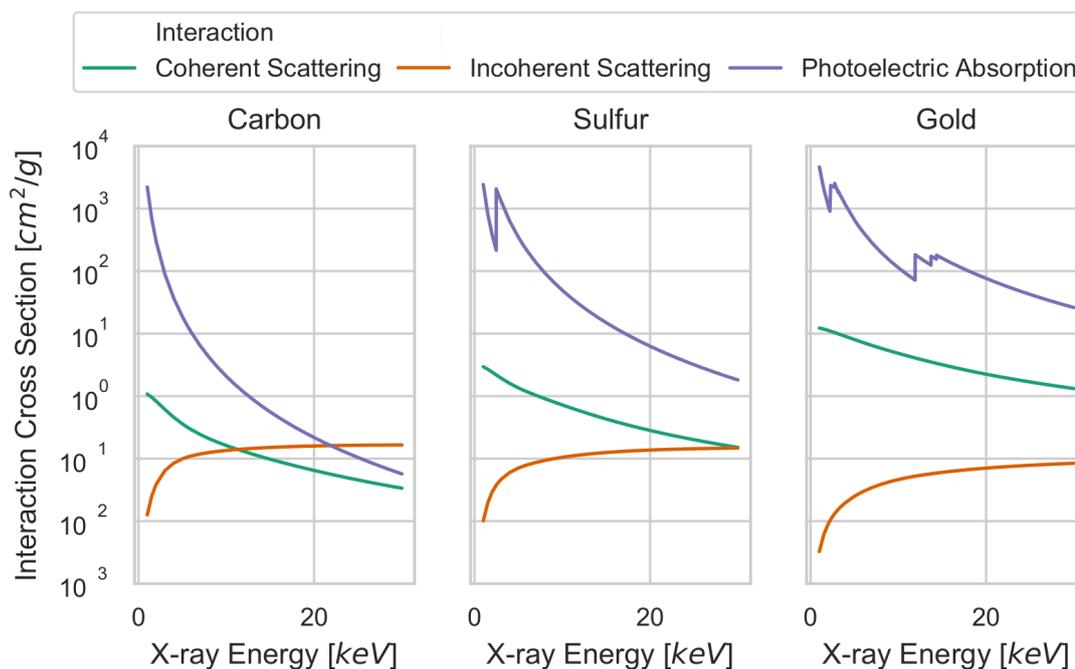


Figure 2.2: **Selected elemental x-ray interaction cross sections.** Cross section for coherent scattering (teal), incoherent scattering (orange), and photoelectric absorption (lavender), between 1 keV and 30 keV are shown for carbon ( $Z = 6$ ), sulfur ( $Z = 16$ ), and gold ( $Z = 79$ ). The  $y$ -axis has a logarithmic scale, i.e. elastic scattering is an order of magnitude greater for gold compared to carbon.

## 2.2 Macromolecular Crystallography

The term crystallography was coined in the 18th century by Maurice Capeller to describe the study of crystals [24]. This discipline initially relied on observation of the *external* regularity of crystals, as an indication of their *internal* order; the discovery, by Roentgen, of x-rays (and a means of producing them) in the late 19th century changed this, allowing the interior of crystalline matter to be probed directly. Since then, crystallography has expanded to be a broad and fundamental scientific discipline, providing insight into the properties of solid materials, both inorganic and organic. In this section, we lay out the fundamental theoretical principles of crystallography and its application to determining protein structure, and provide an overview of the x-ray sources typically used in crystallography experiments.

### 2.2.1 Crystals

Crystals are a category of solid materials that exhibit a regular ordering of their constituent atoms or molecules. Crystalline order can be seen with the naked eye by examining, for example, large salt crystals or crystals of quartz; the faces of the crystals, and the angles between them, are indicative of the crystals' internal order (see fig. 2.3c). Crystals can be described as the combination of a *lattice*, an abstract, regular array of points in space; and a *motif*, the repeating unit of the crystal. By placing a copy of the motif at each lattice point, the crystal is constructed. A 2-dimensional example of this process is shown in figs. 2.3a and 2.3b. In mathematical terms, the crystal is the *convolution* of the motif and the lattice. The motif of a crystal may be a single atom, as is the case for many metals; a small molecule such as NaCl; or something much larger and more complicated, such as a protein.

#### The Bravais Lattices

It was shown by Bravais [25] that there are 14 unique lattices that can describe all crystalline arrangements in three dimensions. A Bravais lattice, which is infinite in extent, is defined by the equation

$$\mathbf{R} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}, \quad (2.2)$$

where  $\mathbf{a}$ ,  $\mathbf{b}$ , and  $\mathbf{c}$  are primitive vectors describing three unique (not necessarily orthogonal) directions, and  $n_i$  are integers. The angles  $\alpha$ ,  $\beta$ , and  $\gamma$  are often used in a description of a Bravais lattice, where  $\alpha$  is the angle between  $\mathbf{b}$  and  $\mathbf{c}$ ,  $\beta$  between  $\mathbf{c}$  and  $\mathbf{a}$ , and  $\gamma$  between  $\mathbf{a}$  and  $\mathbf{b}$ . Bravais lattices are constructed by combining the lattice systems, which define the relative lengths of the primitive vectors and the angles between them; and centring types, which identify the locations of the lattice points with respect to the unit cell. The 14 Bravais lattices can be constructed in two

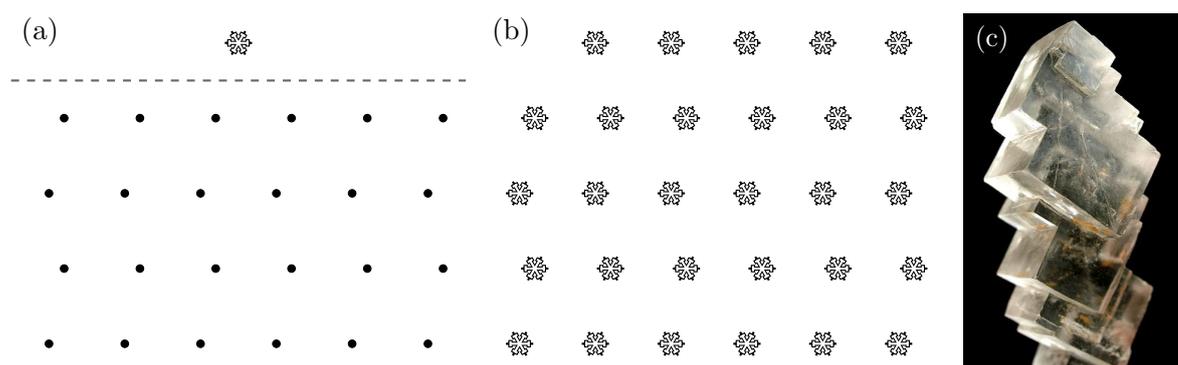


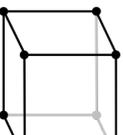
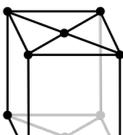
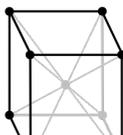
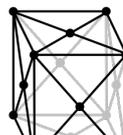
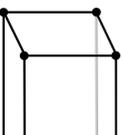
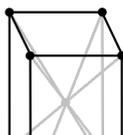
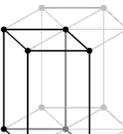
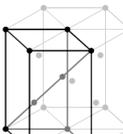
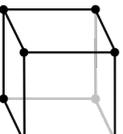
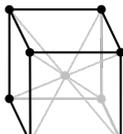
Figure 2.3: **Crystal structure.** A conceptual model of a 2 dimensional crystal is shown in (a) & (b): the motif (above) and lattice (below) in (a) are combined to form the crystal in (b). (c) Example of a halite (salt) crystal, illustrating that the internal order of the crystal can be seen in the outer faces. Image by Rob Lavinsky, used, without modification, under a Creative Commons license (CC BY-SA-3.0).

ways: using 7 lattice systems and 4 centring types, or 6 lattice systems and 5 centring types; the choice affects only a single lattice (the *trigonal* lattice which can be considered a rhombohedrally-centred hexagonal lattice or a primitive rhombohedral lattice). We will use the combination of 6 lattice systems —trigonal, monoclinic, orthorhombic, tetragonal, hexagonal, and cubic— and 5 centring types —primitive, base-centred, body-centred, face-centred, and rhombohedrally-centred— as this description aligns with the *Pearson symbols* commonly used to represent the Bravais lattices. Not every combination of lattice system and centring type is required, as several of them are equivalent to each other. The construction of the 14 Bravais lattices is shown in table 2.1.

### The Unit Cell

The *unit cell* is, like the lattice itself, a geometrical construct; it is a subdivision of a lattice that can be translated in such a way as to fill three-dimensional space, without gaps. A simplified example, in two dimensions, is shown in fig. 2.4. Like the motif, the unit cell is repeated in reference to the lattice, however the unit cell and the motif are distinct entities; it is possible to create different unit cells (in fig. 2.4, an orange cell labelled 'A' and a teal cell labelled 'B') for a given combination of lattice and motif, and a unit cell may contain multiple copies of the motif (this is the case, for example, with tetragonal lysozyme, which has 8 molecules (the motif) to each unit cell). A *primitive unit cell* is one that contains a single lattice point (cell 'A' in fig. 2.4 is primitive). For atomic or small-molecule crystals, a unit cell may contain only a handful of atoms; in *macromolecular crystallography*, the unit cell will contain one or more protein molecules, each containing thousands, or even hundreds of thousands of atoms.

Table 2.1: **The fourteen Bravais lattices used to describe three dimensional crystals.** The Bravais lattices are shown based on their construction via six *lattice systems* and five *centring types*. Each system is defined by a set of restrictions to the unit cell lengths,  $a$ ,  $b$ , and  $c$ , the angles between them,  $\alpha$ ,  $\beta$ , and  $\gamma$ , or both. The centring types dictate the presence of additional lattice points beyond the points at the vertices of the primitive lattice. Images adapted from originals by Bor75, DrBob, Daniel Mayer, Rocha, and Stannered, under a Creative Commons license (CC BY-SA 3.0).

Lattice System	Centring Type				
	Primitive	Base	Body	Face	Rhombohedral
Triclinic -	 aP				
Monoclinic $\alpha = \beta = 90^\circ$	 mP	 mS			
Orthorhombic $\alpha = \beta = \gamma = 90^\circ$	 oP	 oS	 oI	 oF	
Tetragonal $a = b$ $\alpha = \beta = \gamma = 90^\circ$	 tP		 tI		
Hexagonal $a = b$ $\alpha = \beta = 90^\circ,$ $\gamma = 120^\circ$	 hP				 hR
Cubic $a = b = c$ $\alpha = \beta = \gamma = 90^\circ$	 cP		 cI	 cF	

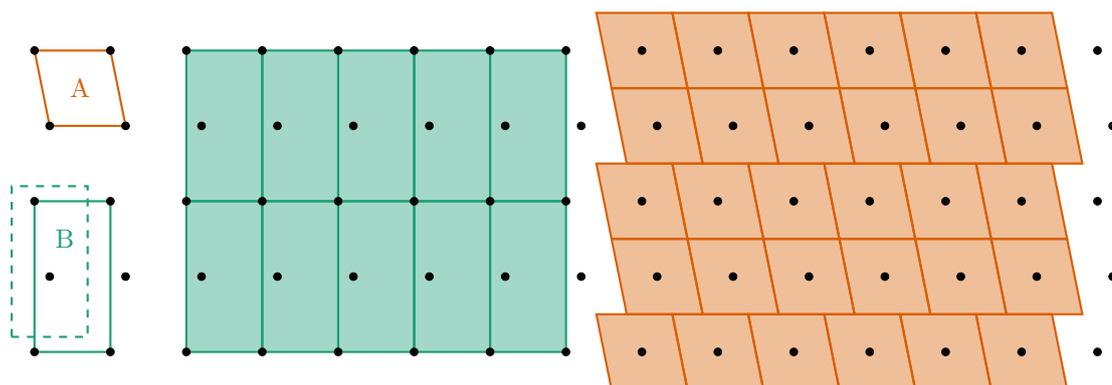


Figure 2.4: **A 2-dimensional representation of the unit cell and crystal lattice.** The array of point represent the lattice points of a crystal. Shown on the left in orange and teal are two possible unit cells, labelled 'A' and 'B'. Both A and B are valid unit cells: they can be used to tile the plane lattice as illustrated in the centre and on the right. Unit cell A is *primitive* as it contains only one lattice point, a fact that is emphasised by translating it so that it is centred on the lattice point (on the right). Unit cell B contains two lattice points, in this case highlighted by the the dashed-line copy of the cell that has been translated up and to the left.

### Symmetry & Space Groups

Symmetry is a property of objects whereby the application of some kind of transformation leaves it apparently unchanged. The repetition of a motif across a lattice shows that crystals have *discrete translational symmetry* in that translating the crystal by an integer number of unit cells returns an identical crystal. Notions of *mirror symmetry* and *rotational symmetry* are familiar from day-to-day life. These are examples of *point symmetries*, as the point about which the transformation occurs does not move. There are five types of point symmetry in three dimensions, if one includes the identity operation, the other four are rotation, reflection (in plane), inversion (mapping  $(x, y, z)$  to  $(-x, -y, -z)$ ), and improper rotation (a rotation followed by a reflection in the plane perpendicular to the rotational axis). In addition, there are two symmetry operations that combine point symmetry operations and translation: the *screw*, which involves a rotation and a translation of a fraction of the unit cell; and the *glide* which combines a reflection and fractional translation.

Different sets of the point symmetries define the 32 crystallographic point groups which, in combination with the 14 Bravais lattices, form the 73 *symmorphic space groups*. Considering also the screw and glide symmetries leads to an additional 157 *non-symmorphic space groups*, for 230 space groups altogether. A crystal can be defined by its space group and its *asymmetric unit*, which is the smallest fraction of the unit cell that can be rotated, translated, etc. according to the symmetries of the space group to recover the complete unit cell. For example, a single lysozyme

molecule, specifying the relative position of all its constituent atoms, forms the asymmetric unit of a tetragonal lysozyme crystal, with the space group  $P4_32_12$ . By applying the space group symmetries, the complete unit cell, containing 8 molecules, is created.

## 2.2.2 X-ray Diffraction by Crystals

Diffraction describes the bending of a wave around an obstacle—or, equivalently, through an aperture—to extend into the obstacle's 'shadow'. It can be readily observed in the everyday world by, for example, watching the interaction of waves with the support posts of a pier or jetty. The diffraction of visible (laser) light through a diffraction grating is a common physics demonstration. Diffraction requires that the wavelength and the obstacle involved be of similar size; a condition which makes hard x-rays ideal for examining the structure of materials, as their wavelengths and the inter-atomic spacing are both on the order of angstroms.

X-ray diffraction occurs due to the interaction between x-rays and charged particles: the atomic electrons of the material. Electromagnetic radiation accelerates charged particles, and that acceleration produces electromagnetic radiation. In the case of elastic scattering, the source of x-ray diffraction, an incident x-ray photon causes oscillation of an atomic electron, which acts as a source for a new x-ray photon of the same energy. The x-ray photons thus scattered from the sample are subsequently incident upon a detector, in which plane they interfere, based on their relative phase, causing a diffraction pattern. When the sample is a crystal, the repetition of the motif (e.g. a protein molecule) acts as an amplifier, producing a diffraction signal that is the coherent superposition of the diffraction patterns of many thousands of individual copies of the motif. This effect is illustrated in fig. 2.5. The model molecule is a pentagonal ring with a short tail; it is shown along with its diffraction pattern in fig. 2.5a. The colours of the diffraction pattern represent the phase information. In fig. 2.5b, a one-dimensional 'crystal' has been created by copying the molecule. The vertical fringes of the diffraction pattern are the reciprocal of the spacing between molecules. In figs. 2.5c and 2.5d the crystal has been extended to two-dimensions with 8 and 15 molecules, respectively. The discrete nature of the diffraction pattern is much stronger, highlighting the reciprocal lattice. There are still significant fringes around the Bragg peaks, as the crystal is spatially limited in extent (see section 2.2.4). The phases (colour) of the Bragg spots are identical to the diffraction of the molecule at the same point in reciprocal space (fig. 2.5a); this highlights the necessity of recovering phase information to determine molecular structure (see section 2.2.4). In fig. 2.5e, the crystal has been extended beyond the repetition shown in the cropped image, to be essentially infinite in spatial extent. This is reflected in the discrete nature of the diffraction pattern, which now lacks the fringes shown in figs. 2.5c and 2.5d. The inter-molecular spacing has also increased, resulting in reduced spacing between the Bragg spots of

the diffraction pattern.

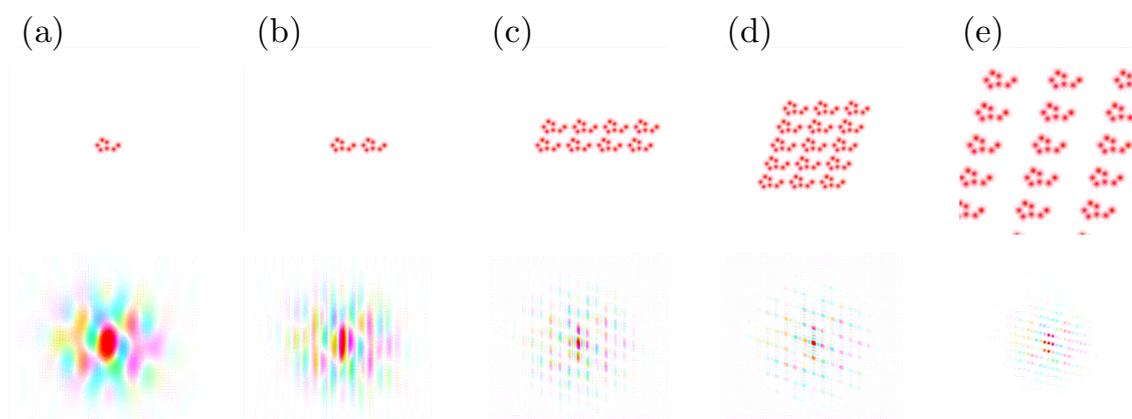


Figure 2.5: **Relationship between the continuous diffraction pattern of a molecule and the discrete diffraction pattern of a crystal.** An example molecule —a pentagonal ring of atoms with a two-atom ‘tail’— is used to demonstrate the relationship between the diffraction of a discrete object and the diffraction of a crystal. The arrangement of molecules (in real space) is shown above, the diffraction (reciprocal space) below. Colours in the lower images indicate phases. (a) A single molecule. (b) Two molecules forming a minimal one-dimensional ‘crystal’. (c) and (d) Two-dimensional crystals of 8 and 15 molecules, respectively. (e) An ‘infinite’ crystal. Images adapted from originals by Kevin Cowtan, used with permission.

### Crystal Planes & Bragg’s Law

A frequently used heuristic in considering the formation of diffraction patterns is to imagine the crystal built up of intersecting sets of parallel planes. These planes run through the lattice sites of the crystal (be they atoms or more complex molecules) and the external faces of the crystal are visible examples of them. The planes are labelled by their *Miller indices*, which are derived from the intersections of the planes with the coordinate system of the lattice, as illustrated in fig. 2.6. For example, the (1 0 0) plane is parallel to both the *b* and *c*-axes: the zeros indicate that it does not intersect them.

Using the ray model of electromagnetic radiation, Bragg’s law describes the conditions under which x-rays that ‘reflect’ off these parallel crystal planes will create the diffraction patterns observed at the detector. As shown in fig. 2.7, when two parallel rays reflect off adjacent crystal planes the difference in their path lengths is related to the plane-spacing,  $d$ , and the angle that the x-rays make to the planes,  $\theta$ . If this path-length difference is an integer number of wavelengths, they constructively interfere at the detector, causing a bright spot. If not, they will destructively interfere. Analysis of the so called *Bragg spots* of a diffraction pattern can thus recover information about the spacing, and orientation, of the crystal planes and, by extension, the crystal

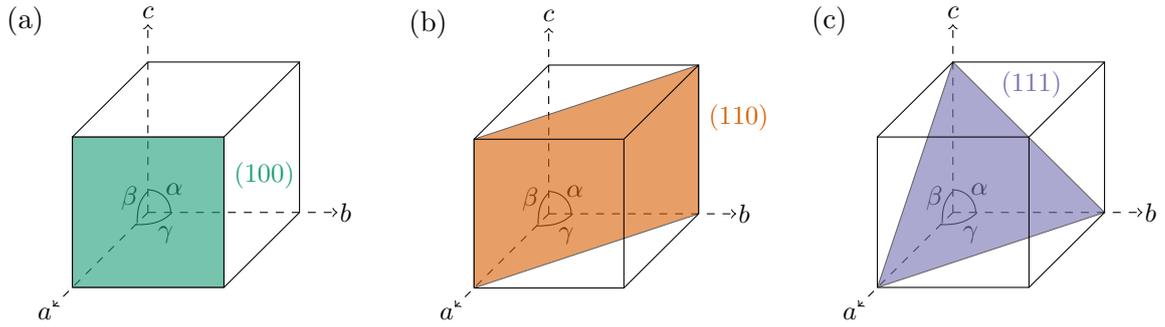


Figure 2.6: A representation of crystal planes and their Miller indices as they relate to the crystal's unit cell. A primitive unit cell is shown in each of (a), (b), and (c), defined by the basis vectors  $a$ ,  $b$ ,  $c$ , and the angles  $\alpha$ ,  $\beta$ ,  $\gamma$ . By definition the edge-lengths of the unit cell are 1. The shaded regions indicate different crystal planes, defined by their intersections with the lattice vectors. The Miller indices are given by  $(1/a, 1/b, 1/c)$ , i.e. the (100) plane (teal) intersects  $a$  at 1, and does not intersect either  $b$  or  $c$  ( $1/\infty = 0$ ); the (110) plane (orange) intersects  $a$  and  $b$  at 1, and does not intersect  $c$ ; the (111) plane (lavender) intersects all three axes at 1.

lattice. Bragg's law is expressed as

$$n\lambda = 2d \sin(\theta), \quad (2.3)$$

where  $n$  is a positive integer, and  $\lambda$  is the wavelength of the x-rays.

### 2.2.3 The Reciprocal Lattice

A reciprocal lattice can be obtained via the Fourier transform of another lattice (e.g. a Bravais lattice). In crystallography, the term is used to denote the *reciprocal space lattice*, where the *real space lattice* is the Bravais lattice of the crystal in question. A useful characteristic of the reciprocal space lattice is that it maps an array of crystal planes onto an array of points. Consider a general set of crystal planes, in real space, identified by Miller indices  $(hkl)$ . The spacing between these planes is  $d$ . Let a vector be positioned perpendicular to a given plane, with length  $1/d$ . This vector defines a point, in reciprocal space, that is labelled  $hkl$  (with no parentheses). Each set of planes in the crystal corresponds to such a point in reciprocal space. Together, the set of points constitute a representation of the crystal itself. These points have coordinates on the reciprocal lattice that correspond to the coordinates of the plane on the real space lattice. The reciprocal lattice has a basis given by the lattice vectors  $\mathbf{a}^*$ ,  $\mathbf{b}^*$ , and  $\mathbf{c}^*$ , which are related to the basis vectors of the real space lattice,  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$ , by two sets of equations:

$$\mathbf{a}^* \cdot \mathbf{b} = \mathbf{a}^* \cdot \mathbf{c} = \mathbf{b}^* \cdot \mathbf{c} = \mathbf{b}^* \cdot \mathbf{a} = \mathbf{c}^* \cdot \mathbf{a} = \mathbf{c}^* \cdot \mathbf{b} = 0, \quad (2.4)$$

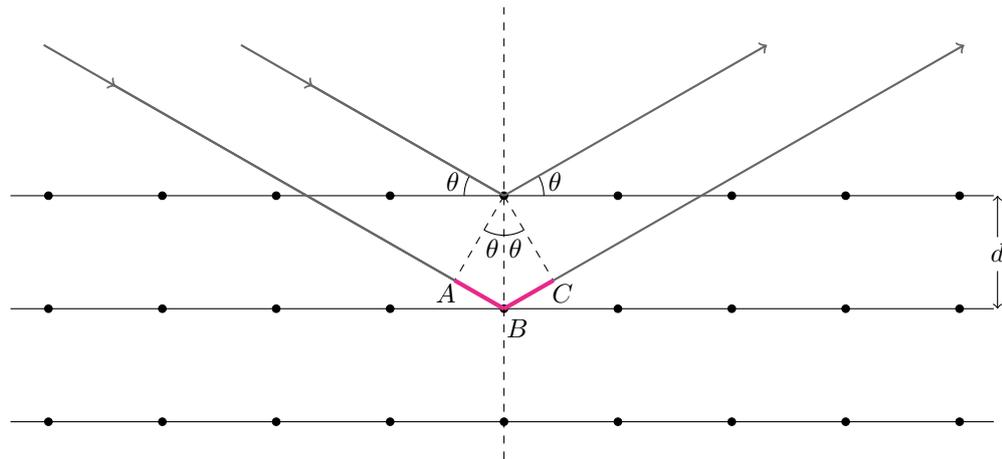


Figure 2.7: A schematic representation of Bragg's law. Parallel x-rays 'reflect' from adjacent crystal planes. When their path length difference ( $ABC$ , highlighted pink) is an integer number of wavelengths they arrive at the detector in phase, constructively interfering, resulting in a Bragg spot.

and

$$\mathbf{a}^* \cdot \mathbf{a} = \mathbf{b}^* \cdot \mathbf{b} = \mathbf{c}^* \cdot \mathbf{c} = 1. \quad (2.5)$$

Equation (2.4) stipulates that  $\mathbf{a}^*$  is orthogonal to both  $\mathbf{b}$  and  $\mathbf{c}$  (and so on for vectors  $\mathbf{b}^*$  and  $\mathbf{c}^*$ ). Equation (2.5) defines the length of the reciprocal space basis vectors as the inverse of their real space counterparts ( $|\mathbf{a}^*| = a^* = 1/a$ , etc). A 2-dimensional representation of this inverse relationship can be seen in 2.5: if the spacing between molecules (each occupying a lattice point in real space) is increased, the Bragg spots in the diffraction pattern (each occupying a lattice point in reciprocal space) is decreased, and vice-versa.

### The Ewald Sphere

The link between real space and reciprocal space, and the utility of the latter in considering diffraction patterns, can be seen via the construction of the *Ewald sphere*, named for Paul Ewald, who conceived of it in the early 20<sup>th</sup> century. Consider the two-dimensional representation in fig. 2.8a. The point  $O$  represents the position of a crystal in real space, oriented such that the line  $AC$  makes the angle  $\theta$  with a set of planes ( $hkl$ ), in the crystal, with plane spacing  $d$ . The point  $C$  is the origin of the reciprocal space lattice. The crystal is illuminated by x-rays of wavelength  $\lambda$ , travelling in the direction of the line  $AC$ . The line  $CP$  is perpendicular to the planes of the crystal and has length  $1/d$ , in other words the point  $P$  is the lattice point  $hkl$  in reciprocal space. A sphere (here represented by a circle) of radius  $1/\lambda$  is drawn through  $C$ ; if it passes through

point  $P$ , as illustrated here, then consideration of the right triangle  $APC$  yields

$$\sin(\theta) = \frac{CP}{AC} = \frac{1/d}{2\lambda} = \frac{\lambda}{2d}, \quad (2.6)$$

which is Bragg's law (eq. (2.3)) with  $n = 1$ . So if a reciprocal lattice point lies on the Ewald sphere, then Bragg's law dictates that diffraction will be observed in the direction of  $OP$ .

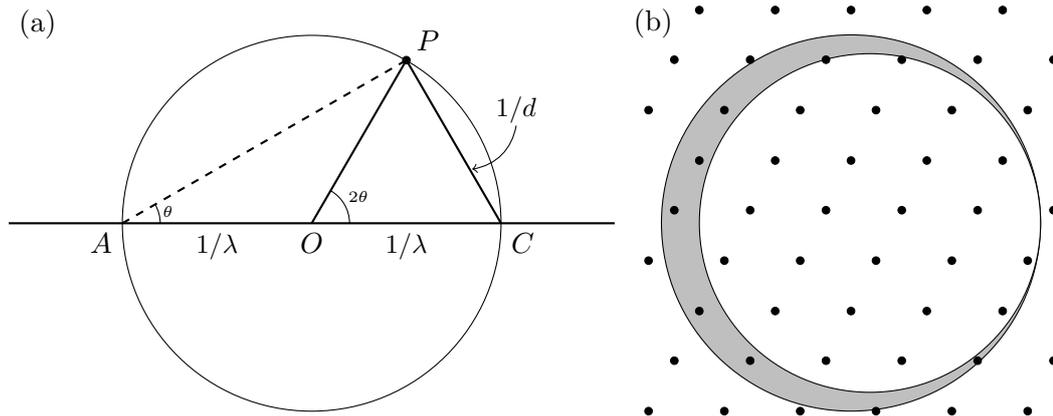


Figure 2.8: **Ewald sphere: construction and effect of bandwidth.** (a) The geometric construction of the Ewald sphere in 2D projection.  $O$  represents the position of a crystal, in *real space*,  $C$  is the origin of *reciprocal space*. X-rays travel in the direction  $AC$ , making angle  $\theta$  to crystal planes, with separation  $d$ .  $CP$ , perpendicular to the crystal planes, with length  $1/d$ , is the reciprocal space vector corresponding to the crystal planes. Intersection of the Ewald sphere (with radius  $1/\lambda$ ) with the point  $P$ , indicates that it fulfils the Bragg condition; diffraction will be observed in the direction of  $OP$ . (b) When the x-rays are not perfectly monochromatic, multiple wavelengths,  $\lambda$ , will result in multiple radii of the Ewald sphere. Any reciprocal lattice points falling within the range of radii (shaded gray) will fulfil the Bragg condition for one of the wavelengths present in the x-ray beam.

In practice, an x-ray beam will always contain at least a small spread of energies (or, equivalently, wavelengths), even if it is highly monochromated. Figure 2.8b illustrates the effect of the bandwidth on the Ewald sphere: two circles are coincident at the origin of reciprocal space; the outer circle corresponds to the lowest frequency present in the beam, the inner circle to the highest frequency. Any reciprocal lattice points that fall between the two circles will fulfil the Bragg condition for one of the wavelengths present in the beam and therefore be present in the diffraction pattern.

## 2.2.4 Crystallography & Structure Determination

The amplitude,  $A(\mathbf{q})$ , of the x-radiation scattered by a crystal is described by the Fourier transform of the crystal's electron density,  $\rho_X(\mathbf{r})$ :

$$A(\mathbf{q}) = \mathcal{F}(\rho_X(\mathbf{r})) = \int \rho_X(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}, \quad (2.7)$$

where  $\mathbf{r}$  is a real-space position vector and  $\mathbf{q}$  is a reciprocal-space position vector, sometimes called the *momentum transfer vector*. As noted in section 2.2.1, the crystal can be described as the convolution of the lattice,  $L(\mathbf{r})$ , and the motif. If we take the motif to be the electron density of the unit cell,  $\rho_{UC}(\mathbf{r})$ , the electron density of the crystal is

$$\rho_X(\mathbf{r}) = L(\mathbf{r}) * \rho_{UC}(\mathbf{r}). \quad (2.8)$$

By the convolution theorem, the Fourier transform of a convolution of two functions is the product of the Fourier transforms of the individual functions, and we can rewrite eq. (2.7) as

$$A(\mathbf{q}) = \mathcal{F}(L(\mathbf{r}) * \rho_{UC}(\mathbf{r})) = \mathcal{F}(L(\mathbf{r}))\mathcal{F}(\rho_{UC}(\mathbf{r})). \quad (2.9)$$

The lattice,  $L(\mathbf{r})$ , is described by a series of dirac delta functions, being zero everywhere except the lattice sites,  $\mathbf{R}_n$ :

$$L(\mathbf{r}) = \sum_n \delta(\mathbf{r} - \mathbf{R}_n). \quad (2.10)$$

The Fourier transform of a series of dirac delta functions is, itself, a series of dirac delta functions (this is merely the mathematical statement that the reciprocal space lattice is the Fourier transform of the real space lattice):

$$\mathcal{F}(L(\mathbf{r})) = L(\mathbf{q}) = \frac{1}{V} \sum_n \delta(\mathbf{q} - \mathbf{Q}_n), \quad (2.11)$$

Where  $V$  is the volume of the (real space) unit cell. The Fourier transform of the electron density of the unit cell is called the unit cell *structure factor*,  $F(\mathbf{q})$ , giving

$$A(\mathbf{q}) = \frac{F(\mathbf{q})}{V} \sum_n \delta(\mathbf{q} - \mathbf{Q}_n). \quad (2.12)$$

The information recorded by the detector is the intensity,  $I(\mathbf{q})$ :

$$I(\mathbf{q}) = |A(\mathbf{q})|^2 = \frac{|F(\mathbf{q})|^2}{V^2} \sum_n |\delta(\mathbf{q} - \mathbf{Q}_n)|^2. \quad (2.13)$$

The dirac delta series indicates that the position of the Bragg peaks is determined by the crystal lattice, and reveals the size and symmetry of the unit cell. The magnitude term outside the series gives the intensities of the Bragg peaks; these intensities relate to the atomic positions within the unit cell. It is this latter information that protein crystallographers are interested in. The unit cell electron density is obtained via the inverse Fourier transform of the square root of eq. (2.13)

### The Phase Problem

The structure factor,  $F(\mathbf{q})$ , is a complex quantity (as is  $A(\mathbf{q})$ ), having both an amplitude,  $|F(\mathbf{q})|$ , and a phase,  $e^{i\phi(\mathbf{q})}$ :

$$F(\mathbf{q}) = |F(\mathbf{q})|e^{i\phi(\mathbf{q})}. \quad (2.14)$$

In measuring the intensity,  $I(\mathbf{q})$ , the phase information is lost; only the amplitude is retained. In order to recover the electron density of the unit cell and thereby solve the protein structure, the phase information must be somehow recovered. This is known as *the phase problem*. Phase recovery<sup>1</sup> can be accomplished in various ways and the process typically involves multiple stages of refinement via a back-and-forth series of Fourier transforms whereby an initial estimate of the electron density with phases is used to produce a simulated diffraction pattern that can be compared to the observed pattern and the difference used to re-estimate the phases, and so forth. This algorithmic refinement is handled by crystallography software, typically incorporated into a detailed, automated software pipeline. The original phase estimate is often obtained via anomalous phasing techniques or taken from a previously-solved homologous structure.

### Finite, Imperfect Crystals

The description of crystal diffraction and derivations up to this point have assumed a perfect crystal of infinite extent. These properties are not possessed by real crystals, though in practice as long as the crystal is very large relative to the unit cell and the degree of crystalline order is high, it can be suitably approximated by an infinite perfect crystal.

The description of an infinite crystal as the convolution of its Bravais lattice,  $L(\mathbf{r})$ , and the electron density of the unit cell,  $\rho(\mathbf{r})$ , used above can be adapted for a finite crystal by the inclusion

<sup>1</sup>In the context of crystallography, the term 'phase recovery' may refer to the process of determining an initial phase estimate. Here, we use it to indicate the recovery of the *final* phase information.

of a shape function,  $s(\mathbf{r})$ , which has the property of being equal to 1 within the crystal bounds, and 0 outside of it:

$$C(\mathbf{r}) = L(\mathbf{r}) * (\rho(\mathbf{r}) \cdot s(\mathbf{r})). \quad (2.15)$$

In reciprocal space the shape function of the crystal manifests in the presence of fringes around the central bright spots of the Bragg peaks. This can be seen in fig. 2.5c - 2.5d: the 'crystal' is finite, composed of only a few molecules, and the fringes around the bright lattice points are very clear, while in fig. 2.5e, they have essentially disappeared as the 'crystal' becomes infinite in extent. Conventional approaches to structure retrieval essentially ignore these fringes and use only the central peaks themselves, as these fringes cannot typically be resolved using a standard MX setup. However, under certain conditions, the real-space information represented by the fringes can be recovered using a technique known as *Bragg coherent diffractive imaging* (BCDI). The application of this technique to protein crystals is relatively new and more challenging compared to BCDI of comparatively radiation-hard materials samples. BCDI has been used to create a three-dimensional reconstruction of a lysozyme crystal [26] and to study the effects of radiation damage in micro-crystals [27, 28]. The technique highlights the possibility of using micro-crystals to not only solve protein structures, but to obtain additional insight that is not available via traditional MX.

Crystal imperfections may arise in a variety of ways and take many forms. They may involve either a variation in the lattice over the extent of the crystal, variation in the contents of the unit cell, or both. In practice, even highly ordered crystals are made up not of a single large region conforming to a particular lattice orientation but of many smaller domains, each locally ordered along the lattice but rotated very slightly to each other. This form of disorder is called *mosaicity*; it has the effect of blurring Bragg peaks as the reflection angles of misaligned domains are all slightly different. Blurring can also result from thermal vibration of the atoms of a crystal as their precise position varies about some central mean, characterised by the debye-waller factor.

The intensities of the Bragg peaks are affected by the relative positions of atoms within the unit cell. This kind of disorder becomes increasingly likely as the molecules that occupy the lattice sites become larger, allowing for greater variation in conformity and rotational alignment. Crystallography software will accommodate a certain amount of variation in Bragg peak size, shape, and intensity, but as the departure from the ideal becomes too great, individual peaks will be rejected in order to minimise the impact of disorder on the analysis. This can limit the resolution of structure solution or even preclude it altogether.

### 2.2.5 Protein Crystals

As noted previously, protein crystallographers are interested in the structure of the protein itself, the crystal is simply an essential step in the process as it acts as an amplifier of the scattered x-rays, allowing the diffraction to be observed. This is in contrast to inorganic crystallography, where the crystalline structure is itself of interest. Although they share many of the properties of inorganic crystals, protein crystals also have many distinguishing features that affect their preparation and use in crystallography.

#### Crystallisation & Characteristics of Protein Crystals

Many inorganic crystals can be found in nature, having formed under conditions appropriate to their adoption of a particular, ordered form. For example, carbon will form diamonds when exposed to the necessary heat and pressure. Others are created through the careful management of their environmental parameters, such as the silicon crystals used to construct the semi-conducting substrates of solid-state electronics. Proteins also rely on a careful tuning of parameters in order to adopt crystalline order. Protein crystals are grown in solutions known as *mother liquors*, which typically contain some combination of salts, polyethylene glycol (PEG), acids, detergents, and other additives. Under the correct conditions, a super-saturation of protein in solution will precipitate out in crystalline form. Changes in pH, temperature, protein concentration, ionic strength and concentration, and other factors will affect the rate of crystal growth and the size of the crystals. In some cases, the type of lattice adopted also varies.

Inorganic crystals are typically composed of densely-packed small molecules (e.g. NaCl) or individual atoms, with strong inter-molecular bonding between lattice sites, rendering them physically robust. A high degree of internal order contributes to stronger diffraction of x-rays, which is also helped by the fact that they tend to be made up of heavier atoms, which have higher cross sections for elastic scattering. By contrast protein crystals are formed by the periodic arrangement of large molecules, composed primarily of light atoms. The inter-molecular bonding between lattice sites is typically weak, and even the intra-molecular bonding that gives the protein its shape is very delicate. The precise alignment of individual proteins can vary relative to their neighbours, limiting the resolution of their diffraction patterns [29]. Proteins are also *chiral* molecules, which limits the symmetries available to protein crystals; only 65 of the 230 space groups may be adopted by protein crystals as all the others allow operations that change the handedness of the molecule.

## Lysozyme

Lysozyme is a component of the innate immune system found in tears, saliva, and human breast milk. It is present in high concentration in chicken egg whites and commercially available lysozyme is often concentrated from this source. Laschtschenko [30] first observed the antibacterial properties of egg whites, for which lysozyme is responsible, in 1909. Alexander Fleming [31], the discoverer of penicillin, observed the same property in nasal mucus in 1922, and is responsible for naming lysozyme. The antibacterial activity of lysozyme, which involves cleaving the peptidoglycan polymers which form a protective mesh around bacteria, is illustrated in fig. 2.9a. Lysozyme was one of the earliest proteins to have its structure solved using MX in 1965 [32, 33]. It is commonly used as a model protein due to the relative ease with which it crystallises. It also has a variety of conditions under which it can be crystallised, and a wide range of crystal sizes can be grown. Examples of tetragonal lysozyme crystals with space group  $P4_32_12$  are shown in fig. 2.9b along with the packing of the tetragonal unit cell (fig. 2.9c). Native lysozyme was the protein used throughout the work presented in this thesis, for both experiments and computational modelling. Lysozyme was chosen as the protein of interest due to its radiation tolerance and relatively low solvent content, characteristics that helped to observe PE escape in our experiments.

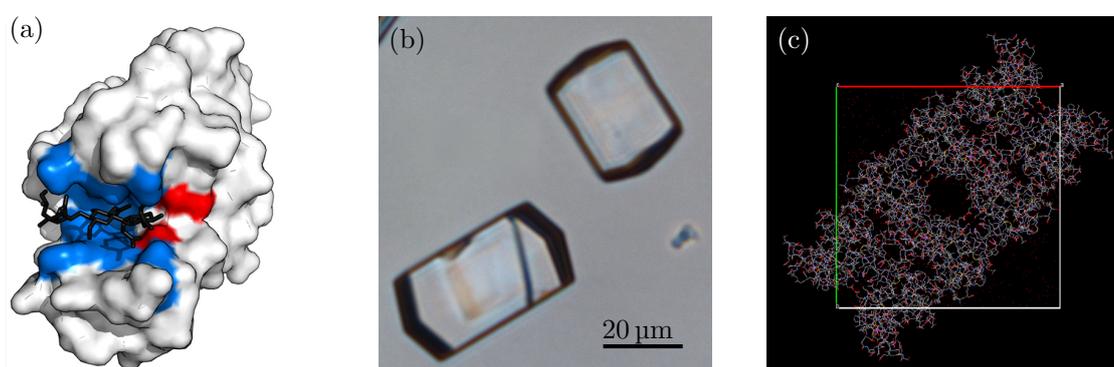


Figure 2.9: **Structure and crystallisation of lysozyme: A model protein.** (a) Molecular surface model rendering of the lysozyme structure showing the binding sites coloured in blue and active (catalytic) site in red. A peptidoglycan substrate is shown in black, bound by the protein. Image by Thomas Shafee, used under a Creative Commons license (CC BY 4.0). Image has been modified by cropping. (b) Optical microscopy image of tetragonal lysozyme crystals with space group  $P4_32_12$ . (c) Model of protein-packing in tetragonal lysozyme crystals. The image is a projection along the  $c$ -axis (37.77 Å) of the unit cell. The  $a$ -axis (red) and  $b$ -axis (green) are of equal length (78.54 Å). Eight protein monomers are visible, highlighting the rotational-, and translational-symmetries of the crystal. A solvent channel is visible in the centre of the unit cell. Image made by the author using CCDC Mercury [34].

## 2.2.6 Synchrotron Light Sources

The majority of MX experiments are now conducted at synchrotron facilities, dozens of which are available to scientists around the world. Modern 3<sup>rd</sup> and 4<sup>th</sup> generation synchrotrons represent the culmination of more than half a century of theoretical development and technological implementation; combined with developments in x-ray detection, robotic automation, and software for interpreting diffraction data, synchrotrons allow experiments that once took months to be completed in as little as several hours.

### Early X-ray Sources

The earliest x-ray sources were vacuum tube devices that incorporated an electron gun (cathode) and a metal target (anode). Electrons emitted from the cathode would be accelerated in the vacuum to collide with the anode, wherein they underwent continuous deceleration and collision with atomic electrons. These tubes would produce a broad-spectrum *Bremsstrahlung* (braking radiation) and fluorescent x-rays at the characteristic energies of the anode. Modern versions of such devices are in common use in laboratories today, though the low intensity and broad spectrum of the radiation they produce makes them a poor choice for MX experiments, compared to the high-intensity, highly monochromated radiation available at synchrotrons.

### Synchrotron Radiation

Acceleration of charged particles results in the emission of electromagnetic radiation. *Synchrotron radiation* is a label given to radiation produced by relativistic charged particles—typically electrons—travelling in curved paths. Synchrotron facilities use linear accelerators (LINACs) to accelerate electrons to close to the speed of light. Magnetic field arrays are then employed to further boost the speed of the electrons, to maintain them at a stable energy and current while containing them to a roughly circular path in the *storage ring*, and to generate radiation with specific characteristics.

### Four Generations of Synchrotron Light Sources

Synchrotron radiation was originally viewed as an inconvenient by-product of the particle accelerators built in the 1940s and '50s for nuclear physics experiments. The observation that this radiation could be harnessed for other types of experiments led to the adaptation of this first generation of synchrotrons to allow for the parasitic use of the radiation produced. From

the 1960s, a second generation of synchrotrons was built, specifically designed to produce synchrotron radiation across a broad spectrum from the infrared to x-rays, for use in experiments across a variety of fields.

In early synchrotron designs electrons travelled in circular paths. These circles were later replaced with a series of straight sections, between which bending magnets were used to accelerate the electrons around corners, into the next straight section. The design of *insertion devices* to be placed in these straight sections heralded the arrival of the third generation of synchrotrons in the late 1980s and early '90s. *Undulators* are linear magnetic arrays that accelerate electrons in oscillatory paths. Undulators produce significantly higher radiated flux than bending magnets; that flux is highly collimated and restricted to a narrow bandwidth, which can be tuned for a given experiment.

MAX-IV [35] in Lund, Sweden, was the first fourth generation synchrotron to come on-line, reporting first light in 2015 [36]. This new generation is defined by the use of *multi bend achromats* (MBAs), increasing the number of dipole magnets in bending magnet arrays from 2 or 3 to 5 or 7, thereby reducing the angle of deflection per magnet and consequently reducing the horizontal emittance of the radiation produced (see below). Several green-field facilities are currently in the planning and/or construction stages, and many more third generation synchrotrons are planning or currently implementing MBA upgrades to their storage rings.

### **Brilliance**

Each generation of synchrotrons has achieved an increase in *brilliance* (sometimes also called *brightness*), a figure of merit that accounts for photon flux, beam source size, beam divergence, and relative energy bandwidth. The units of brilliance are photons  $\text{s}^{-1} \text{mm}^{-2} \text{mrad}^{-2} (0.1\% \text{ BW})^{-1}$ . Brilliance is proportional to photon flux and can thus be increased by increasing flux via an increase in storage ring current. This also leads to an increase in radiated power, however, which places an increased thermal load on accelerator components and beamline optics, often mitigating the gains achieved. The emerging fourth generation of synchrotrons are making gains in brilliance by reducing beam *emittance*,  $\epsilon$ , a measure of beam spread that accounts for size and divergence.

### Ultra Low Emittance Storage Rings

Average brilliance (at a given wavelength,  $\lambda$ ),  $B(\lambda)$ , is related to photon flux,  $\Phi(\lambda)$ , and total horizontal and vertical emittance,  $\varepsilon_{X,Y}$ , by

$$B(\lambda) \propto \frac{\Phi(\lambda)}{\varepsilon_X \varepsilon_Y}. \quad (2.16)$$

The total emittance is the convolution of the diffraction-limited radiation emittance,  $\varepsilon(\lambda)$  and the transverse electron emittance,  $\varepsilon_{x,y}(e^-)$ :

$$\varepsilon_{X,Y} = \varepsilon(\lambda) * \varepsilon_{x,y}(e^-). \quad (2.17)$$

$\varepsilon(\lambda)$  is proportional to  $\lambda/4\pi$  for Gaussian beams [37], and  $\lambda/2\pi$  for typical undulator beams [38]. When  $\varepsilon_x(e^-) < \varepsilon(\lambda)$ , the storage ring is diffraction limited at that wavelength. This is true for all storage rings for long wavelengths (low energies), the label *diffraction limited storage ring* (DLSR) might be said to apply when it is true at wavelengths that are representative of the storage ring; for modern medium- to high-energy storage rings, this means energies on the order of 10 keV. By this definition, the 200 pm rad to 300 pm rad emittance of the MAX-IV [39] and Sirius [40] light sources are still an order of magnitude off the diffraction limited emittance for 1 Å (12.4 keV) x-rays,  $\varepsilon(1 \text{ Å}) \approx 10 \text{ pm rad}$ , though these storage rings are referred to as being diffraction limited.

### MX at Synchrotrons

Early synchrotron MX experiments were conducted in the 1970s [41–43] on the parasitic beamlines of first generation synchrotrons at the Stanford Synchrotron Radiation Project (SSRP, now SSRL) and the Deutsches Elektronen Synchrotron (DESY) [44, 45]. Exposure times were measured in hours. This still represented an improvement over the laboratory sources available at the time, with the achievable flux being higher by two orders of magnitude [41]. In the half-century since, the history of MX has been inextricably linked to the history of synchrotrons. In 2021, there are over 100 MX beamlines operating at synchrotrons around the world [46, 47], and their use is responsible for the vast majority of new submissions to the Protein Data Bank (PDB) [5, 46, 47], as summarised in fig. 2.10. Measurements that once took hours are now routinely accomplished in seconds (see fig. 2.11).

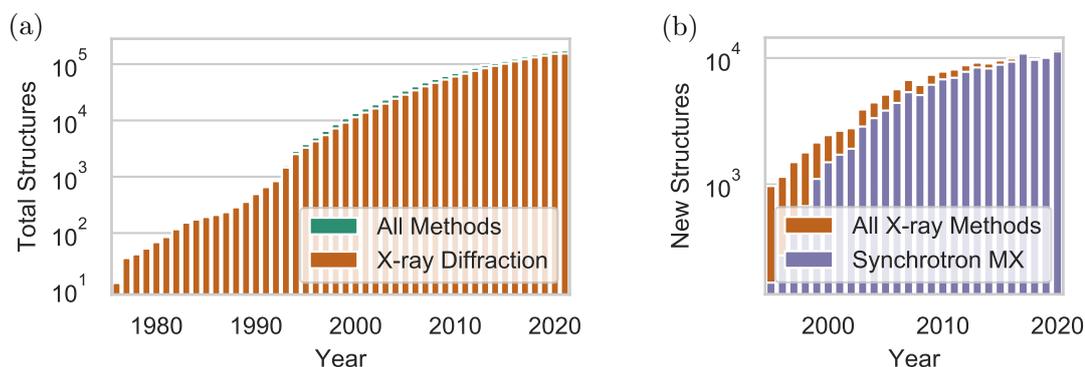


Figure 2.10: **PDB structure depositions.** The graphs demonstrate the importance of x-ray diffraction in general, and synchrotron MX in particular, in determining protein structures. Both graphs use a log scale on the  $y$ -axis. (a) Total structures listed on the PDB as determined by x-ray diffraction (orange) and all other methods (e.g. NMR, electron diffraction) (teal). The data were obtained from the PDB Statistics site [5]. (b) New structure depositions per year as determined by synchrotron MX (lavender) and other x-ray diffraction methods (orange). The data were obtained from BioSync [46, 47].

### MX Beamlines

An example of a typical MX beamline optics chain is shown in fig. 2.12; it consists of a source, monochromator, focusing optics, and collimator(s), the precise layout of which varies from facility to facility, as does the total length of the beamline. The source may be either an undulator or a bending magnet. The monochromator, responsible for selecting a single beam energy (with some nominal bandwidth), is typically a silicon crystal, aligned so that the chosen x-ray energy reflects from the (111) plane [45]. Other crystal planes are sometimes used, though less commonly. By using a pair of crystals, the output beam can be kept parallel to the input beam, and sagittally bending the crystal can provide some auxiliary beam focusing. Focusing is primarily accomplished with curved mirrors that take advantage of the total external reflection of x-rays at grazing angles. A pair of mirrors is commonly configured in the Kirkpatrick-Baez (KB) arrangement, mounted orthogonally to each other to provide independent focusing in the vertical and horizontal planes. The radii of curvature of these mirrors can be anywhere from 200 m to 10 km, adjustable via piezo-electric actuators; the mirrors themselves are typically around a metre long. A collimator placed at the end of the optical chain can be used to restrict beam size beyond the capacity of focusing optics and/or to provide a top-hat beam profile, rather than the native near-Gaussian profile. Several sets of slits may be used through the optical chain to define secondary sources. Any beam shaping using the slits comes at the expense of flux [45].

The endstation of an MX beamline is responsible for sample mounting and positioning, temperature control, and potentially automated sample exchange. Oscillation is a key component of

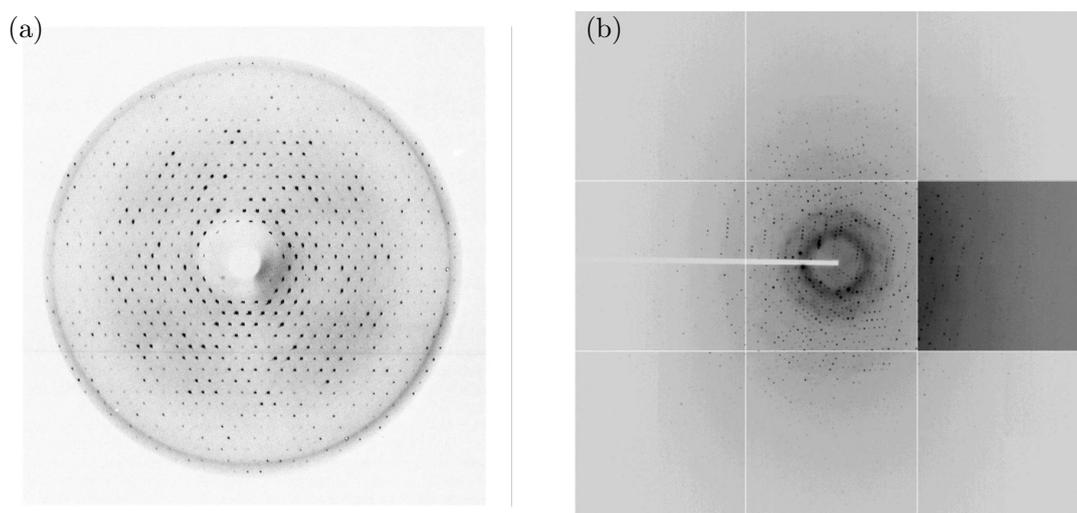


Figure 2.11: **Comparison of diffraction patterns from 1<sup>st</sup> and 3<sup>rd</sup> generation synchrotrons.** (a) Diffraction pattern collected at one of the earliest MX beamlines at Stanford SSRL in 1975. Exposure time was 5 hours. (b) Diffraction pattern collected at the 19-ID beamline of the Advanced Photon Source (APS). Exposure time was 5 seconds with the beam attenuated to 30%. The middle-right panel has been contrast-adjusted to highlight fainter spots at higher resolution. Images from Dauter *et al.* [44]. Reproduced with permission of the International Union of Crystallography.

MX, accomplished by mounting the sample on a goniometer, typically having a single axis of rotation, though multi-axis goniometers are also available [48, 49]. Positioning of the sample can be achieved with the use of on-axis video microscopes (OVAs), which employ mirrors to achieve a ‘beam’s-eye-view’ of the sample. The utility of this approach is limited by beam and sample size, however, when misalignment on the order of a micron can mean the difference between an excellent exposure and missing the sample altogether [45]. In this case, low-intensity raster scanning of the sample can be used to produce a 2D grid of diffraction patterns, which are automatically scored to provide a visual indication of where in the grid the best crystals are located [50, 51]. For cryogenic crystallography, the sample is cooled with a gaseous flow of nitrogen or, less commonly, helium. In the case of room temperature crystallography, a humidity control device can be used to prevent dehydration of the sample [52]. At beamlines optimised for high throughput, sample mounting can be accomplished with the aid of a multi-axis robot, reducing downtime and the need for frequent human presence in the experimental hutch [45].

Several designs for sample holders have been used for MX experiments. Until the advent of cryo-crystallography in the 1990s (see section 2.3.5), crystals were typically mounted in glass or

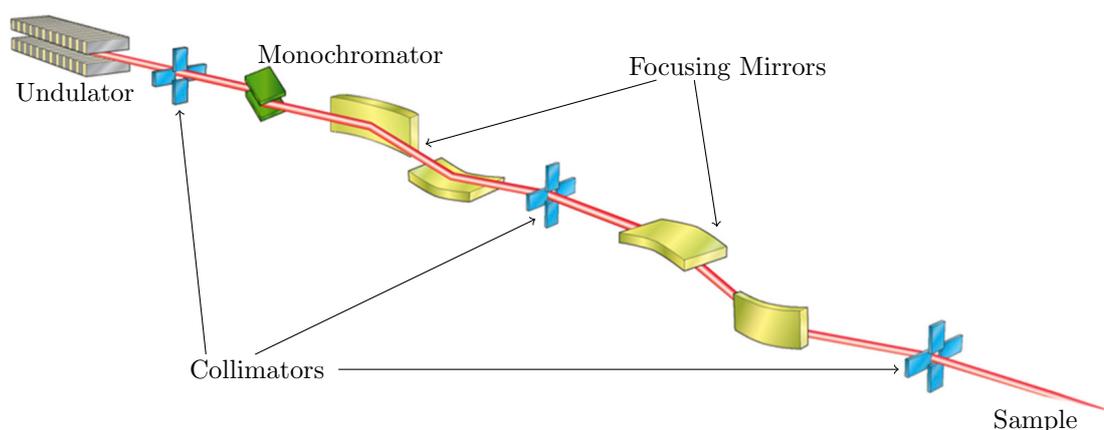


Figure 2.12: **An schematic showing an example of a typical optics setup at an MX beamline.** The key components of an MX beamline optics chain are shown, including source (undulator), monochromator, focusing mirrors, and collimating apertures/shutters. The diagram does not represent the precise layout of a particular beamline. Image from Owen *et al.* [45], used under a Creative Commons license (CC BY 4.0)

similar capillaries to maintain a hydrated environment during data collection [13]. Modern cryo-compatible mounts include nylon loops with micro-meshes [53] and chips with perforated sample windows [54] that hold crystals and a small quantity of surrounding solvent/cryoprotectant. These can be attached to the goniometer with a magnetic mount, allowing fast sample changing. Chips can also be used for SSX by measuring diffraction from several individual crystals at different positions on the chip. Serial techniques also inspired alternative sample delivery systems, where a large quantity of crystals is continually delivered to the beam via a liquid stream. These include jet-designs such as the gas dynamic virtual nozzle (GDVN) [55], though this is suited to XFEL, rather than synchrotron application. Of use in synchrotron MX are high viscosity extrusion jets [56–59] and microfluidic devices [60].

Individual synchrotrons often have multiple MX beamlines; the Advanced Photon Source (APS) at Argonne National Laboratory (ANL) in Illinois has 19 operational MX beamlines. Individual beamlines may be optimised for particular techniques such as MAD/SAD for phasing, powder diffraction, or high throughput. Beam characteristics such as available energy (or energies), flux, and spot size can vary significantly. BioSync [46, 47] publishes a thorough survey of MX beamlines world wide, and similar services are offered, on a more local scale by organisations such as Way For Light [61] which provides a catalogue of European light sources.

## X-ray Detectors

Prior to the late 1990s, diffraction patterns were recorded on photographic films and plates, requiring subsequent development prior to analysis. These have since been replaced by solid-state imaging detectors which operate on the same principles as the visual light detectors used in modern digital cameras. The sensing face of a pixel array detector (PAD) is essentially a 2-dimensional array of photodiodes. X-ray photons are absorbed in the diode, releasing electron-hole pairs; a biasing voltage across the diode causes the charges to move to the cathode/anode, generating an electrical pulse. The pattern of pulses read out from each pixel is then a counter for the number of photons striking it.

The performance of x-ray detectors can be quantified by several metrics, two particularly important ones being the read out time — how long it takes the timing circuits to interpret the pattern of electrical pulses being output by the pixels— and the quantum efficiency (QE) —the percentage of photons of a given energy that are absorbed by the detector’s sensing region. Pixel array detectors (PAD) such as the Dectris Pilatus and Eiger models have read out times of less than a millisecond, allowing data to be collected continuously, rather than having to constantly pause and resume data collection as was the case with earlier semiconductor detectors, which had read out times of 2–3 seconds. Quantum efficiency depends on the photoelectric absorption cross section of the detector material —itself relative to the  $Z$ -number— and its thickness. Until recently, PADs have had silicon ( $Z_{Si} = 14$ ) sensors, whose QE drops rapidly for energies greater than around 10 keV. This can be somewhat mitigated by using thicker sensor panels, but there are practical limits to how thick the panels can be. Recently, a new generation of PADs has emerged, based on cadmium-telluride (CdTe,  $Z_{Cd} = 48$ ,  $Z_{Te} = 52$ ) sensors, which maintain high QE at higher energies. These CdTe detectors are, as of 2021, in the early stages of commercialisation and are not yet common at MX beamlines.

### 2.2.7 X-ray Free Electron Lasers

In simplified terms, an x-ray free electron laser (XFEL) can be thought of a very long undulator, fed by a linear accelerator (LINAC), rather than a synchrotron storage ring. The LINAC produces electron bunches that are much smaller and denser than the bunches that traverse a synchrotron storage ring. As these electron bunches are accelerated through an undulator, they produce radiation; this radiation subsequently impacts the motion of the electrons (in addition to the magnetic field array of the undulator itself) as they continue to move through the undulator. The longer the undulator, the stronger the radiation field is and the more impact it has on the motion of the electrons. The wavelength of the radiation spatially modulates the electron interaction with the radiated field, causing the electron bunch to separate into micro-bunches that are

spaced out from each other by the wavelength of the radiation they are producing. At this point, it is not only the radiation of a single given electron that sums coherently, but the radiation of *all* the electron micro-bunches. This mechanism, called self amplified stimulated emission (SASE), is, as the name suggests, self-enhancing, producing highly coherent radiation with a brilliance ten orders of magnitude beyond what is possible at 3<sup>rd</sup> generation synchrotrons. An undulator design to exploit the SASE principle was first proposed in the 1980s [62, 63]; the LINAC Coherent Light Source (LCLS) was the first operational hard x-ray FEL, reporting first lasing in April 2009 [64].

Due to their incredibly high brilliance, XFELs present the possibility of obtaining useful diffraction data from very small, poorly-diffracting crystals that have previously been unsuitable for use in MX experiments at synchrotrons. *Serial crystallography*, discussed in the next section, is a technique developed at XFELs to achieve this goal. Another key characteristic of XFEL radiation is the incredibly short (on the order of 10 fs), high-intensity pulses of x-ray photons produced by SASE operation, this opens the door for time-resolved studies on a temporal scale not previously available, and has implications for the impact of radiation damage on diffraction data, which is discussed in section 2.3.5.

There are now several FEL facilities in operation, construction, and planning stages around the world, though only a few of them operate in the hard x-ray regime that is useful for MX studies. By comparison to synchrotrons, the number of XFEL facilities is very limited. They make a useful complement to synchrotron facilities but are not intended to replace them as the workhorse facilities for the majority of x-ray science experiments, including MX.

### 2.2.8 Serial Crystallography

In classical MX, a single large ( $\sim 10 \mu\text{m}$ ), highly ordered crystal is mounted on a goniostat and rotated in the x-ray beam, capturing exposures in small increments ( $0.1^\circ$  to  $1.5^\circ$ ), to cover a total rotation of at least  $30^\circ$ . A much higher rotation may be necessary, depending on the symmetries of the crystal's space group. The collated exposures, combined with refined phase information, are used to produce a three-dimensional map of the reciprocal lattice, which can be inverse Fourier transformed with the aid of crystallographic software to determine a real-space model of the protein. Until recently, if a sufficiently large, well-diffracting crystal was not available, structure solution was not possible. Holton and Frankel [65] calculated that under theoretically ideal conditions a complete data set could be obtained from a crystal on the order of 100 nm to  $1 \mu\text{m}$ , allowing that real-world conditions placed the practical limit somewhere between one and two orders of magnitude higher. This reliance on a single large crystal has been somewhat mitigated in recent years via the introduction of *serial crystallography*, a technique that replaces a

single crystal under controlled rotation with a supply of small crystals (on the order of 100 nm to 1  $\mu\text{m}$ ), each of which contributes (at best) a single diffraction image. The diffraction of many such crystals is combined in order to assemble the complete data set previously obtained from a single larger crystal. Between the two extremes described here, there exist other approaches such as combining small oscillation wedges, each composed of several exposures, from multiple crystals. The radiation damage mechanics and potential mitigants explored in this thesis can be applied to multiple methodologies.

Serial crystallography was first achieved at the Linac Coherent Light Source (LCLS) XFEL, where Chapman *et al.* [9] recorded more than  $3 \times 10^6$  diffraction patterns from crystals, ranging in size from 200 nm to 2  $\mu\text{m}$ , of the membrane protein complex photosystem I. The XFEL-based technique is referred to as *serial femtosecond crystallography* (SFX) as each diffraction pattern is produced by a single roughly 30 fs x-ray pulse. Serial crystallography has since been performed at synchrotron sources [10]. The slightly adapted technique is called *serial synchrotron crystallography* (SSX). Both variants of serial crystallography use liquid-based delivery systems to pass a continuous beam of crystals-in-suspension into the path of the beam; a representation of the configuration used by Chapman *et al.* is shown in fig. 2.13.

### 2.3 Radiation Damage in Synchrotron MX

As discussed in section 2.2.4, crystallography relies upon an approximation of perfect crystal structure. With increasing disorder, the crystal deviates further from perfection, and it becomes more difficult to determine its structure. Radiation damage may be thought of as *acquired disorder*; the energy deposited in the crystal by the beam causes changes to the individual structure and position of the proteins, causing the samples to become less crystalline. Given a sufficiently long exposure, even a large crystal with an initially high degree of perfection will cease to provide useful diffraction data.

Understanding radiation damage, especially the rate at which it accrues relative to the rate at which useful diffraction data is recorded, has utility for experimental design and planning. Given the long and costly processes required to actually obtain a crystallised sample of a new target for structure solution, determining the amount of sample required for a complete data set prior to the experiment is critical. A sufficiently detailed understanding of the progression of radiation damage might go further by permitting the retroactive correction of intensity data to facilitate structure solution from highly damaged crystals that are currently unusable.

Radiation damage in MX—and in related imaging fields—is a broad and varied topic, subject to considerable active research. In this section, we cover the key underlying physical principles,

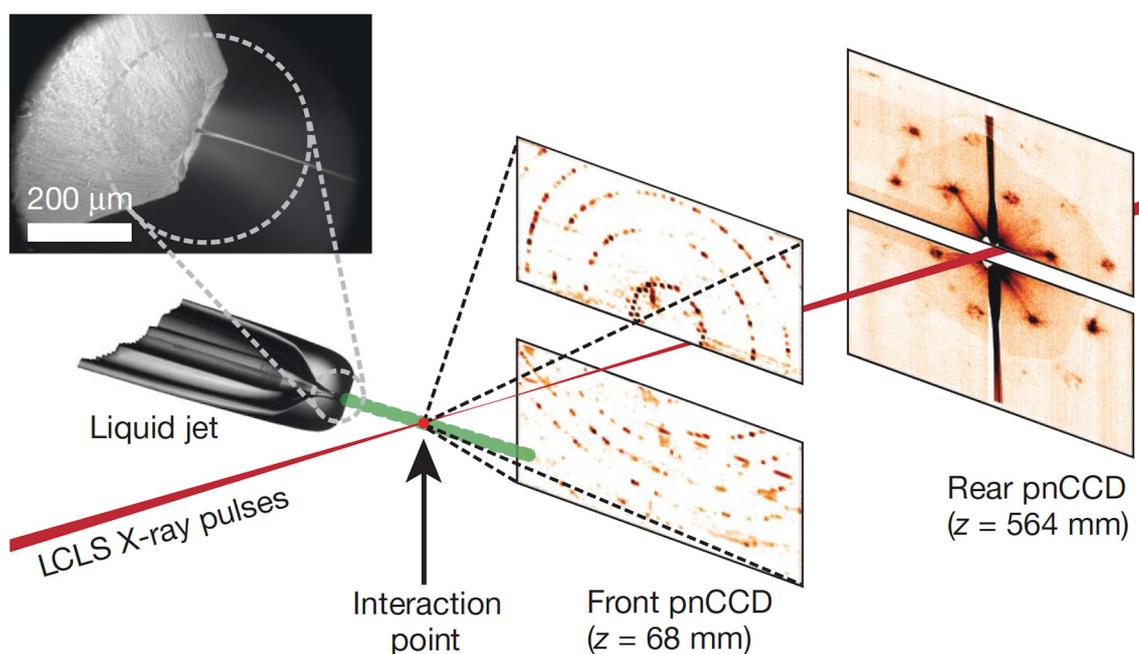


Figure 2.13: **Diagram showing the setup of serial femtosecond crystallography setup at LCLS.** Nano- and micro-crystals in suspension are injected in a continuous stream into the beam interaction region using a liquid jet. High- and low-resolution diffraction data are recorded separately, downstream. The insert shows a close-up of the sample being emitted from the liquid jet. Image from Chapman *et al.* [9], used with permission.

and present a review of the literature that is most relevant to the topics of dose limits, damage modelling, and PE escape in MX with micro-crystals.

### 2.3.1 Dose

Of the small percentage of  $x$ -rays that will interact with a protein crystal, most will contribute to radiation damage through either photoelectric absorption or inelastic scattering. At typical MX energies (around 12 keV), photoelectric absorption is the dominant source of damage, with inelastic scattering becoming important above 20 keV (see fig 2.2). Damage manifests in a variety of ways, some of which are easier to observe than others. Metrics for tracking radiation damage are measured against radiation *dose*, being the energy absorbed per unit mass. The unit of dose is the gray ( $1 \text{ Gy} = 1 \text{ J kg}^{-1}$ ); doses in MX experiments are typically on the order of  $1 \times 10^7 \text{ Gy}$  and are therefore expressed in MGy.

Dose is not typically measured but calculated based on experimental parameters such as the flux density of the beam, and density and interaction cross-section of the sample. As these properties

can be determined (within generally acceptable error-bounds) prior to the experiment taking place, it is possible to estimate the dose a crystal will receive during an experiment and, by extension, how long it will provide usable diffraction data. RADDPOSE-3D [66, 67] is a commonly used software application for predicting dose. The current version of RADDPOSE-3D (v.4) is the result of almost two decades of continuous, iterative development by the Garman lab at Oxford University [68–70].

### 2.3.2 Radiation Damage Metrics

The manifestations of radiation damage can be divided into *specific damage*, which relates to changes in real-space, and *global damage*, which relates to changes in reciprocal-space. Specific damage accrues roughly 60 times faster than global damage [71] and tends to occur in a predictable manner as a function of dose [72–74]: metal centres are reduced, disulfide bonds elongate and then break, and aspartates and glutamates undergo decarboxylation. The earliest manifestations of specific damage involve the heavier elements (metals and sulphur) that have much higher cross-section for interaction. Changes at these important sites lead to a loss of detail/definition in electron-density maps and may quickly impede structure solution [73]. Global damage is seen in unit cell expansion, an increase in scaling B-factors, and a reduction in the integrated intensity of diffraction patterns. Loss of diffracted intensity can be observed in real time, making it a convenient metric to monitor during experiments [71]; it is the primary metric examined throughout this thesis. Figure 2.14 shows an example of intensity loss due to radiation damage.

#### Loss of Diffracted Intensity & Half Dose

The intensity of diffraction patterns diminishes as a function of dose (and, at room temperature, as a function of *dose rate*). As damage accumulates, the arrangement of individual proteins changes relative to each other, reducing the amplifying effect of the crystal structure. Short-range (high-detail) order diminishes more quickly, meaning higher resolution (higher scattering angle) Bragg spots fade more quickly than lower resolution spots. Observations of intensity loss are often made for entire integrated diffraction patterns, but may also be made for resolution shells, individual Bragg peaks, and repeat rotational wedges [74]. In general, this observed loss of intensity is exponential, but there are departures from this behaviour [75, 76], where observed fading is steeper than exponential, shows a plateau prior to exponential fading, or even increases prior to fading. Observing the fading of individual peaks is most problematic in this regard, as any given peak may deviate notably from the average behaviour that would be observed in a resolution shell or entire diffraction pattern containing many peaks. In some cases, deviation

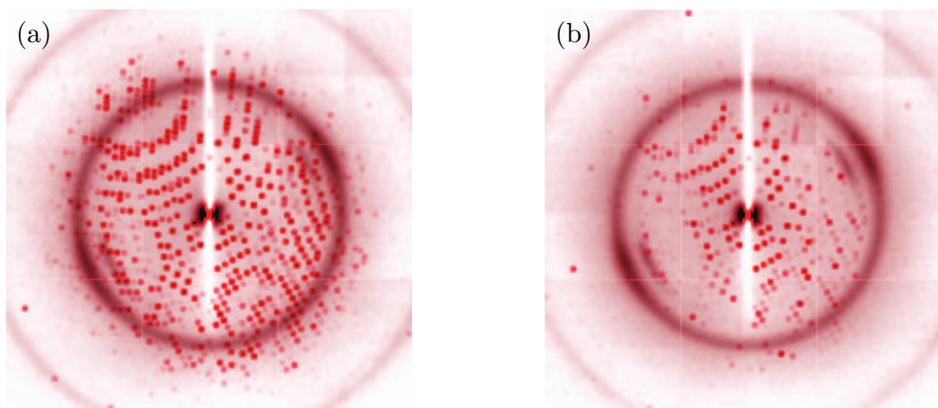


Figure 2.14: **Loss of diffracted intensity due to radiation damage.** Diffraction patterns collected at 100 K showing loss of intensity due to radiation damage. In (a) the dose is approximately 0 MGy, in (b) dose is approximately 80 MGy. (b) highlights the loss of information due to radiation damage, especially at higher resolutions (closer to the edge of the detector). The resolution at the edge of the image is 2 Å. Images were post-processed by adding red circles of appropriate intensity over all dark spots (Bragg peaks). Images from Garman & Weik [74], used with permission.

from an exponential trend in a small number of bright peaks may be sufficient to cause non-exponential fading in the entire diffraction pattern [76]. Deviation from exponential spot fading is discussed in detail in Chapter 5.

Loss of diffracted intensity is typically quantified by the so called *half dose* ( $D_{1/2}$ ), which is the dose required to reduce a diffraction pattern to half its initial intensity [71]. This dose will depend on the range of resolutions present in the pattern, so it is typically stated (for entire diffraction patterns) along with the initial maximum resolution. When referring to the fading of individual peaks or resolution shells, the term *local half dose* is sometimes used, however this terminology is confusing as one might also refer to the dose-state of a subsection of a crystal that is subject to non-uniform dose by the same term. We use the terms *shell half dose* and *spot half dose*, which are interchangeable as (on average) one spot at a given resolution should fade at the same rate as several spots at the same resolution.

### 2.3.3 Tolerable Dose

A tolerable dose could, although with limited utility, be defined as the maximum dose a crystal can absorb prior to being rendered unsuitable for structure solution. In reality, there is no hard cut off as the effects of dose accumulation are continuous and while a certain dose may preclude a structure solution to 1.5 Å, for example, it may still be possible to obtain a structure at 2 Å; whether this is considered success or failure is dependent upon the precise goals of the

researcher. Radiation tolerance can also vary significantly from sample to sample. Particularly with regard to specific damage, how much dose is considered acceptable before useful data can no longer be obtained can be highly variable. Simple estimates of tolerable dose are often based solely on loss of intensity in diffraction patterns, and the term is often used interchangeably with half dose.

The introduction of cryogenic cooling systems in the early 1990s greatly increased dose tolerance in MX; samples held at temperatures from 40 K to 100 K can tolerate roughly 70 times the dose of the same samples at room temperature [77]. In 1990, Henderson [16] proposed a dose limit of 20 MGy, noting that for the synchrotron sources of the time, this dose would require incredibly long exposures to reach. Since then, however, the flux density at synchrotrons has increased to the point that this dose will typically accrue in a matter of seconds, if not less than a second. In the early 2000s, Burmeister [72] measured a half dose of 21 MGy for myrosinase crystals diffracting to 2.0 Å; and Teng and Moffat [78, 79] observed a half dose of 17 MGy for lysozyme crystals diffracting to 1.6 Å. Considering other metrics of damage in addition to spot fading, the latter recommended a dose limit of 10 MGy. Owen *et al.* measured a half dose of 43 MGy for crystals of apoferritin and holoferritin diffracting to 2.3 Å. They also suggested a lower maximum dose of 30 MGy, which corresponded to the diffracted intensity dropping to 70% of its initial value. 30 MGy, termed the *Garman Limit*, has since been frequently cited as a generally advised dose limit, though it was never intended to be taken as such [80]. Leibschnner *et al.* [75] measured half doses of 16 MGy, 18.5 MGy, and 20 MGy for thaumatin crystals diffracting to 1.22 Å, 2.08 Å, and 2.45 Å, respectively. This spread of results highlights the variability in dose tolerance that is seen at different resolutions and across different samples, even when restricted to cryogenic temperatures, serving as a reminder that a maximum dose recommendation is at best a crude, if useful, measure. Based on the recommendations of Teng and Moffat, and Owen *et al.*, the maximum tolerable dose should likely be lower than the half dose, possibly by a factor of up to 2.

Although the majority of modern MX experiments are conducted at cryogenic temperatures, there is a resurgent interest in room temperature MX. Structural models based on room temperature data may be more biologically meaningful [81], and liquid delivery systems for SSX cannot function at cryogenic temperatures. The question of tolerable dose at room temperature is more complicated than at cryogenic temperatures as the rate of damage progression seems to depend on dose rate as well as absolute dose. As noted above, when cryogenic cooling was introduced it extended crystal dose tolerance by a factor of 70, however modern synchrotrons are capable of delivering much higher dose rates and this may reduce the gap somewhat. Recently, De la Mora *et al.* [82] compared half doses of crystals which diffracted to a similar resolution limit at 100 K (1.9 Å) and 300 K (2 Å), using a dose rate of 40.3 MGy s<sup>-1</sup>. The half dose at room temperature

was 0.57 MGy, roughly 4% of the half dose at 100 K (15.3 MGy). This represents a considerable improvement over the factor of 70 difference previously observed. Warkentin *et al.* [76] observed a similar difference between half doses at room temperature and 100 K. Their room temperature results for thaumatin saw a factor of 2 improvement in dose tolerance between dose rates of  $0.095 \text{ MGy s}^{-1}$  and  $35.6 \text{ MGy s}^{-1}$ , however their lysozyme crystals only saw a factor of 1.2 improvement in dose tolerance across a similar range. Given the popularity of cryogenic MX over the last several decades, there is understandably little data available for the impact of dose rate on tolerable dose at room temperature but it is likely to become better quantified in coming years if room temperature experiments continue to increase in popularity.

### 2.3.4 Empirical Models of Spot Fading

Empirical models of radiation damage-induced spot fading have been used since the early days of MX in the middle of the 20<sup>th</sup> century [83]. The chief utility of these models is to estimate, ahead of time, the rate at which diffraction will fade and, based on this, the quantity of sample required for an experiment; and to assist in retroactively correcting peak intensities to improve structure solution. As the relative intensity of Bragg peaks reveals information about the contents of the unit cell (i.e. the electron distribution of the protein), if this relationship changes significantly, it will not be possible to recover the protein structure. Blake and Phillips [83] proposed a model in 1962 that served as the basis for interpreting spot fading for several decades, during which time it was refined by Hendrickson [84], and Sygusch and Allaire [85]. This model was based on the assumption that a partially damaged crystal was composed of several regions: undamaged (virgin, highly-ordered crystal), partially disordered, and amorphous. The second region had a reduced, resolution-limited contribution to diffraction, and the third region did not contribute at all. The evolution of the state of the crystal was therefore described by a rather unwieldy equation involving several exponential terms to account for the transition rates between the different damage states. The details of the equation are presented in Chapter 5. Although the Blake and Phillips model could be fit (with several optimisable parameters) to well describe many experimental results, its assumptions regarding the underlying physical behaviour of the crystal are not consistent with modern understanding and the spot fading behaviours it models can also be modelled by simpler equations, taking into account details of crystal illumination such as Gaussian beam shape [76, 86].

In 2009, Holton [71] suggested a spot fading model based on the comparison of recently published data with a meta-analysis across several disciplines by Howells *et al.* [87], that indicated a linear relationship between resolution and spot half dose. For a Bragg spot with resolution  $d$  Å (or, equivalently, several spots in a shell with the same resolution), the Howells-Holton model

is expressed

$$I(t) = I_0 \exp\left(-\ln(2)\frac{D(t)}{Hd}\right), \quad (2.18)$$

where  $I$  and  $D$  are the spot intensity and dose (in MGy), respectively, after time  $t$ ,  $I_0 = I(0)$  is the initial spot intensity, and  $H = 10 \text{ MGy } \text{\AA}^{-1}$  is the *Howells criterion*. The accuracy of the Howells-Holton model has, more recently, been discussed further in the literature. Atakisi *et al.* [86] claim, based on Gaussian models of radiation damage and a meta-analysis of recent spot fading studies, that spot half dose is proportional to  $d^\alpha$ , with  $\alpha > 1$ . Their model is expressed in terms of the scattering vector  $q = 2\pi/d$ :

$$I(q, D) = I_0 \exp\left(-\frac{D}{K}q^\alpha\right), \quad (2.19)$$

with  $\alpha$  (and  $K$ ) determined by fitting, and expected to be approximately 2. Chapter 5 presents a detailed comparison of these models.

### 2.3.5 Mitigation of Radiation Damage

Various techniques and methodologies have been pursued over the many decades since the inception of MX, with the goal of reducing, or working around the impact of radiation damage. In this section we briefly address the use of cryogenic cooling, radical scavengers, and the use of XFELs to provide diffraction before destruction, before dealing in depth with damage mitigation via PE escape in section 2.3.6.

#### Cryogenic Cooling

As noted in the previous section, cooling crystals to cryogenic temperatures can increase their dose tolerance significantly, making cryogenic cooling (or cryocooling) incredibly effective as a means of mitigating radiation damage. Prior to the 1990s, almost all MX experiments were conducted at room temperature (RT, approx. 298 K). Two technical developments enabled the routine use of cryocooling in MX experiments: the loop mounting method [53], where the protein crystal is held by surface tension in a small diameter (0.1 mm to 1 mm) nylon loop, rather than being sealed in a glass (or similar) capillary, as was typical for RT MX; and the availability of a suitably reliable cryostat [88]. In addition to the radically increased dose tolerance and by extension sample lifetime, cryocooling improved diffraction resolution by preserving crystal order for longer, and reduced background scattering by eliminating the need to seal the sample in a capillary to prevent dehydration [13]. By 2009, it was estimated that over 90% of MX experiments at synchrotrons were performed at cryogenic temperatures [14].

### Free Radical Scavengers

Over the course of an x-ray exposure, the concentration of highly-reactive radical species in the sample increases, leading to a damage cascade. Damaging species include mobile electrons, both PEs and secondary electrons, positively charged holes, and hydroxyl radicals (HO·), though the latter are only mobile at RT [89]. Several studies [90–97] have attempted to mitigate the damage cascade caused by these reactive species by adding *radical scavengers* to the crystal—either by co-crystallisation or subsequent soaking—with the aim that radicals will preferentially react with the scavengers rather than the crystallised protein. The results of these studies are, to date, inconclusive and occasionally contradictory, even when comparing crystals of the same type that have received nominally identical treatment. As such, there remains no consensus on the efficacy of using scavengers to mitigate damage, though it remains an option that may yet yield advantages for specific sample types.

### Diffraction Before Destruction

In section 2.2.8 it was noted that XFELs have been used to conduct serial femtosecond crystallography experiments. The intensity of a single  $\sim 10$  fs XFEL pulse is sufficient to collect a usable diffraction pattern from a single crystal with a diameter as small as 100 nm, prior to the sample being destroyed. This ‘diffraction before destruction’ method [98–100], allows MX experiments at XFELs to ‘outrun’ conventional radiation damage, provided a sufficient quantity of micro- or nano-crystals is available for the SFX approach to be viable (see section 2.2.8). Synchrotrons, like XFELs are pulsed sources, however the pulses are orders of magnitude less intense and of much longer duration (on the order of 100 ps). This means that the signal from many individual synchrotron pulses needs to be integrated to produce a measurable diffraction pattern.

### 2.3.6 The Role of Photoelectrons in Radiation Damage

It was noted in section 2.1 that photoionisation is the primary driver of radiation damage in MX due to both the high cross section relative to Compton scattering, and the fact that *all* of the incident x-ray photon’s energy is transferred to the sample. In this section, we outline the theory behind PE motion within a sample, and examine the possibility that said motion may result in a reduction in radiation damage under specific conditions. The potential for damage mitigation via PE escape is at the core of the discrete simulation presented in Chapter 3 and the experiments discussed in Chapter 4.

### Electron Transport & The Electron Inelastic Mean Free Path

In a photoionisation event, an electron is emitted from a core shell (most likely the  $K$ -shell) with significant kinetic energy. For a typical MX experiment with beam energy,  $E_\Phi \geq 12 \text{ keV}$ , and an average  $K$ -shell binding energy,  $E_K \sim 1 \text{ keV}$ , the resulting PE will have energy  $E_{PE} = E_\Phi - E_K \sim 10 \text{ keV}$ . The PE will subsequently travel through the sample, depositing its energy via a series of inelastic collisions with atomic electrons until it is eventually reabsorbed by the material, or travels beyond its bounds. The penetration depth, or range, of an electron in a material can be estimated by various methods. The continuous slowing down approximation (CSDA), as its name suggests, assumes that the electron loses energy at a continuous rate, which is equivalent to its stopping power within the material. The stopping power is a function of the electron energy and electronic configuration of the medium in which it is travelling. By integrating the reciprocal of the stopping power with respect to energy, the mean penetration depth can be estimated [101]. The distribution of PE energy within a sample will then be a function of radial distance from the point of photoionisation. Penetration depth can also be estimated based on its relationship to the electron's inelastic mean free path.

An electron's inelastic mean free path (IMFP) is the average distance it travels between (inelastic) scattering events. The IMFP is dependent upon electron energy and the density of the material in which the electron is moving. Tanuma *et al.* [102] calculated electron IMFPs for 41 elemental solids, observing that, for all materials, the IMFP curves exhibit a distinctive shape: beyond a minimum—in the region of 10s of eV, up to approximately 100 eV—the IMFP increases with electron energy. At 12.09 keV, the shortest IMFP (for iridium) was 9.16 nm and the longest (for potassium) was 45.78 nm. Shinotsuka *et al.* [103] extended this analysis to liquid water, which is sometimes used to provide a coarse approximation of the behaviour expected of biological samples. They observed a minimum IMFP of 0.954 nm at 73.7 eV and a maximum of 61.6 nm at 29.73 keV; at 12.09 keV the IMFP of liquid water was 29.32 nm. The results of Shinotsuka *et al.* and a representative sample of the results of Tanuma *et al.* are plotted in fig. 2.15a.

Each time a free electron collides with a bound electron, it will lose some energy and be scattered in a new direction. It will then travel a shorter distance (due to its decreased energy) before undergoing another collision. Its total penetration depth from the original photoionisation event, will then be the vector sum of several, gradually shorter IMFPs, as shown in fig. 2.15b. Though the IMFP for all materials considered by Tanuma *et al.* and Shinotsuka *et al.* is on the order 10 nm for energies around 12 keV, the penetration depth of the electrons is much greater (on the order of  $1 \mu\text{m}$  [11, 104]), as they will undergo many collisions, and hence travel many short distances, prior to losing all their energy.

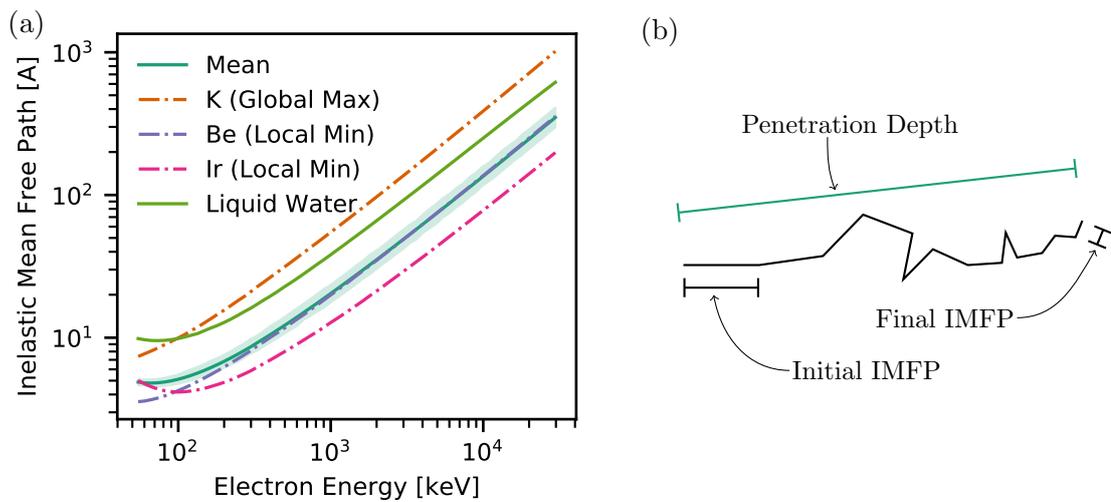


Figure 2.15: **Electron inelastic mean free path (IMFP) and penetration depth.** (a) Electron IMFP for several elemental solids and liquid water, at energies between 50 eV and 30 keV. The solid teal line and shaded region is the mean with 95% confidence interval of 41 elemental solids, as calculated by Tanuma *et al.* [102]. Dashed orange (K), purple (Be), and pink (Ir) lines are from the same study and represent individual elements with the longest and shortest IMFPs; potassium had the longest IMFP at both ends of the energy spectrum, beryllium had the shortest IMFP at low energy, and iridium had the shortest IMFP at high energy. Shown in solid lime is the IMFP for liquid water, as calculated by Shinotsuka *et al.* [103] using the same method as Tanuma. (b) Representation of electron penetration depth as the vector sum of individual IMFPs. As the electron loses energy with each collision, the IMFP is reduced. Energy is deposited in the sample at each collision (change of direction).

### CASINO: Monte Carlo Modelling of Electron Transport in Solids

The transport of electrons through solid material can be modelled with CASINO [105–107], a software application developed to model the behaviour of electrons during scanning electron microscopy (SEM) experiments. CASINO is based on the so called *Monte Carlo method*, named in reference to the Monte Carlo casino in Monaco. The Monte Carlo method uses repeated random sampling and deterministic computation to predict the behaviour of stochastic systems. In the case of electrons traveling through solid matter, the Monte Carlo simulation begins by assigning a direction vector to an electron at a given position, with a given kinetic energy. The length of this vector will be the IMFP; its direction will be random, though potential bounded by theoretical and experimental constraints. The new position defined by this vector represents the first collision of the free electron with an atomic electron; at this point the energy lost to the collision is calculated, based on the material properties being modelled (depending on the complexity of the simulation this value may be a deterministic average or it may be randomly chosen from a

range of possibilities). The energy deposited at this position is recorded and a new direction vector is determined. This new vector will have a shorter length, as the free electron now has less kinetic energy. The process is repeated until the kinetic energy of the electron crosses a defined threshold. This process defines a single possible path, called a random walk, that an electron might follow through the sample. By generating a sufficiently large number of random walks (tens of thousands to millions, depending on the circumstances being modelled and the precision required) the simulation can build a robust stochastic model of how the electrons spread through the material and where their energy is deposited.

### **Photoelectron Escape**

A consequence of PE motion within a sample is that it may escape the sample entirely, given the appropriate conditions, i.e. if the distance between photoionisation and the boundary of the sample is less than the penetration depth of the PEs. If this were to occur, the amount of energy (and by extension dose) deposited in the sample would be reduced, and radiation damage to the sample would be concomitantly mitigated. Until relatively recently, the conditions required for PE escape to lead to a reduction in radiation damage have not been relevant to MX. This situation has changed due to technological advancements at synchrotrons (see section 2.2.6) that have facilitated the use of much smaller crystals (on the order of 1  $\mu\text{m}$ ) and the capacity to focus the x-ray beam to a similar size.

The possibility of leveraging PE escape to reduce damage in MX experiments was proposed by Nave and Hill [11] in 2005. They used the CSDA and Monte Carlo methods (using CASINO) to estimate the penetration depth of photoelectrons with energies of 10 keV, 20 keV, 30 keV and 40 keV. The results of their simulations indicated that the maximum penetration depth of 10 keV PEs is less than 2  $\mu\text{m}$ , increasing to around 8  $\mu\text{m}$  at 20 keV and close to 16  $\mu\text{m}$  at 30 keV. They suggested that energies above 20 keV would be necessary to extract significant benefit for crystals larger than 4  $\mu\text{m}$ , noting that this requirement was outside the optimal parameters of most MX beamlines then operating or in development.

They suggested the possibility of performing a serial MX experiment with many small crystals mounted on a grid, similar to those used in electron microscopy. Key challenges noted were the potential issue of detector inefficiency at high energy negating gains from PE escape; and the need to minimise the volume of solvent surrounding the crystal to avoid PEs being generated there and scattered back into the crystal, off-setting the benefit derived from PEs escaping the crystal. Holton and Frankel [65] included Nave and Hill's two models of PE escape —CSDA and Monte Carlo— in their model of the minimum crystal size required to record a complete data set. The two models suggested ideal beam energies of 16 keV or in excess of 40 keV.

Over the subsequent decade there was growing interest in the possibilities, and specifics, of damage reduction via PE escape. In 2008 Cowan and Nave [12] proposed optimum beam energies in the range 20 keV to 30 keV for a range of crystal sizes from 1  $\mu\text{m}$  to 20  $\mu\text{m}$ , and noted that, for micro-crystals, leveraging PE escape at synchrotron sources might provide a suitable alternative to seeking to outrun damage by using the ultra short pulses at XFELs. Cowan and Nave noted, but did not explore, the fact that PE emission would be concentrated in the direction of the x-ray beam's polarisation vector, which might be useful when working with certain beam and crystal geometries. This idea was subsequently explored in some detail by Stern *et al.* [108] and Finfrock *et al.* [109, 110]. The former proposed focusing x-rays in narrow stripes as a way of separating the diffracting region of the crystal (within the stripes) from the damaged region (between the stripes); based on simulations they predicted an increase in crystal lifetime by a factor of 1.6 in the general case, and 14 in the optimal case of perfectly aligned platelet crystals. Finfrock *et al.* [109, 110] used an 18.6 keV, vertical line-focus beam (with FWHM 2.3  $\mu\text{m}$  and 0.7  $\mu\text{m}$ ) to minimise damage to the diffraction region, while maximising diffraction signal. They observed that the damage profile extended beyond the beam profile and estimated an average penetration depth of  $(5.0 \pm 0.5) \mu\text{m}$  for PEs with an average energy of 17.35 keV.

Sanishvili *et al.* [104] and Coughlan [111] extended the Monte Carlo approach used by Nave and Hill [11] to further explore the effect of PE motion within, and around, the illuminated volume of a protein crystal, in particular the effect of anisotropic emission of PEs due to the polarisation of the x-ray beam. Sanishvili *et al.* also performed experiments with large crystals and micro-focus beams (FWHM 0.87  $\mu\text{m}$  to 15.6  $\mu\text{m}$ ), observing a reduction in radiation damage by a factor of 3 between the largest and smallest beams. They estimated the PE penetration depth to be 4  $\mu\text{m}$  for a 18.5 keV x-ray beam, which concords with the observations of Finfrock *et al.*.

The nature of PE escape in micro crystallography and, in particular, the validation of theoretical models via experimental results, is very much an active area of research; our review here is intended to encompass the current literature, during the early stages of the research presented in this thesis (2017 through early 2018). Where new evidence has been published during the course of this research, we have included in-depth discussion in Chapters 3, 4, and 5.

## 2.4 Summary

We have presented, in this chapter, an overview of the fundamental theory and review of the relevant literature, intended for a non-expert graduate in the physical sciences. Section 2.1 covered the nature of x-ray–matter interactions that underpin the interpretation of macromolecular x-ray crystallography experiments. Section 2.2 expanded on the link between the underlying physics

discussed in the previous section and actual practice of MX, including an overview of the nature of protein crystal samples, and the synchrotron facilities that are responsible for the majority of MX experiments. Finally, in section 2.3, we discussed the nature and effects of radiation damage in MX, which is the primary focus of the research presented in this thesis.

In Chapter 3, we present the design and execution of a discrete simulation of radiation damage, based on the theory discussed in section 2.3.6.

## Chapter 3

# A Spatially- and Temporally-Resolved Model of Radiation Damage in Macromolecular Crystallography

In this chapter we describe the development of a spatially- and temporally-resolved computational simulation that models the effect of radiation damage on a crystal during an MX experiment under synchrotron conditions. This simulation was used to examine the effects of beam size, beam energy, and crystal size on the rate at which radiation damage accrues under two opposing assumptions: a) PEs do not escape the crystal —or, more generally, the diffraction volume—, and b) PEs *do* escape the diffraction volume. In the discussion, we explore the implications of our simulated results for MX experiments employing micro crystals and micro beams. The bulk of this chapter comprises a paper entitled ‘The Influence of Photoelectron Escape in Radiation Damage Simulations of Protein Micro-Crystallography’ that was originally published in 2018 in the journal *Crystals* [21]. Page-number references in this chapter are to the page numbers of the paper, not the page numbers of the thesis.

### 3.1 Introduction

In section 2.1, we noted that x-rays interact with matter in three ways in the energy regime relevant to MX: elastic scatter, Compton scatter, and photoionisation. Elastic scatter is responsible for the diffraction patterns that reveal the electron-density of the proteins, while Compton scatter and photoionisation transfer energy to the sample, damaging it. This damage manifests in various ways, as discussed in section 2.3; of interest in this chapter is the loss of intensity in Bragg peaks. In the context of diffraction, crystals can be thought of as *amplifiers*: if the molecules are perfectly aligned then the diffraction signal from each molecule sums coherently. However,

if the molecules making up the crystal do not well resemble each other (e.g. by being aligned differently, or slightly different conformations), then their individual molecular transforms are different to each other and the amplifying effect is reduced and eventually lost. Compton scattering and photoionisation change the electron-density of the molecules that they affect, resulting in a gradual decline in the order of the crystal and therefore diffracted intensity. In addition to the overall intensity of diffraction decaying over time, the rate at which individual Bragg peaks fade is proportional to their resolution, with higher resolution peaks fading more quickly. As the electron density of the molecule manifests in reciprocal space in the relative intensity of Bragg peaks —while the crystalline structure manifests in the spacing between the peaks— loss of intensity can affect the chances of structure solution even before the overall intensity is significantly diminished.

The usable lifetime of protein crystals is radically extended by holding them at cryogenic temperatures ( $\leq 100$  K) during x-ray exposure. Cryocooling has the effect of restricting the motion of protein molecules, keeping them better aligned throughout the exposure, and slowing the diffusion of radical species that deposit energy in the crystal. At cryogenic temperatures, fading of Bragg peaks is proportional to dose (energy absorbed per unit mass), measured in gray ( $1 \text{ Gy} = 1 \text{ J kg}^{-1}$ ), independent of the rate at which dose accrues [71]. However, even as routine cryocooling was becoming possible, 3<sup>rd</sup> generation synchrotrons began to come on line (beginning with ESRF in 1994 [112]), offering an increase in flux density that allowed experimenters to overcome the weaker diffracting power of smaller, less perfect biological crystals, while simultaneously delivering killing doses in seconds [45]. Indeed, the flux density of 3<sup>rd</sup> generation synchrotrons is sufficient to induce radiation damage even in small-molecule crystallography [18]. With 4<sup>th</sup> generation synchrotrons now coming on line [36], flux density has further increased, bringing a concomitant reduction in crystal lifetimes even under cryogenic conditions. XFELs may offer the promise of outrunning radiation damage via high intensity femtosecond x-ray pulses, but in synchrotron MX, the workhorse methodology for structure solution, radiation damage remains a pressing concern and an area of active research.

With radiation damage remaining a limiting factor in structure solution, it is important to be able to estimate, prior to an experiment, the rate at which damage is likely to accrue and therefore the amount of sample that is likely to be required. RADDPOSE-3D [66, 67] is the most recent iteration of RADDPOSE, a mature application that has been shown to produce realistic estimates of dose under the conditions typical in classical MX, e.g. crystals and beam sizes on the order of  $10 \mu\text{m}$  to  $100 \mu\text{m}$ . While such conditions are still used in MX, technological improvements at synchrotrons mean that much smaller crystals and beams (on the order of  $1 \mu\text{m}$ ) are now becoming common, relying on serial techniques [10] to generate a complete crystallographic diffraction dataset.

A growing body of research [11, 12, 104, 109, 110] has suggested that when crystal samples and beam foci are small enough (around 2  $\mu\text{m}$  to 5  $\mu\text{m}$ ) some of the PEs produced in the crystal will escape it prior to depositing all their energy, resulting in a reduction of dose to the crystal. The details of this process are described in section 2.3.6. At the time that the simulations presented in this thesis were developed, RADDPOSE-3D produced a spatially-resolved map of damage, enabling the user to account for non-uniform illumination of the sample, but did not include PE escape, and thus potentially had limited accuracy when dealing with micron-sized beams and crystals.

As the IMFP of electrons, and therefore their penetration depth, is energy dependent, higher energy PEs are more likely to leave the crystal with a higher portion of their initial energy. The ionisation potential of a protein does not change significantly between 10 keV and 30 keV (see table 3.1), being on the order of 1 keV, so any increase in the beam energy translates to the same increase in PE energy. MX experiments have typically used roughly 12 keV x-rays as this translates to a maximum theoretical resolution of 1  $\text{\AA}$ , the length scale of inter-atomic bonds. In practice, resolution doesn't meet this theoretical maximum due to technological constraints, and any increase to theoretical resolution limit via reduced x-ray wavelength is not deemed to be useful. The possibility of PE escape however, suggests that there is potential utility in conducting experiments at higher energies, though this is not directly related to improving resolution.

The detailed simulation of radiation damage presented in this chapter examines the implications of PE escape on MX experiments with micro-crystals through parallel simulation streams, two of which neglect PE escape, and one of which includes it. We examine the impact of changes in crystal and beam size (in the range 1  $\mu\text{m}$  to 5  $\mu\text{m}$ ) and beam energy (in the range 9 keV to 30 keV). Our results can help inform a decision as to whether the additional difficulties of working at higher energies and with smaller crystals and beams are worth pursuing.

## 3.2 Simulation Design

An overview of the simulation design and execution is given in the 'Materials and Methods' section (section 4 (p. 13) of the paper, itself section 3.4 of this chapter); here we describe the top-level design and discuss some components of the simulation in further detail, intended to be accessible to a less-specialised audience than the paper itself.

### 3.2.1 Top-Level Structure

The top-level logical structure of our simulation is presented in a simple flow diagram in fig. 3.1. At this level, the simulation is divided into three modules, which are executed sequentially:

1. Monte Carlo simulation of electron trajectories in the sample,
2. Spatially-resolved modelling of dose distribution,
3. Time-resolved modelling of the loss of integrated intensity of Bragg spots.

Within each of these modules, one or more processes are executed —sequentially, in parallel, or both— in order to provide the input for the next module. A detailed view of the modules, and their inputs and outputs, is shown in fig. 6 of the paper (p. 14). Two third-party applications were incorporated into our simulation pipeline: CASINO [105–107], an application developed to model electron transport in solids during scanning electron microscopy (SEM) experiments; and RADDPOSE-3D [66, 68–70], an existing –and mature– application for predicting dose in MX experiments. In both cases, the latest version of the software available at the time (mid 2017) was used. MATLAB was used to script the majority of the simulation.



Figure 3.1: **Flow diagram showing the top-level organisation of our simulation of radiation damage.** The first module (teal) of our simulation is the Monte Carlo component, executed via the third-party application CASINO. The second module (orange) processes the MC outputs to produce spatially-resolved maps of dose. It incorporates MATLAB scripts and RADDPOSE-3D. The third module (pink) is entirely scripted in MATLAB. It uses the dose maps from the second module to produce time-resolved spot fading predictions.

### 3.2.2 Monte Carlo Electron Trajectory Model: CASINO

CASINO [105–107] is a Monte Carlo modelling application for scanning electron microscopy (SEM) and electron-beam lithography (EBL), designed to model low-energy electron beam interactions in bulk and thin foil samples [107]. CASINO has been used previously to model the energy deposited by PEs in protein crystals [11, 12, 104], which is the primary cause of global damage during x-ray crystallography experiments. Although not designed specifically for this purpose, CASINO’s Monte Carlo simulation of electron paths in solids can provide an approximation of the behaviour of PEs in a protein crystal once they have been produced. Previous studies have obtained rough estimates of penetration depth and the approximate spread of PEs

but have restricted modelling to CASINO alone, rather than post-processing the CASINO outputs to create a detailed spatially-resolved model as we have done here.

CASINO takes inputs that define the electron beam and the sample. Beam inputs include the electron energy (with the option to run several different energies in a single execution), beam radius (the beam is modelled as a Gaussian distribution), the number of electrons to be simulated, and the angle at which the beam is incident upon the sample. The sample is defined by chemical composition, density, and thickness. Several outputs are available; the only one relevant to our simulation pipeline is a spatially resolved map of energy deposition by position.

In the context of our simulation pipeline, the electron trajectories produced by CASINO are interpreted in the following way: the simulated trajectories represent the statistical distribution of paths followed by electrons produced *at a particular point in space* (in CASINO, the intersection of the electron beam with the sample) and *moving initially in a particular direction* (in CASINO, straight down, along the  $z$ -axis). Post processing is required to account for the fact that:

1. PEs can be produced *anywhere* within the x-ray beam footprint, with the likelihood of photoemission at a given point being a function of the beam intensity at that point and the sample composition;
2. PEs emitted at a given point in space have a probability of being emitted in a given direction that is dependent upon the polarisation of the x-ray beam. i.e. for a linearly polarised beam, the probability of emission in a particular direction goes as  $\cos^2(\theta)$ , where  $\theta$  is the angle made to the polarisation vector.

Both these requirements are handled in our simulation by the second module: ‘Space-Resolved Dose Mapping’. In reference to the flow diagram in the paper, the relevant processes are ‘Model Pt Src’, ‘Model Beam’, and ‘Model Energy Deposition’.

### 3.2.3 Space-Resolved Dose Mapping

This module forms the bulk of our simulation pipe-line. In order to construct a spatially-resolved picture of crystal damage, we modelled the interaction region—and a buffer region beyond it—with arrays, where each element (voxel) represented the flux, dose, etc. in a small volume of the crystal and its surroundings. This approach allowed us to account for the (Gaussian) beam profile and the spread of PE energy about the the origin of a photoionisation event (by post-processing the energy deposition data produced by CASINO). The spatially-resolved damage maps produced were then used in the final stage of our simulation to analyse the effect of damage accrued over time.

There are two key sub-steps that contribute to the spatially resolved model of energy deposition in the entire interaction region:

1. Modelling the average energy deposition of a single PE about the point at which it is produced;
2. Modelling the beam and, by extension, the interaction rates between the beam and the sample (and surrounding material).

### Modelling the Average Energy Deposition of A Photoelectron

Our model of the deposition of energy due to a photoionisation event takes as input the *energy deposition by position* output of CASINO, the x-ray beam energy,  $E_\gamma$ , and the average  $K$ -shell energy of the sample,  $E_K$ . The script's output is a 2-dimensional map of the average energy deposited in a planar cross-section of the sample —perpendicular to the beam axis— due to a single photoionisation event. The map is centred on the point at which the photoionisation event occurs.

The input and output of this script are illustrated as heat maps in fig. 3.2 for a beam energy of 12 keV. The input array,  $\mathbf{E}_0$ , is a copy of the CASINO output, translated such that the electron origin is centred in the array. In steps of  $1^\circ$ ,  $\mathbf{E}_0$  is rotated and scaled by  $\cos^2(\theta)$ , where  $\theta$  is the angle made to the  $x$ -axis (i.e. simulating the emission probability of a horizontally-polarised x-ray beam). A sequence of arrays,  $\mathbf{E}_\theta$ , is thus generated for  $\theta = 0, 1, 2, \dots, 359$ . The output array,  $\mathbf{P}$ , is the sum of these  $\mathbf{E}_\theta$ , scaled so that its elements sum to the average PE energy,  $E_{e^-} = E_\gamma - E_K$ :

$$\mathbf{P} = \frac{E_{e^-}}{\sum_0^{359} \|\mathbf{E}_\theta\|} \sum_{\theta=0}^{359} \mathbf{E}_\theta \quad (3.1)$$

### Modelling the Beam

We assumed a Gaussian beam. In the execution of our simulation, the beam was circularly-symmetric, but the sub-routine was scripted to permit different beam widths in the horizontal and vertical directions. Inputs to this script are beam intensity,  $I_\phi$ , and beam size (FWHM) in each direction,  $\sigma_X$  and  $\sigma_Y$ . Output is an array representing the spatially-resolved photon flux of the beam footprint. Our simulation assumes that transmission is sufficiently high to ignore absorption, so the same beam cross-section is applied at any depth of the sample.

MATLAB's built-in normal probability distribution function is used to generate two one-dimensional Gaussian distributions, one in a column vector,  $\mathbf{g}_Y$  the other in a row vector,  $\mathbf{g}_X$ . The

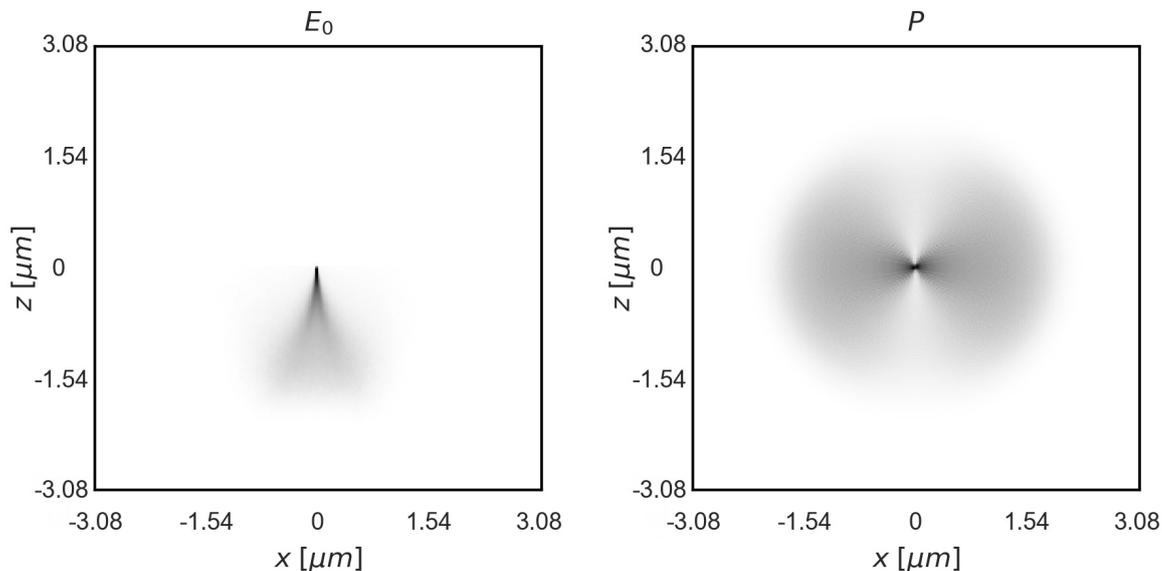


Figure 3.2: **Modelling a point source of photoelectrons (PEs) in the sample.** The heat maps show the input and output of the subroutine that models the average energy deposition of a single PE. Heat maps have been post processed to simultaneously provide good visibility of the shape of the energy distribution and the variation in energy between the centre of the distribution and its edges. If the value of a given pixel is  $x_{i,j}$ , the corresponding value shown in the plot is  $\ln(x_{i,j} \times 10^4 + 1)$ .  $E_0$  is a copy of the energy deposition by position map produced by CASINO, translated to place the PE origin at the centre of the array.  $P$  is the processed energy deposition map, representing the spatial distribution of energy deposited (on average) by a single PE produced in the centre of the array, taking into account the effect of beam polarisation on the likelihood of PE emission in a given direction.

cross-product of these vectors is a two-dimensional Gaussian array,  $G = g_Y \times g_X$ , which is scaled to the beam intensity,  $I_\phi$ , to produce the array representing the beam:

$$B = \frac{I_\phi}{\sum_{i,j} G^{(i,j)}} G \quad (3.2)$$

### Modelling Interaction Rates

We modelled the three types of interaction (photoionisation, Compton scattering, and elastic scattering) that are most relevant to MX. The rates at which these interactions occur are a function of the beam flux, and the sample cross-sections and thickness. This script takes as input the Gaussian beam array,  $B$ , produced in the previous step, the (energy- and sample-dependent) interaction cross-sections for photoionisation,  $\sigma_P$ , coherent scatter,  $\sigma_C$ , and incoherent scatter,  $\sigma_I$ , and the sample density,  $\rho$ . The thickness of the sample is determined by the pixel size, i.e. if the individual pixels are  $20 \text{ nm} \times 20 \text{ nm}$ , then  $t = 20 \text{ nm}$ . It outputs three arrays,  $R_P$ ,  $R_C$ , and  $R_I$ ,

being the respective rates at which these events occur, given the sample and beam parameters used in the simulation.

The general case of the computation is

$$\mathbf{R} = \sigma \rho t \mathbf{B}. \quad (3.3)$$

This computation is repeated for each interaction.

### Modelling the Total Energy Deposited in the Interaction Region in the Absence of Photoelectron Escape

This script takes as input the interaction rate arrays for photoionisation and inelastic scattering,  $\mathbf{R}_P$  and  $\mathbf{R}_I$ , produced by the previous subroutine; the PE energy,  $E_{e^-}$ ; the Compton recoil energy,  $E_{Compt}$ ; and an energy-dose conversion factor,  $\kappa$ , which accounts for the ‘mass’ of a voxel. The output is an array,  $\mathbf{X}_{NoPE}$ , representing the dose deposited in the beam footprint when PE escape is not taken into account.

As no spreading of PE energy is modelled in this case, if a photoionisation event—or an inelastic scattering event—occurs in a given voxel, all of the energy involved is deposited in that same voxel. The output array is therefore calculated as

$$\mathbf{X}_{NoPE} = \kappa (E_{e^-} \mathbf{R}_P + E_{Compt} \mathbf{R}_I) \quad (3.4)$$

### Modelling the Total Energy Deposited in the Interaction Region Accounting for Photoelectron Escape

The inputs to this script are the same as the inputs to the script that ignores PE escape, with the addition of the array,  $\mathbf{P}$ , representing the energy deposition due to a single photoionisation event. The output is an array,  $\mathbf{X}_{PE}$ , representing the dose deposited in the beam footprint when PE escape is taken into account.

This case is more complex than the case where PE escape is ignored as it must account for both the likelihood of a photoionisation event occurring (given by the values of  $\mathbf{R}_P$ ) and the distribution of energy around the point where a photoionisation event does occur (given by the values of  $\mathbf{P}$ ). If  $\mathbf{P}$  is an  $(m \times n)$  array and  $\mathbf{R}_P$  is a  $(p \times q)$  array, with  $m < p$  and  $n < q$ , the output

array, which is also  $(p \times q)$ , is calculated as

$$\mathbf{X}_{PE} = \kappa \left( E_{Compt} \mathbf{R}_I + \sum_{i=m/2}^{p-m/2} \sum_{j=n/2}^{q-n/2} R_{p(i,j)} \mathbf{P} \right). \quad (3.5)$$

The execution of the double sum in eq. (3.5) can be understood as a discrete two-dimensional convolution (or cross-correlation) of  $\mathbf{P}$  and  $\mathbf{R}_P$ .

### RADDOSE-3D

RADDOSE-3D was used as a point of comparison to the results produced by our simulations when PE escape is not taken into account. The dose maps output by the version of RADDOSE-3D tested here are three-dimensional and do not account for Compton scattering, but are otherwise expected to be similar to the 2D dose maps produced using our MATLAB code.

RADDOSE-3D takes a large number of inputs that define the sample, beam, and any rotation to be modelled. Key inputs are the crystal dimensions, heavy atoms in the protein and solvent, solvent fraction, beam energy, and beam profile. A complete set of input parameters can be seen in Appendix A, which contains an example RADDOSE-3D input file. RADDOSE-3D produces as output a three dimensional array of the crystal volume,  $\mathbf{X}_{RD}$ , providing a spatially resolved map of dose distribution. It also produces a summary file showing average global dose, total diffracted intensity, and several other figures of merit.

### 3.2.4 Time-Resolved Spot Fading

Our simulation models the intensity of an individual Bragg spot with resolution  $d$  (or a shell of spots with the same resolution), according to the *Howells-Holton* empirical spot fading formula (eq. (2.18)). At the cost of extended computation time, our simulation can be extended to model the intensity of the entire diffraction pattern by repeated calculation of eq. (2.18) for many values of  $d$ .

The inputs of the spot fading script include the resolution of the Bragg spot(s) being modelled,  $d$ , and the dose arrays,  $\mathbf{X}_{NoPE}$ ,  $\mathbf{X}_{PE}$ , and  $\mathbf{X}_{RD}$ , produced by the subroutines for modelling energy deposition in the interaction region, and RADDOSE-3D. More strictly, the input arrays are sub-slices of the dose arrays. The dose arrays are originally calculated for the entire interaction region which extends to the edge of the beam or beyond; the arrays used at this step are sliced based on the size of the crystal being simulated. Our scripts require that the crystal be no larger than the interaction region, and default to the assumption that the crystal is centred in the

beam, though it is possible to select a different position for the crystal, within the bounds of the interaction region. Output of this subroutine is a collection of time series, recording the relative intensity of the Bragg spot(s), with one series for each input dose array.

Based on eq. (2.18), a critical dose,  $D_C = Hd$ , is calculated. For each dose array,  $\mathbf{X}$ , a global dose rate,  $\Delta D_G$  is calculated for the whole crystal, and an array to track the damage rate,  $\mathbf{D}$ , is created, with  $\dim(\mathbf{D}) = \dim(\mathbf{X})$ . For the outputs of our own scripts,  $\mathbf{X}_{NoPE}$  and  $\mathbf{X}_{PE}$ , the global dose is the same for a single slice as our simulation ignores beam attenuation through the sample, therefore the energy absorbed by a slice relative to the whole crystal scales by the same factor as the mass. The total simulation time,  $t_{MAX}$ , is defined as the maximum time required for the crystal to absorb twice the critical dose (i.e. the time taken to reach this dose at the lowest dose rate):

$$t_{MAX} = \frac{2D_C}{\min(\Delta D_G)}. \quad (3.6)$$

The total simulation time is then split into 300 individual time steps ( $\Delta t = t_{MAX}/300$ ).

The damage-tracking arrays,  $\mathbf{D}$ , are initiated as matrices of ones. As the time simulation progresses, each element takes on a value between 0 and 1, with 1 indicating that section of the crystal is entirely undamaged and 0 indicating that it no longer contributes at all to the diffracted intensity of the Bragg peak(s). At each time step,  $t_n = n\Delta t$  for  $n \in [1, 300]$ , the value of each element of  $\mathbf{D}$  is calculated using eq. (2.18), which, with the relevant substitutions, becomes

$$D_{i,j} = \exp\left(-\ln(2)\frac{t_n X_{i,j}}{D_C}\right). \quad (3.7)$$

The elements of  $\mathbf{D}$  are then summed to give the damage state of the crystal at  $t_n$ , and this value is normalised with respect to the value at  $t = 0$ . This damage state is taken as a representation of the relative Bragg peak intensity produced by the crystal, and decays exponentially as a function of time.

The time simulation is run for each of the dose rate arrays,  $\mathbf{X}_{NoPE}$ ,  $\mathbf{X}_{PE}$ , and  $\mathbf{X}_{RD}$ . The implementation is slightly different for  $\mathbf{X}_{RD}$ , as this is a 3-dimensional array, while  $\mathbf{X}_{NoPE}$  and  $\mathbf{X}_{PE}$  are 2-dimensional arrays, but the logical process and output are identical for all three cases.

### 3.3 Simulation Execution

We used our simulation pipeline to model the loss in Bragg peak intensity of lysozyme crystals, a common model-system used in protein crystallography. We varied crystal size, beam size, and

Table 3.1: **Sample parameters used in the simulation.** The upper entries show the chemical composition and density of the solvent and the crystal, including the solvent fraction. These parameters defined the sample in the MC stage of the simulation. The lower entries are the ionisation potentials of the sample at each of the beam energies that were simulated. These values are the weighted average of the  $K$ -shell ionisation potentials of the sample constituents.

Material	Chemical Composition	Density
Solvent	$C_{32}H_{134}O_{52}NaCl$	$1.184 \text{ g cm}^{-3}$
Crystal	$C_{766}H_{1603}N_{193}O_{435}S_{10}NaCl$	$1.259 \text{ g cm}^{-3}$

X-ray Energy [keV]:	9	12	15	20	25	30
Ionisation Potential, Solvent [keV]:	1.11	1.15	1.17	1.20	1.23	1.25
Ionisation Potential, Crystal [keV]:	1.19	1.23	1.27	1.31	1.33	1.36

beam energy in order to analyse the effect of these parameters on the rate at which radiation damage accrues under the opposing assumptions that PEs escape the diffraction volume or not. In the paper (section 3.4), the design and execution of the simulation are covered together. Here, we highlight key components of the execution in greater detail.

### 3.3.1 Simulating a Crystal Sample: Chemical Composition & Weighted Cross Sections

Our simulated sample was crystallised lysozyme, suspended in a 1:1 mixture of polyethylene glycol with an average molecular weight of 400 (PEG400) and 0.5 M acetic acid with 9% NaCl (w/v), following crystallisation and sample preparation protocols by Coughlan [27]. The solvent and the crystal —itself containing a solvent fraction of 36%— were modelled as two separate samples. The modelled parameters, which are summarised in table 3.1, were chemical composition, density,  $\rho$ , and average  $K$ -shell ionisation energy.

The chemical composition and density of the solvent-cryoprotectant mixture are simply calculated based on the proportions of the constituents. The formula is an average, as the formula for PEG400 is, itself, an average. We calculated the chemical composition and density of the crystal based on the contents of the unit cell. The sample we modelled was tetragonal lysozyme, with space group  $P4_32_12$ , containing 8 lysozyme molecules per unit cell. Based on the amino-acid sequence of lysozyme, the volume of the unit cell, and the chemical composition of the solvent, we determined the chemical composition of the hydrated crystal. The density was calculated based

on the solvent fraction, the density of the solvent, and the density of dehydrated lysozyme, determined by White *et al.* [113] to be  $1.36 \text{ g cm}^{-3}$ .

The average ionisation potential of a complex chemical compound or mixture is not simply the numerical average of the ionisation potentials of the constituent atoms; some atomic species are much more prevalent than others, and some have much higher cross sections for interaction. For example hydrogen is more common than the other atoms in a protein but barely interacts with x-rays at all, and is frequently ignored in simulations. In contrast, sulphur has a cross section for photoionisation (at 12 keV) that is 10,000 times that of hydrogen, and 20 times that of carbon, and its ionisation potential is two orders of magnitude greater than hydrogen and an order of magnitude greater than carbon, however its prevalence is roughly  $1/70^{\text{th}}$  that of carbon. To determine an average  $K$ -shell ionisation potential,  $E_{K,\varepsilon}$  at each beam energy,  $\varepsilon$ , weighted by elemental cross section and chemical composition, we used the  $K$ -shell binding energies of each relevant element,  $E_{K,i}$ ; the elemental cross sections,  $\sigma_{\varepsilon,i}$ ; and calculated the mass density of the individual elements within the crystal,  $\rho_i$ . The average binding energy is then

$$E_{K,\varepsilon} = \sum_i \frac{\sigma_{\varepsilon,i} \rho_i}{\sum_i (\sigma_{\varepsilon,i} \rho_i)} E_{K,i}. \quad (3.8)$$

### 3.3.2 CASINO Execution

In the initial stage of our pipeline, modelling electron trajectories, we ran separate simulations for two different sample compositions: the solvent alone, and the crystal containing 36% solvent. In CASINO, each sample was defined by chemical composition, density, and thickness. The first two parameters are as described above, the third was set to be thick enough that no electrons passed through the far edge; this value varied with energy and was determined through a process of trial and error.

For each sample material, we ran simulations at 6 electron beam energies,  $E_{e^-}$ . In each case  $E_{e^-}$  was determined based on the x-ray beam energy,  $E_X \in \{9, 12, 15, 20, 25, 30\}$ , in keV, and the average  $K$ -shell energy of the sample, weighted by chemical composition and elemental cross-section for photoelectric absorption. The electron beam energies and sample thicknesses used in execution are listed in table 2 of the paper (p. 15). As the intent was to simulate a point source of PEs within the sample, we used the minimum electron beam radius of 1 nm. For each sample and beam configuration, 100,000 electrons were simulated.

### 3.3.3 Space-Resolved Dose Map Execution

After generating the arrays,  $P$ , representing the averaged energy deposition of a single photoionisation event, for both the crystal and solvent samples, we compared them and observed that they did not differ greatly. We averaged the results and used this average subsequently as this simplified the simulation, reducing execution time. We modelled the beam with a range of widths (FWHM): 1  $\mu\text{m}$ , 1.5  $\mu\text{m}$ , 2  $\mu\text{m}$  and 5  $\mu\text{m}$ . In each case, the  $x$  and  $y$  widths were the same. The total flux modelled was  $1 \times 10^{10}$  photons  $\text{s}^{-1}$  in each case.

### 3.3.4 Time-Resolved Spot Fading Execution

All the parameters required by the time-resolved module of our simulation are hard-coded or pulled directly from the outputs of the preceding module.

## 3.4 Paper

**Manuscript title:** 'The Influence of Photoelectron Escape in Radiation Damage Simulations of Protein Micro-Crystallography'

**Publication status:** Published June 2018, doi:10.3390/cryst8070267

**Authors:** Hugh Marman, Connie Darmanin, Brian Abbey

**Notes:** There is a typo on page 9, in the captioning of fig. 5. The sub-figures at the top of the page are labelled (b) and (c). They should be labelled (c) and (d).

In fig. 6, the colouring of the processes is incorrect. All the rectangles are pink. The rectangle labelled "RADDPOSE-3D" is correctly coloured, whereas the rectangle labelled "CASINO" should be red, and all the other rectangles should be orange. RADDPOSE-3D and CASINO are third-party applications. All the remaining rectangles are processes scripted in MATLAB as part of this thesis.

Article

# The Influence of Photoelectron Escape in Radiation Damage Simulations of Protein Micro-Crystallography

Hugh Marman, Connie Darmanin \* and Brian Abbey \*

ARC Centre of Excellence in Advanced Molecular Imaging, Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia; hmarman@students.latrobe.edu.au

\* Correspondence: c.darmanin@latrobe.edu.au (C.D.); b.abbey@latrobe.edu.au (B.A.); Tel.: +613-9479-5367 (C.D.); +613-9479-2645 (B.A.)

Received: 1 May 2018; Accepted: 24 June 2018; Published: 27 June 2018



**Abstract:** Radiation damage represents a fundamental limit in the determination of protein structures via macromolecular crystallography (MX) at third-generation synchrotron sources. Over the past decade, improvements in both source and detector technology have led to MX experiments being performed with smaller and smaller crystals (on the order of a few microns), often using microfocus beams. Under these conditions, photoelectrons (PEs), the primary agents of radiation-damage in MX, may escape the diffraction volume prior to depositing all of their energy. The impact of PE escape is more significant at higher beam energies (>20 keV) as the electron inelastic mean free path (IMFP) is longer, allowing the electrons to deposit their energy over a larger area, extending further from their point of origin. Software such as RADDOSE-3D has been used extensively to predict the dose (energy absorbed per unit mass) that a crystal will absorb under a given set of experimental parameters and is an important component in planning a successful MX experiment. At the time this study was undertaken, dose predictions made using RADDOSE-3D were spatially-resolved, but did not yet account for the propagation of PEs through the diffraction volume. Hence, in the case of microfocus crystallography, it is anticipated that deviations may occur between the predicted and actual dose absorbed due to the influence of PEs. To explore this effect, we conducted a series of simulations of the dose absorbed by micron-sized crystals during microfocus MX experiments. Our simulations spanned beam and crystal sizes ranging from 1  $\mu\text{m}$  to 5  $\mu\text{m}$  for beam energies between 9 keV and 30 keV. Our simulations were spatially and temporally resolved and accounted for the escape of PEs from the diffraction volume. The spatially-resolved dose maps produced by these simulations were used to predict the rate of intensity loss in a Bragg spot, a key metric for tracking global radiation damage. Our results were compared to predictions obtained using a recent version of RADDOSE-3D that did not account for PE escape; the predicted crystal lifetimes are shown to differ significantly for the smallest crystals and for high-energy beams, when PE escape is included in the simulations.

**Keywords:** macromolecular crystallography; radiation damage; photoelectron escape; Monte Carlo simulation; nanocrystals

## 1. Introduction

Technological advances over the last few decades have helped to mitigate, though not eliminate, radiation damage in protein crystallography. They have also facilitated the use of smaller, less perfect crystals in structure determination. Experiments can be performed using microfocus X-ray beams with sufficient flux density to cause rapid degradation of crystals. The use of such highly-focused beams is

set to become routine in the near future, as diffraction-limited storage rings, currently in construction or planning at several facilities worldwide, begin to come on-line [1,2]. The characteristics of small crystals, combined with these developments, create new challenges for understanding the effects of radiation damage. For an overview of the field, the reader is directed toward the recent reviews by Garman [3–6] and Holton [7].

Several metrics have been proposed to track the progression of radiation damage. These include loss of integrated intensity in the diffraction patterns, increase in scaling B-factors, expansion of the unit-cell, reduction of metal centres, and loss of resolution in susceptible residues [3]. These metrics are measured against the *dose* (in MGy) received by a crystal, being the energy absorbed per unit mass. Dose is not measured directly but is calculated based on the relevant beam and sample parameters.

The total diffracted intensity is one of the simplest metrics used for tracking X-ray induced damage to macromolecules [7]. The integrated intensity of the complete diffraction volume, a specific wedge of diffraction data, or the intensity at a single Bragg peak, have all been used to study radiation damage [3]. The integrated intensity of an individual Bragg spot can be modelled with the empirical formula

$$I(t) = I_0 \exp\left(-\ln(2) \frac{D(t)}{Hd}\right), \quad (1)$$

where  $I$  and  $D$  are the spot intensity and dose received at time  $t$ , respectively,  $I_0$  is the intensity at  $t = 0$ ,  $d$  is the resolution of the Bragg spot, and  $H = 10 \text{ MGy } \text{\AA}^{-1}$  is the *Howell's criterion* [8]. Howell's criterion was derived using spot fading to half the initial intensity as an indication of dose limit [7]. As higher resolution spots fade faster, Equation (1) therefore provides a resolution-dependent estimate of the dose that a crystal can absorb before the diffracted intensity falls to half its initial level. Holton has shown that Equation (1) shows good agreement with recent damage studies [7].

Dose prediction is, therefore, a critical factor when planning an MX experiment, especially if the amount of crystalline sample is limited and/or the available crystals are limited in size. RADDDOSE [9–12] is a software package that has been demonstrated to produce realistic dose estimates for MX experiments under typical conditions (e.g., crystal and beam sizes  $>5 \mu\text{m}$  and X-ray energies around 12 keV). Improvements in source and detector technology, however, mean that increasingly MX experiments are being carried out using micron-sized crystals and beams.

At the beam energies typically used in MX, X-ray–matter interactions are dominated by photoionisation, with incoherent scatter becoming appreciable at energies above 20 keV. During photoionisation, a photoelectron (PE) is ejected from a core shell (principally the K-shell since this has the largest cross-section) by an incident X-ray photon. The primary PE has a kinetic energy equal to the energy of the incident X-ray photon minus the core shell binding energy. The ejected PE undergoes deceleration via a series of inelastic collisions with the atomic electrons within the solid. The average distance travelled by the PE between these collisions is its inelastic mean free path (IMFP). Provided the energy of the absorbed X-ray photon is large enough and the size of the X-ray beam small enough, the PE may escape the diffracting volume of the sample prior to depositing all of its energy, leading to a reduction in radiation damage, relative to diffracted intensity. The beam size, rather than the crystal size, is the critical element here: if the beam is much larger than the crystal, approximately the same number of PEs will enter the crystal from the solvent as will escape it, offsetting the benefit due to PEs escape. Although the basic principles behind this damage reduction have been established the circumstances under which these effects may become important in MX experiments are still a topic of active research [13–18]. At the time this study was undertaken, the most widely used programmes for determining dose in MX experiments did not yet account for the possibility of PE escape, a fact that has only recently changed with the release of a new version of RADDDOSE-3D [19]. Neglecting the possibility of PE escape is expected to result in a lower-than-predicted dose rate to the crystal, provided the beam and crystal size are small enough and the X-ray energy high enough [20,21]. Establishing the conditions under which PE escape is expected to have a significant impact on dose rates for micron-sized crystals is a key aim of the current work.

The last decade has seen a growing body of literature exploring the potential benefits of PE escape for MX through both theoretical models and experiments. These studies have explored the length scales and energy ranges over which a reduction in radiation damage might be realised due to PE escape, providing some practical solutions for minimising radiation damage [20,21]. Suggestions generally rely on making the beam footprint smaller (in at least one dimension) than the damage footprint. This may be accomplished by using microfocus or line-focus X-ray beams [16–18]. Some special cases have also been investigated, where PE escape might be enhanced, these include imaging of platelet or needle-like crystals [22]. Previous studies [17,20,21] of the effect of PE escape have often incorporated Monte Carlo (MC) modelling of electron trajectories, performed using programmes such as CASINO [23,24]. In the current paper, we present the first systematic simulation study of PE escape in MX accounting for:

- (1) the effects of a finite beam size comparable to the PE stopping range;
- (2) the effects of a finite crystal size matched to the FWHM of the beam;
- (3) the effect of variation in X-ray beam energy; and
- (4) the influence of the beam profile as a function of the dose received by the crystal over time.

In contrast to previous work, we explicitly simulated the distribution of dose due to both PE escape and Compton scattering in *the whole beam footprint*, accounting for PEs produced in the crystal and the surrounding solvent. These spatially-resolved simulations of dose distribution were used to update the time-resolved damage state of the crystal. By comparing spatially-resolved dose maps produced with and without PE motion incorporated, we arrived at an estimate of the experimental conditions under which the effects of PE escape could lead to a substantial reduction in radiation damage.

## 2. Results

We simulated the dose absorbed by a lysozyme crystal that was fixed in place and did not rotate. We used 2D (cross-sectional) dose maps produced according to Section 4.5, and 3D dose maps produced using RADDPOSE-3D [9], as inputs for a spot-fading model (Section 4.7) that simulated the loss in integrated intensity of a 3 Å Bragg spot under a range of beam- and sample-conditions. In this section, we present:

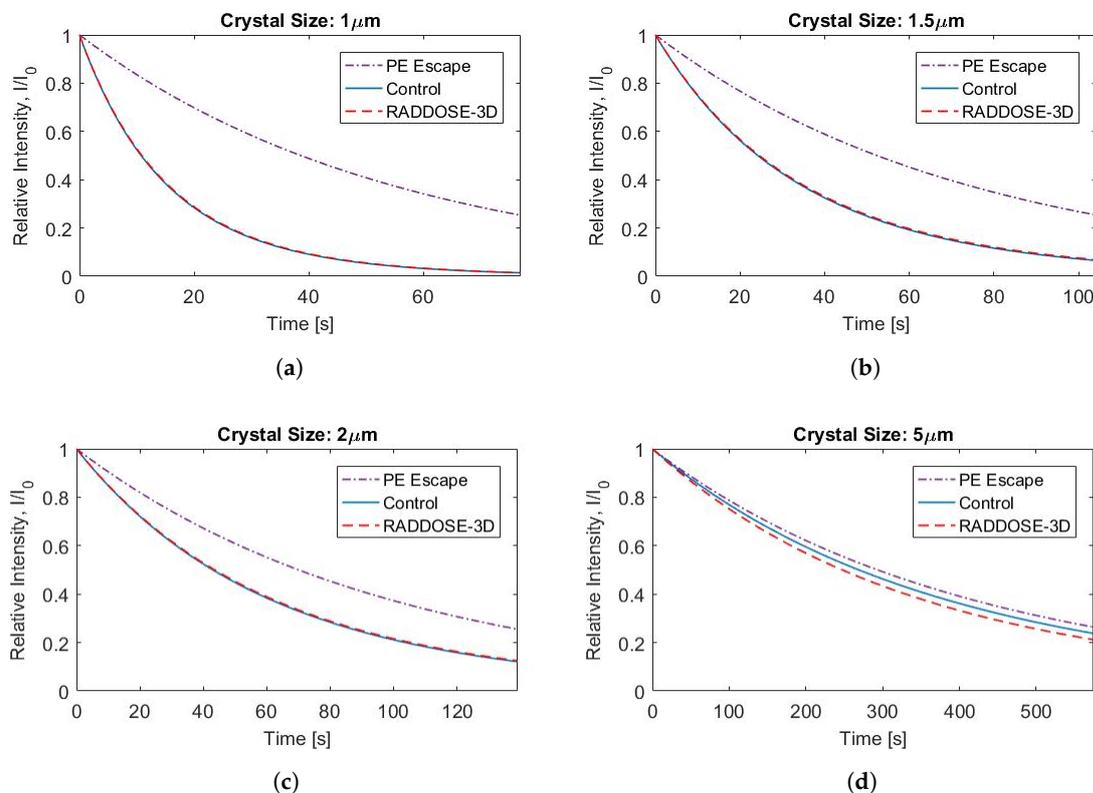
- the results of our spot-fading model, comparing the effects of variation in crystal size, beam size, and beam energy;
- calculations of the change in X-ray interaction cross-sections under our simulation parameters, and the implications this has for diffraction efficiency in micro-crystallography; and
- a representation of a 2D slice through the real-space diffraction volume, illustrating how the contribution to diffracted intensity within this volume evolves as the crystal absorbs a dose of up to 60 MGy.

We note that RADDPOSE-3D produces several aggregate dose metrics; when comparing global dose rates, we used the *average dose, whole crystal* (ADWC).

### 2.1. Crystal Size

Figure 1 shows plots of the ratio of integrated intensity to initial intensity,  $I/I_0$  (Equation (1)), as a function of time for cubic crystals with edge lengths of 1  $\mu\text{m}$ , 1.5  $\mu\text{m}$ , 2  $\mu\text{m}$  and 5  $\mu\text{m}$ . Simulations based on three sets of dose maps are compared. We produced one set of dose maps that accounted for PE escape and a second (control) set that did not; both included the effect of Compton scattering. The dose maps produced using RADDPOSE-3D did not include either PE escape or Compton scattering. In all cases, the beam energy was fixed at 12 keV with a beam flux of  $1 \times 10^{10}$  photons  $\text{s}^{-1}$ ; the beam FWHM was matched to the side length of the crystal. The influence of PE escape on the rate of spot-fading was found to be significant for the 1  $\mu\text{m}$  crystal, with simulations that account for PE escape predicting that the peak intensity falls to half its initial value in approximately 40 s, four times longer than predictions that do not account for PE escape. The discrepancy between simulations with and without PE escape

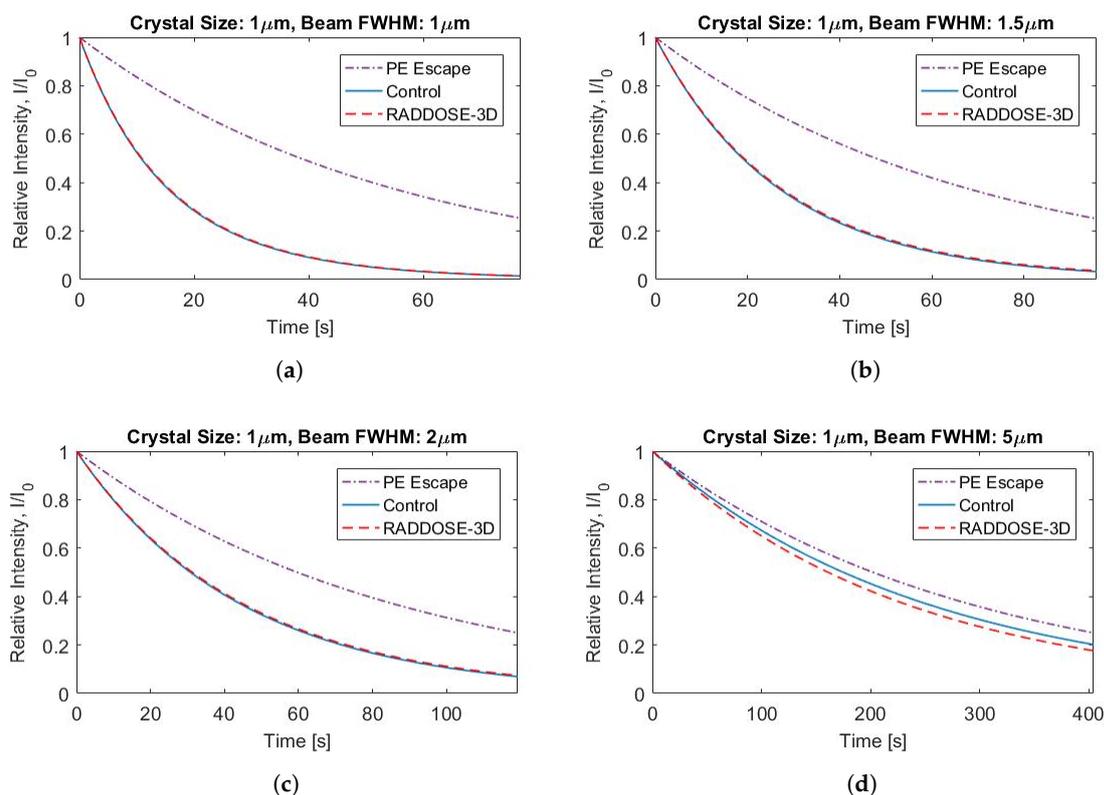
was found to decrease steadily as the crystal/beam size increased up to 5  $\mu\text{m}$ , beyond which no significant difference was seen between the two approaches. No significant difference was observed, for any crystal size, between the two simulations that did not account for PE escape.



**Figure 1.** Relative intensity,  $I/I_0$ , as a function of time, for a 3 Å Bragg spot for crystal size: (a) 1  $\mu\text{m}$ ; (b) 1.5  $\mu\text{m}$ ; (c) 2  $\mu\text{m}$ ; and (d) 5  $\mu\text{m}$ . Each crystal is assumed to be centred in a 12 keV beam with incident flux of  $1 \times 10^{10}$  photons  $\text{s}^{-1}$  and a beam FWHM matched to the crystal size. Three sets of dose maps were used in the simulations: Purple dot-dash lines and solid blue lines show results based on our simulated dose maps (see Section 4.5). The former accounted for PE escape, the latter (control) did not; both accounted for Compton scattering. Our model that accounted for PE escape also included the effect of PEs being scattered *into* the crystal from the solvent. Dashed red lines show results based on RADDOS-3D [9] dose maps; the version used did not account for either PE escape or Compton scattering. The discrepancy between the simulation that accounts for PE escape and those that do not becomes more significant as the crystal size decreases from 5  $\mu\text{m}$  to 1  $\mu\text{m}$ . No significant variation is observed between the two simulations that did not account for PE escape.

## 2.2. Beam Size

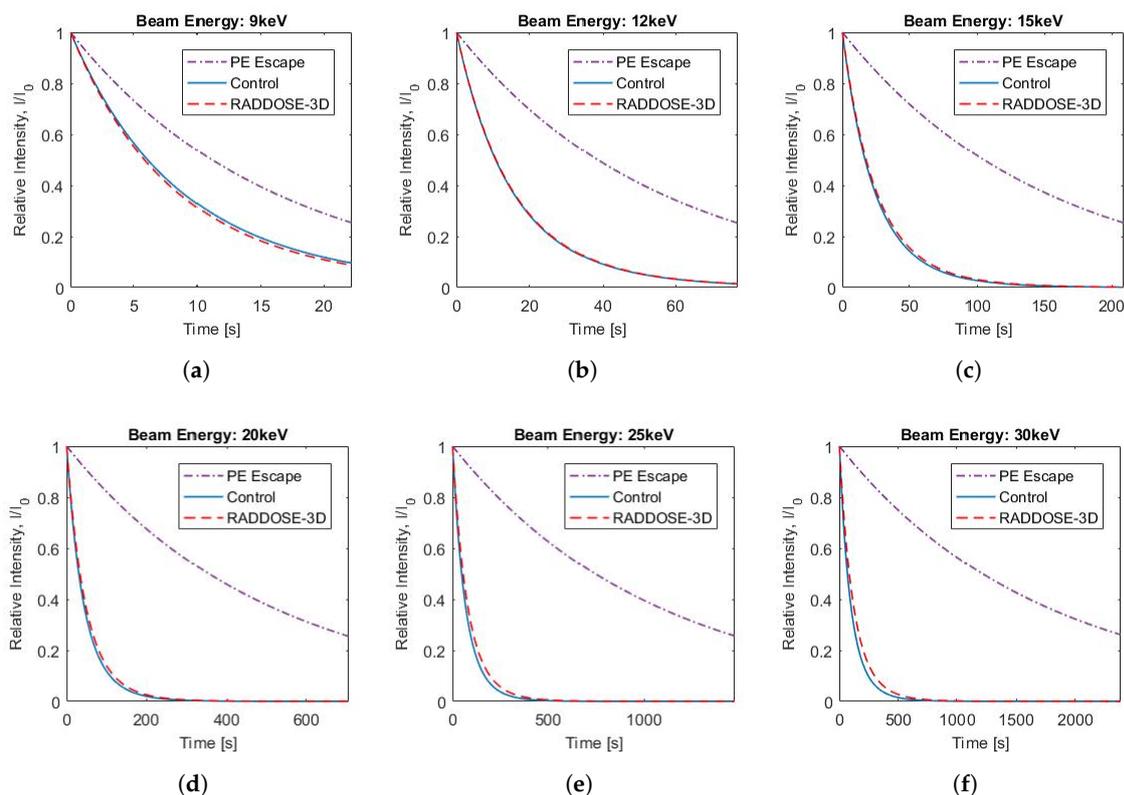
Figure 2 shows plots of the relative integrated intensity,  $I/I_0$ , as a function of time, with the crystal side-length fixed at 1  $\mu\text{m}$  and the beam FWHM varying from 1  $\mu\text{m}$  to 5  $\mu\text{m}$ . The beam energy and incident flux were fixed at 12 keV and  $1 \times 10^{10}$  photons  $\text{s}^{-1}$ , respectively. As expected, when the beam FWHM was matched to the 1  $\mu\text{m}$  crystal size, there was a significant discrepancy between the simulations that included PE escape and those that did not. This discrepancy reduced as the beam FWHM was increased up to 5  $\mu\text{m}$ , at which point it was no longer significant. No significant variation was seen between the two simulations that did not include PE escape, for any beam size.



**Figure 2.** Relative intensity,  $I/I_0$ , as a function of time, for a  $3 \text{ \AA}$  Bragg spot generated by a  $1 \mu\text{m}$  crystal centred in X-ray beams of varying width. All beams had energy of  $12 \text{ keV}$  and incident flux of  $1 \times 10^{10} \text{ photons s}^{-1}$ . The beam FWHM was: (a)  $1 \mu\text{m}$ ; (b)  $1.5 \mu\text{m}$ ; (c)  $2 \mu\text{m}$ ; and (d)  $5 \mu\text{m}$ . Three sets of dose maps were used in the simulations: Purple dot-dash lines and solid blue lines show results based on our simulated dose maps (see Section 4.5). The former accounted for PE escape, the latter (control) did not; both accounted for Compton scattering. Dashed red lines show results based on RADDOS-3D [9] dose maps; the version used did not account for either PE escape or Compton scattering. The discrepancy between the simulation that accounts for PE escape and those that do not becomes more significant as the beam size decreases from  $5 \mu\text{m}$  to  $1 \mu\text{m}$ . No significant variation is observed between the two simulations that did not account of PE escape.

### 2.3. Beam Energy

Figure 3 shows plots of the variation in  $I/I_0$  as a function of time for a range of different beam energies ( $9 \text{ keV}$ ,  $12 \text{ keV}$ ,  $15 \text{ keV}$ ,  $20 \text{ keV}$ ,  $25 \text{ keV}$  and  $30 \text{ keV}$ ). The beam FWHM was fixed at  $1 \mu\text{m}$ , matched to the  $1 \mu\text{m}$  crystal size. For all energies simulated, slower spot fading was predicted when PE escape was included in the model. The difference between the two different simulations was small at  $9 \text{ keV}$ , but increased steadily up to  $20 \text{ keV}$  for the simulated lysozyme crystal. Beyond  $20 \text{ keV}$ , the discrepancy continued to increase but at a much reduced rate. Between  $9 \text{ keV}$  and  $20 \text{ keV}$ , no significant variation was seen between the two simulations that did not include PE escape. A small difference between the two was observed at  $25 \text{ keV}$  and  $30 \text{ keV}$ , with our control (that did account for Compton scattering) predicting slightly faster spot-fading than RADDOS-3D (which did not account for Compton scattering).



**Figure 3.** Relative intensity,  $I/I_0$ , as a function of time, for a  $3 \text{ \AA}$  Bragg spot generated by a  $1 \text{ \mu m}$  crystal centred in a  $1 \text{ \mu m}$  beam with an incident flux of  $1 \times 10^{10} \text{ photons s}^{-1}$  and beam energy of: (a) 9 keV; (b) 12 keV; (c) 15 keV; (d) 20 keV; (e) 25 keV; and (f) 30 keV. Three sets of dose maps were used in the simulations: Purple dot-dash lines and solid blue lines show results based on our simulated dose maps (see Section 4.5). The former accounted for PE escape, the latter (control) did not; both accounted for Compton scattering. Dashed red lines show results based on RADDOSE-3D [9] dose maps; the version used did not account for either PE escape or Compton scattering. Incorporating PE escape resulted in lower damage rates being predicted at all energies, with increasing discrepancy seen between results as the energy increased from 9 keV to 20 keV, and minimal change in discrepancy observed above 20 keV. Inclusion of Compton scattering did not result in any significant discrepancy between the two simulations that ignored PE escape for energies up to 20 keV and resulted in only a small deviation at 25 keV and 30 keV.

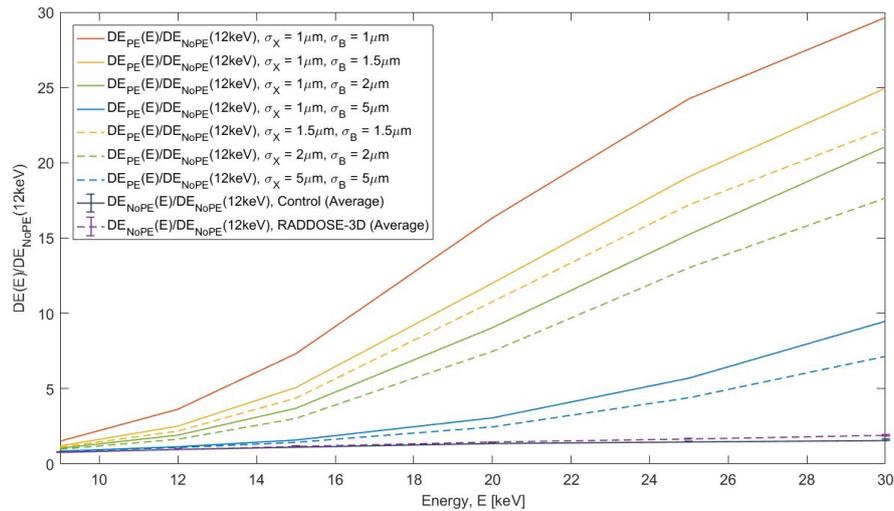
#### 2.4. Variation of Diffraction Efficiency with Energy

We estimated the diffraction efficiency,  $DE(E, \sigma_X, \sigma_B)$ , of our simulated lysozyme crystals (see Section 4.1 for sample details) as a function of crystal size,  $\sigma_X$ , beam size,  $\sigma_B$ , and beam energy,  $E$ , by two methods that are described in Section 4.8. Results are grouped into three categories, which are summarised in Table 1:

- (1)  $DE_{NIST}$  values were calculated using cross-section data for coherent scatter,  $\sigma_C$ , incoherent scatter,  $\sigma_I$ , and photoelectric absorption,  $\sigma_P$ , obtained from the NIST XCOM database [25]. These results did not account for PE escape and were independent of beam and crystal size.
- (2)  $DE_{NoPE}$  values were calculated as the ratio of the number of coherently scattered photons per unit volume,  $N_C/\sigma_X^3$ , to the average global dose,  $D_G$ , based on results of our simulations that did not account for PE escape—this included both our control simulations and simulations that used RADDOSE-3D dose maps.  $DE_{NoPE}$  was insensitive to changes in beam and crystal size; within each set (RADDOSE-3D, control) results at 30 keV showed a standard deviation of 3% from the average, with averages at all other energies showing a standard deviation of less than

1%. Results from our control and RADDOSÉ were within 1–2% of each other at energies up to 15 keV but diverged at higher energies, differing by 4% to 10% between 20 keV and 30 keV. A full table of  $DE_{NoPE}$  values is shown in Appendix A; average values were used in Table 1 and average values with error bars showing the standard deviation were used in Figure 4.

- (3)  $DE_{PE}$  values were calculated the same way as  $DE_{NoPE}$  values but using the results of the simulations that did include PE escape. They showed significant dependence on beam size and some dependence on crystal size for a given beam size.  $DE_{PE}$  values for a 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  beam are shown in Table 1. A plot of all  $DE_{PE}$  values is shown in Figure 4.



**Figure 4.** Diffraction efficiency,  $DE(E)$ , of simulated results with PE escape ( $DE_{PE}$ ) and without PE escape ( $DE_{NoPE}$ ) incorporated. Red, orange, green, and blue lines (both solid and dashed) are for  $DE_{PE}$ , crystals and beams between 1  $\mu\text{m}$  and 5  $\mu\text{m}$ , respectively, as shown in the legend. The solid dark blue line is the average of  $DE_{NoPE}$ , as determined by our control simulations, across all crystal and beam sizes, with error bars showing standard deviation. The dashed purple line shows the average for RADDOSÉ-3D results. As the standard deviation within each set is  $\leq 3\%$  for all energies, the error bars are only resolvable on some data points. The reference value was the value of  $DE_{NoPE}(12\text{ keV})$  calculated with RADDOSÉ-3D; all values of  $DE_{PE}(E)$  and  $DE_{NoPE}(E)$  are relative to this reference.

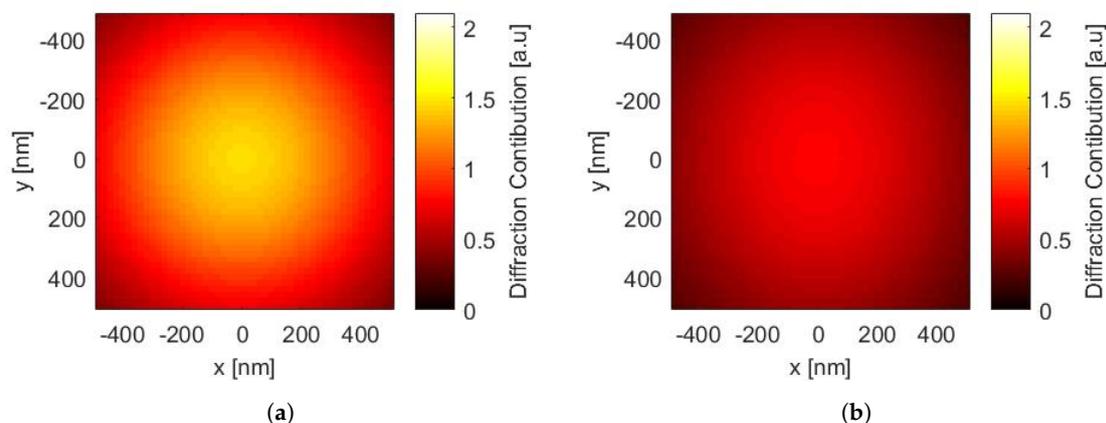
As shown in Table 1, values of  $DE_{NoPE}$  (based on both RADDOSÉ-3D and our controls) differed from those of  $DE_{NIST}$ , for a given energy, by between 15% and 35%. Both  $DE_{NIST}$  and  $DE_{NoPE}$  were found to be essentially independent of beam and crystal size, and showed an approximate doubling between 12 keV and 30 keV, although this factor was lower for our control simulations. As shown in Figure 4, values of  $DE_{PE}$  were only similar to those of  $DE_{NoPE}$  for the 1  $\mu\text{m}$  and 5  $\mu\text{m}$  crystals in the 5  $\mu\text{m}$  beam at energies  $\leq 15$  keV. Across all the simulated results that included PE escape, the ratio  $DE_{PE}(E, \sigma_X, \sigma_B) / DE_{NoPE}(12\text{ keV})$  varied between 1.09 ( $\sigma_X = \sigma_B = 5\text{ }\mu\text{m}$ ) and 3.63 ( $\sigma_X = \sigma_B = 1\text{ }\mu\text{m}$ ) at  $E = 12\text{ keV}$ , with the greatest difference occurring at 30 keV (7.15 ( $\sigma_X = \sigma_B = 5\text{ }\mu\text{m}$ ) and 29.64 ( $\sigma_X = \sigma_B = 1\text{ }\mu\text{m}$ )).

**Table 1.** Cross-section data and diffraction efficiency for a lysozyme crystal with 36% solvent, as described in Section 4.1, relative to the reference values  $DE_{NIST}(12\text{ keV})$  and the RADDOS-3D [9] value for  $DE_{NoPE}(12\text{ keV})$ . Cross-sections for coherent scatter,  $\sigma_C$ , incoherent scatter,  $\sigma_I$ , and photoelectric absorption,  $\sigma_P$ , were calculated using the NIST XCOM database [25]. Diffraction efficiency was calculated as the ratio  $\sigma_C/(\sigma_I + \sigma_P)$  from cross-section data and the ratio of coherently scattered photons per unit volume,  $N_C/\sigma_X^3$ , to average global dose,  $D_G$ , for simulations.  $DE_{NoPE}$  values are averages across all beam and crystal sizes simulated.  $DE_{PE}$  values are for a 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  beam.

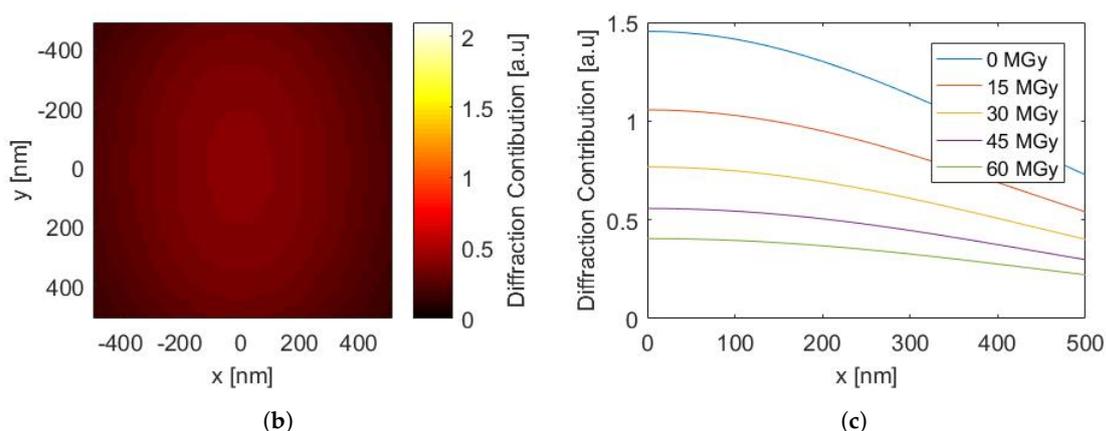
Energy, $E$ , [keV]:	9	12	15	20	25	30
$\sigma_C$ :	1.44	1.00	0.74	0.49	0.35	0.26
$\sigma_I + \sigma_P$ :	2.35	1.00	0.53	0.25	0.15	0.11
$DE_{NIST}(E)/DE_{NIST}(12\text{ keV})$ :	0.61	1.00	1.41	1.98	2.29	2.32
$DE_{NoPE}(E)/DE_{NoPE}(12\text{ keV})$ , Control :	0.81	0.99	1.15	1.37	1.51	1.59
$DE_{NoPE}(E)/DE_{NoPE}(12\text{ keV})$ , RADDOS-3D :	0.79	1.00	1.19	1.47	1.69	1.91
$DE_{PE}(E)/DE_{NoPE}(12\text{ keV})$ :	1.51	3.63	7.32	16.33	24.25	29.64

### 2.5. Evolution of Diffraction Volume as a Function of Dose

Figure 5 shows a series of 2D slices through the diffraction volume of a 1  $\mu\text{m}$  crystal centred in a 12 keV, 1  $\mu\text{m}$  X-ray beam for global doses ranging from 0 MGy to 60 MGy. We observe that at lower doses the shape of the diffraction volume is dominated by the Gaussian profile of the beam. As the crystal continues to absorb radiation, the diffraction volume changes in several ways: the overall volume shrinks; the strength of the diffracted signal (both peak and average) drops; and the shape of the volume becomes broader, an effect that is stronger along the axis perpendicular to the polarisation direction of the X-ray beam.



**Figure 5.** Cont.



**Figure 5.** Heat map showing diffraction volume, as a function of global dose, for a 1  $\mu\text{m}$  crystal centred in a 12 keV, 1  $\mu\text{m}$  beam (FWHM). The heat maps show the contribution to diffracted signal of a 2D slice perpendicular to the beam direction, when the crystal has absorbed a global dose of: (a) 0 MGy; (b) 30 MGy; and (c) 60 MGy. Panel (d) summarises the drop in diffraction volume along the  $x$ -axis (parallel to the polarisation axis of the X-ray beam), including additional calculations after the crystal has absorbed doses of 15 MGy and 45 MGy.

### 3. Discussion

A multi-stage approach was used to explicitly model the average distribution of energy in a crystal (and the surrounding solvent) due to both photoelectric absorption and Compton scattering. Simulations that accounted for PE escape produced predictions of damage progression in MX that agreed with results of our control simulation and RADDOSE-3D (both of which neglected PE escape) when the crystal and beam sizes were greater than 5  $\mu\text{m}$  and the beam energy was less than 12 keV, but diverged when both the crystal and beam size were reduced below 5  $\mu\text{m}$  and/or the beam energy was increased. As discussed below, this divergence in results is consistent with existing models of damage-reduction via PE escape. Moreover, our approach permits a qualitative examination of the progression of crystal deterioration in real-space, a research topic that is the subject of intense investigation, both theoretically [26] and experimentally [14,15].

For simplicity, we did not include rotational effects in our simulations. For many of the configurations we simulated, rotation would result in a more even spread of dose across the crystal volume (particularly at the edges) but the key effect being explored (the impact of PE escape on damage reduction) would not be significantly affected by rotation of the crystal.

#### 3.1. Prediction of Extended Lifetimes for Small Crystals and Beams, and Higher Energies

Our findings suggest that in some cases the lack of inclusion of PE escape leads to deviations from the behaviour predicted by RADDOSE-3D [9] (see, for example, [14,16,17]). This conclusion is supported by both the similarity in output to RADDOSE-3D under conditions where PE escape is unlikely to affect dose (or when it has been ignored in our models), and the divergence in output under conditions where PE escape is expected to be significant.

A comparison of Figures 1 and 2 shows that results that account for PE escape and those that do not exhibit a similar convergence as the beam width is increased, whether or not the crystal size is also increased so that it matches the beam width. The reason for this is that PEs are produced everywhere within the beam footprint, in both crystal and solvent. If the crystal is large, the majority of the PEs produced within it will not escape it. By contrast, if the crystal is small, many PEs will escape but if the beam footprint includes a significant portion of the surrounding solvent, a roughly equal number of PEs will be scattered into the crystal, offsetting the effect of those that escape it. For the sample that

we modelled, the chemical compositions and densities of the crystal and surrounding solvent are not significantly different in terms of the production and propagation of PEs. This is a consequence of the fact that the cross-sections and average binding energies between solvent and crystals are quite similar, leading to only a small difference in energy-deposition profile. Our choice of sample, lysozyme, contains no atoms heavier than sulphur and the only heavy atoms in the solvent we modelled are sodium and chlorine. The presence of more heavy atoms in either the crystal (e.g., metalloproteins) or the solvent (e.g., cacodylate) could lead to a more significant difference between the crystal and the solvent in the production of PEs.

The most striking discrepancy between results that included PE escape and those that did not occurred when simulating a 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  beam (FWHM) at high energies ( $\geq 15$  keV) (Figure 3). Here, our simulations predict increased diffraction efficiency compared to RADDPOSE-3D by a factor of close to 170% at 15 keV, 300% at 20 keV, 380% at 25 keV, and 400% at 30 keV. This discrepancy occurs because the IMFP of the PEs increases with energy and the resulting energy deposition is therefore spread over a larger area. The implication is that for micron-sized crystals illuminated by similarly sized beams, simulations that neglect PE escape over estimate the dose received, becoming increasingly less accurate as the incident X-ray photon energy increases. It has been previously observed that conducting MX experiments with smaller crystals at higher energies may allow experimenters to maximise the reduced radiation damage afforded by PE escape [17,20,21]. However, a systematic experimental study, isolating the effect of beam-energy in radiation damage has not, to our knowledge, been carried out and will be the focus of further work.

### 3.2. Simulations Enabling Qualitative and Semi-Quantitative Interpretation of Published Data

In early simulation work carried out by Nave and Hill [21], they used simple CASINO models to estimate that an 8  $\mu\text{m}$  crystal illuminated with 30 keV X-rays would benefit from a three-fold increase in diffraction efficiency due to PE escape. Although not included in their model, they noted the need to minimise the material surrounding the crystal to prevent PEs being scattered into it from the solvent, and the possibility of leveraging the anisotropic emission of PEs due to the beam polarisation to minimise damage. The largest crystal we simulated was 5  $\mu\text{m}$ , which saw an approximate four-fold improvement in diffraction efficiency at 30 keV when PE escape was accounted for; which is consistent with Nave and Hill's result. By including the solvent in our models, we have been able to gauge the effect of illuminating the volume around the crystal and, as discussed in the previous section, our results support the conclusion that PE production outside the crystal must be minimised if appreciable damage-reduction is to be achieved. Although Nave and Hill assumed larger beams and suggested physically limiting the volume of the solution surrounding the crystal, the increasing availability of microfocus X-ray beams may permit the same ends to be achieved through focusing, provided the crystal can be positioned within the beam with suitable precision to limit PE production outside the crystal.

Following, and extending upon, the work of Nave and Hill, Cowan and Nave [20] used coarse Monte Carlo models to study the impact of PE escape on diffraction efficiency. When neglecting PE escape, but accounting for Compton scattering, they observed a variation in diffraction efficiency of up to 40% between 1 keV and 40 keV, with a plateau of maximum efficiency between 20 keV and 30 keV. Under the same conditions, our model shows a similar trend, though the calculated values vary over a slightly wider range. Incorporating a model for PE escape, Cowan and Nave observed an improvement in diffraction efficiency at higher energies over a wide range of crystal sizes, though the effect was stronger for smaller crystals. In their model, a 1  $\mu\text{m}$  crystal saw a roughly five-fold increase in diffraction efficiency at 12 keV when accounting for PE escape, with the improvement above 20 keV being around 15-fold. This is consistent with our results (Figure 4) for the 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  beam, where the effect of PEs being generated in the solvent (which has been ignored in their model) is minimal. As in the model that excluded PE escape, Cowan and Nave's results show a broad plateau in diffraction efficiency for the 1  $\mu\text{m}$  crystal between 20 keV and 30 keV, before it drops off slightly at

higher energies. Our models also show a similar trend, though the diffraction efficiency (across all beam and crystal sizes) continues to improve as the energy is increased up to 30 keV, albeit at a reduced rate. This suggests that the increased cross-section for Compton scattering at higher energies offsets the reduction in dose due to PE escape. However, as shown in Figure 3, the discrepancy between our control simulation (that did include Compton scattering) and the RADDOSE-3D data (that did not include Compton scattering) is small, even at higher energies, and the discrepancy between them and the simulations that did include PE escape grows much more slowly above 20 keV than it did between 9 keV and 20 keV. Since this study was undertaken, RADDOSE-3D has been updated to include incoherent scatter and Bury et al. [19] observed that this inclusion only begins to have a significant effect on dose prediction above 40 keV. These results suggest that the increase in Compton scattering alone does not account for the plateau in diffraction efficiency at higher energies. A complementary interpretation is that much of the PE energy is already escaping the crystal at 20 keV and the increase in electron IMFP above this energy simply serves to scatter the PEs further into the solvent. There is some support for this idea in Figure 4, as the gradient of the diffraction efficiency of the 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  beam (the configuration compared in Figure 3) drops off slightly at higher energies compared to the gradients for the larger crystals. Cowan and Nave [20], who simulated crystals up to 20  $\mu\text{m}$ , observed a similar effect, with larger crystals continuing to benefit from increased diffraction efficiency as the beam energy was increased above the 20 keV to 30 keV where it plateaued for the 1  $\mu\text{m}$  crystal.

In Holton and Frankel's [27] paper on the minimum crystal size required to collect a complete data set, they estimated the effect of PE escape by first calculating the dose in the absence of PE escape and then scaling this dose by a so-called *Nave-Hill capture factor*,  $f_{NH}$ , which they approximated as

$$f_{NH} = \frac{D_{PE}}{D_{NoPE}} \simeq \frac{R}{R + R_{PE}(E)} \quad (2)$$

where  $R$  was the radius of the spherical crystals they modelled,  $R_{PE}$  was the range of a PE with energy  $E$ , and  $D_{PE}$  and  $D_{NoPE}$  were the dose values with and without accounting for PE escape, respectively. This capture factor was designed to equate to unity for "large" crystals at 12.4 keV and was intended to provide only a coarse indication of the benefit of PE escape. We used global dose values from our simulations to calculate  $f_{NH}$  as the ratio  $D_{PE}/D_{NoPE}$ . We also calculated  $f_{NH}$  using the RHS of Equation (2), taking  $R_{PE}$  as the maximum penetration depth of electrons in our CASINO simulations and substituting half the crystal edge-length for  $R$ . The two methods of calculation produced values for  $f_{NH}$  within a factor of 2 of each other. Our simulations did not have the spherical-symmetry of those performed by Holton and Frankel, but, given the coarse nature of their estimate, this substitution is expected to provide a reasonable point of comparison. The similarity in values of  $f_{NH}$  indicate that our simulation results are comparable to the Holton and Frankel model.

In addition to qualitative comparisons to previous modelling studies, a semi-quantitative assessment of our model can be conducted by comparison to recent results obtained by Finrock et al. [16,18] and Sanishvili et al. [17], who performed experimental studies of radiation damage in large crystals by exposing sub-volumes using micro-focus and line-focus beams. They observed that the radiation damage footprint extended between 1.5  $\mu\text{m}$  and 5  $\mu\text{m}$  beyond the actual beam footprint, showing the spread of PEs beyond the diffraction volume. Both experimenters used beam energies close to 18 keV. In the Monte Carlo stage of our models (implemented using CASINO), energy-deposition extended approximately 5  $\mu\text{m}$  beyond the interaction point at 20 keV, which is in-line with these previous observations.

Since this study was conducted, a new version of RADDOSE-3D [19] has been released that includes Compton scattering and PE escape, neither of which was included in the version we used. As noted above, Bury et al. [19] found that the inclusion of Compton scattering made negligible difference below 20 keV, and only minimal difference between 20 keV and 40 keV. Conversely, Bury et al. found the inclusion of PE escape to have a very significant impact on dose predictions for microcrystals. They simulated crystal sizes from 1  $\mu\text{m}$  to 100  $\mu\text{m}$  for a single beam energy of 12.4 keV

and observed a non-negligible reduction in dose for crystals as large as 20  $\mu\text{m}$  under these conditions. For a 10  $\mu\text{m}$  crystal, they observed a reduction in dose of approximately 20% when accounting for PE escape, and reduction of over 95% for a 1  $\mu\text{m}$  crystal. In our simulations, we observed a reduction of 72% for the 1  $\mu\text{m}$  crystal in a 12 keV beam, when the beam FWHM was matched to the crystal (and a much lower reduction when the beam was larger). This discrepancy can be accounted for by the fact that we simulated the solvent around the crystal, thereby accounting for PEs scattered into the crystal as well as out of it, while the new version of RADDOSE-3D assumes an isolated crystal, neglecting the effect of the surrounding solvent. In the future, a comprehensive comparison of our simulations and this new version of RADDOSE-3D could provide useful insights.

### 3.3. Permitting a Real Space Interpretation of Crystal Damage

The idea that diffracted intensity is approximately proportional to crystal volume is well-established [7] and the implication is that damage leading to a reduced intensity of diffracted radiation can be interpreted in real-space as an apparent “shrinking” of the diffraction volume of the crystal. Note, this does not mean that any actual protein material is being lost from the volume of the initial crystal, but that it becomes so disordered that it no longer contributes meaningfully to the diffraction signal. Recently, Coughlan et al. [14,15] reported such shrinking of the diffraction volume when using Bragg coherent diffractive imaging (BCDI) to produce 2D and 3D real-space images of lysozyme crystals that had been subjected to high dose burns (several hundred MGy).

Warkentin et al. [26] also recently modelled the change in the intensity of diffracted radiation per unit area for a crystal in a Gaussian beam, as it undergoes radiation damage. They observed that, due to the flux being highest in the beam centre, the portion of the crystal nearest the beam centre produced the largest number of diffracted photons, but also underwent radiation damage the fastest. Crucially, however, their model did not account for PE escape. To estimate which regions of the crystal contributed the most diffracted photons at the detector, they weighted each contribution by both the flux density in that part of the beam and the degree of damage already suffered. They observed a broadening of this contribution over time; initially, the central portion of the crystal contributed most strongly to the diffracted signal but by the time the spot intensity had dropped to half its initial value, their models predicted that the contribution to diffracted intensity was more or less equal across the whole FWHM of the beam. After sufficient exposure, radiation damage progressed to the point that only the fringes of the crystal (those in the lower flux edge of the Gaussian beam) contributed coherent scattered photons.

The snapshots of the diffraction volume in Figure 5 show a reduction in the total contribution to diffracted signal and shrinking and flattening of the diffraction contribution—meaning photons contributing to Bragg peaks are coming from different regions of the crystal. The flattening and shrinking of the diffraction volume is anisotropic as PEs are preferentially scattered at lower angles to the polarisation axis of the X-ray beam (here the  $x$ -axis), resulting in slower damage rates along the axis perpendicular to polarisation. As highlighted in the summary graph in Figure 5, our simulations show the same trend as Warkentin’s models [26], with an initial peak in the centre of the crystal and a flattening of the diffraction contribution across the whole crystal with increased dose. The overall shrinking of the diffraction volume is also in-line with the observations of Coughlan et al. [14,15] and may provide an initial model for future experimental examination of this phenomenon.

### 3.4. Micro-Crystallography Could Benefit Significantly from Higher X-Ray Energies

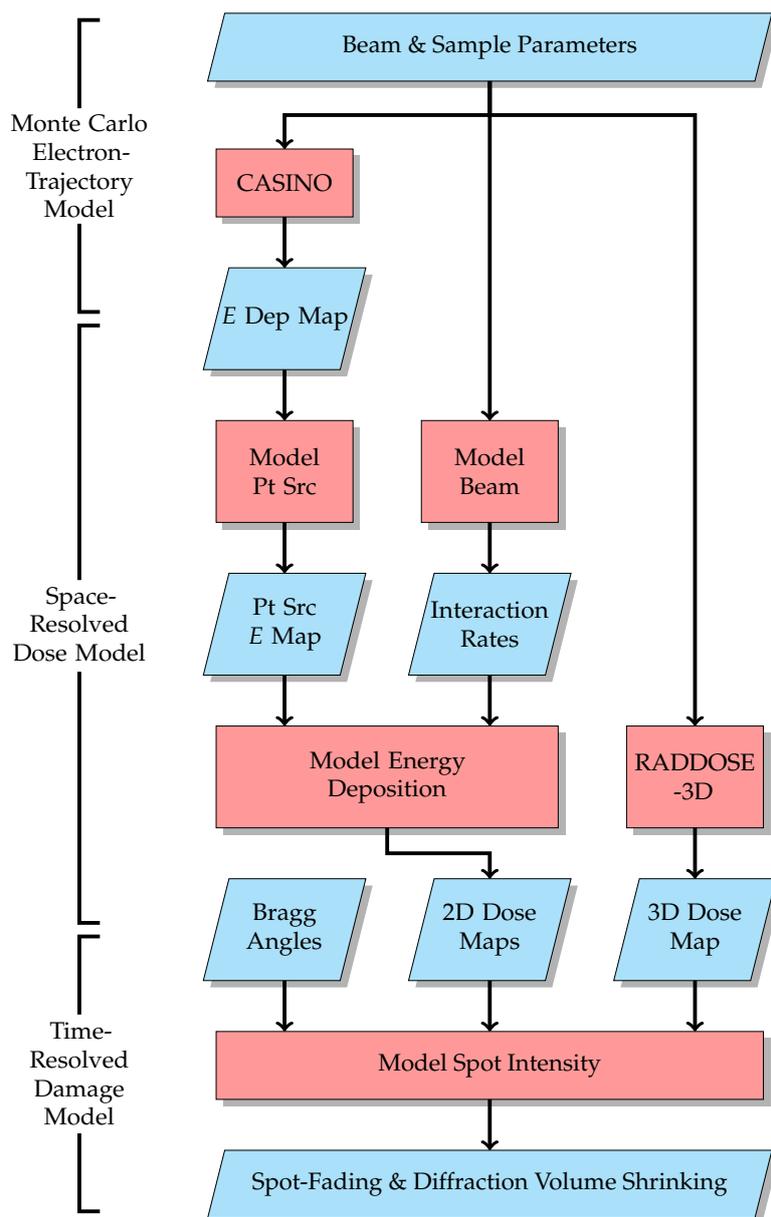
Predictions produced by our model suggest significant benefits could be gained by optimising MX experiments along multiple parameter-axes, particularly when using smaller crystals and higher beam energies. The increase in electron IMFP at higher energies, coupled with a decrease in damage-causing interactions relative to elastic scatter, suggests that, all other things being equal, the greatest increase in diffraction efficiency could be gained from using higher energy beams. As highlighted in Table 1, although the coherent scattering cross-section falls with increasing beam energy, it does so more slowly

than the combined cross-sections for incoherent scattering and photoelectric absorption. Although experiments at higher energy take longer (roughly twice as long at 20 keV and four times as long at 30 keV) due to the reduced cross-section for coherent scattering, the ratio of diffracted intensity to dose-accumulation (damage) would be roughly doubled for experiments above 20 keV, relative to those at 12 keV. This is even before accounting for the effect of PE escape; as shown in Figure 4, our simulations predict that PE escape could be responsible for increasing diffraction efficiency by up to 30 times, relative to the diffraction efficiency at 12 keV in cases where PE escape is not considered. This is assuming the use of micro-crystals and micro-focus beams so that the effect of PE escape is large, but even for a 5  $\mu\text{m}$  crystal/beam, our calculations show PE escape contributing to an increase in diffraction efficiency. Given the relative rate at which the cross-sections change and the longer IMFP for higher-energy PEs, conducting experiments at higher energies may be advantageous even with larger crystals in more conventional MX experiments.

Several experiments [16–18] have shown reduced damage occurring due to PE escape, but have only explored a narrow parameter space thus far; a more thorough examination, particularly focussed on the benefit of using higher energy beams, could be guided by our results and data obtained could subsequently be used to validate and improve our model. Should experiments show quantitative agreement with our predictions, our model could be used to provide more accurate predictions of dose rate, and therefore crystal lifetime in micro-crystallography experiments, compared to currently available radiation damage software packages.

#### 4. Materials and Methods

We simulated energy deposition within a lysozyme crystal and surrounding solvent (1:1 mixture of polyethylene glycol 400 (PEG400) and 0.5M acetic acid with 9% NaCl (*w/v*)). Crystal size was varied between 1  $\mu\text{m}$  and 5  $\mu\text{m}$ . The incident beam was simulated with a fixed flux of  $1 \times 10^{10}$  photons  $\text{s}^{-1}$  at energies of 9 keV, 12 keV, 15 keV, 20 keV, 25 keV and 30 keV, with FWHM from 1  $\mu\text{m}$  to 5  $\mu\text{m}$ . These simulations produced 2D dose maps, showing slices through the diffraction volume perpendicular to the beam direction. By incorporating Monte Carlo modelling of electron trajectories (using CASINO) we produced dose maps that accounted for the motion of PEs and the spreading of their energy over an area surrounding their point of origin. These dose maps accounted for the generation of PEs everywhere in the beam footprint, including the solvent surrounding the crystal. We produced a control set of dose maps that did not simulate PE motion but included Compton scattering. We used a version of RADDOSE-3D [9] that did not include PE escape or Compton scattering to produce 3D dose maps with the same range of sample- and beam-parameters. All three sets of dose maps produced assumed that the crystal was not rotated and was centred in the beam. We subsequently used both 2D and 3D dose maps as inputs for time-resolved simulations that modelled the rate at which radiation damage accrued and the effect of this damage on the diffraction volume. We simulated the drop in integrated intensity of a 3 Å Bragg spot, based on Equation (1). Variation in the diffraction volume of the crystal was modelled based on its spatially-resolved damage state and intensity of coherent scatter. A flow diagram of the simulations is shown in Figure 6; a sectioned explanation of the stages and input/output is given below; subheadings correspond to the labels of individual processes in the flow diagram.



**Figure 6.** Flow diagram of our simulations, which were performed using a combination of 3rd-party applications, CASINO (red) and RADDOSE-3D [9] (pink), and our own MATLAB scripts (orange). Rectangles represent processes, parallelograms are inputs/outputs. CASINO output (*E Dep Map*) gives (ensemble-averaged) energy-deposition of an electron with a particular initial direction; post-processing of this data in MATLAB produces a dose map for a point source of photoelectrons with emission-direction weighted for a horizontally polarised X-ray beam. MATLAB scripts are further used to determine the dose footprint of the X-ray beam, accounting for spatially-distributed energy deposition by photoelectrons. RADDOSE-3D [9] is used to produce a dose map of the crystal that does not account for photoelectron escape. Our final MATLAB script models the loss in integrated intensity of a 3 Å Bragg spot with and without accounting for photoelectron escape.

#### 4.1. Beam and Sample Parameters

Our simulated sample was crystallised lysozyme, suspended in a 1:1 mixture of PEG400 and 0.5 M acetic acid with 9% NaCl (*w/v*), following crystallisation and sample preparation protocols by Coughlan [15]. The solvent content of the crystal was calculated to be 36%. The sample parameters used in the simulation were chemical composition, density, and average ionisation potential (weighted by

elemental cross-section and chemical composition). Crystals were modelled as cubes, with edge lengths of 1  $\mu\text{m}$ , 1.5  $\mu\text{m}$ , 2  $\mu\text{m}$  and 5  $\mu\text{m}$ .

Different X-ray beam widths and energies were modelled; in each case, the total beam flux was  $1 \times 10^{10}$  photons  $\text{s}^{-1}$ . X-ray beam widths modelled were 1  $\mu\text{m}$ , 1.5  $\mu\text{m}$ , 2  $\mu\text{m}$  and 5  $\mu\text{m}$ . Energies modelled were 9 keV, 12 keV, 15 keV, 20 keV, 25 keV and 30 keV. Cross-sections for photoionisation, Compton scattering, and elastic scattering were calculated using NIST XCOM online cross-section database [25].

#### 4.2. CASINO

Monte Carlo simulations of electron trajectories were performed with CASINO (version 2.4.8.1) [23,24]. We simulated two different materials: the crystal, permeated by solvent, and the solvent alone. Both are described above in Section 4.1. For each material, we ran the simulation at six different (electron) energies. The energy in this case was the calculated average photoelectron energy, determined by the X-ray beam energy (9 keV, 12 keV, 15 keV, 20 keV, 25 keV and 30 keV) and the average photoionisation energy of the sample. CASINO is configured using up to five menus under the “Settings” tab; the settings we used are detailed below, under headings that correspond to the menus. Any parameters that are not explicitly stated below were left at their defaults. We made no adjustments from default values to either the “Runtime Options” or “Change Physical Models” menus.

##### 4.2.1. Modify Sample

A single layer was used for each material. The thickness was varied depending on the beam energy used, as detailed in Table 2. This was done to ensure that all electron paths were entirely within the material (this was determined by trial and error in early test runs). The chemical composition of the solvent was  $\text{C}_{32}\text{H}_{134}\text{O}_{52}\text{NaCl}$ , with a density of  $1.184 \text{ g cm}^{-3}$ ; the chemical composition of the crystal was  $\text{C}_{766}\text{H}_{1603}\text{N}_{193}\text{O}_{435}\text{S}_{10}\text{NaCl}$ , with a density of  $1.259 \text{ g cm}^{-3}$ .

##### 4.2.2. Set Up Microscope

CASINO permits simulation of a range of energies in a single execution of the programme, provided the step size is constant. As we had a non-constant step size, we simulated a single energy per execution (see Table 2). For each execution, we simulated 100,000 electrons with a beam radius of 1 nm. We intended the electron beam to represent a point-source origin of PEs in the material, and with a maximum resolution of 20 nm in the later steps of our simulation, this was considered a suitably small beam to approximate a point.

**Table 2.** CASINO input parameters. The thickness of the sample was adjusted to ensure electron pathways did not exit the sample. The electron energy was slightly different for the solvent and the crystal as differences in chemical composition resulted in a slightly different average ionisation energy.

X-ray Energy, $E_X$ , [keV]:	9	12	15	20	25	30
Photoelectron Energy, $E_{PE}$ , (Solvent) [keV]:	7.89	10.85	13.83	18.80	23.77	28.75
Photoelectron Energy, $E_{PE}$ , (Crystal) [keV]:	7.81	10.77	13.73	18.69	23.67	28.64
Sample Thickness [nm]:	2000	4000	5000	10,000	15,000	20,000

##### 4.2.3. Select Distributions

We enabled “Distribution of Maximum depth of electrons” with 1000 divisions, and “Division of energy by position” with 200 divisions in each axis ( $x$ ,  $y$ , and  $z$ ). All other distributions were disabled as their output was not relevant to the remainder of our simulation.

### 4.3. Model Pt Src

Energy deposition by position data (*E Dep Map* in Figure 6) were exported from CASINO and subsequently processed using MATLAB. The Monte Carlo simulations of electron trajectory were interpreted in the following way: the simulated trajectories represented the statistical distribution of paths followed by electrons produced *at a particular point in space* (in CASINO, the meeting of the electron beam with the sample) and *moving initially in a particular direction* (in CASINO, straight down, along the z-axis). The data output by CASINO therefore required post-processing to account for the fact that photoelectrons may be produced at any point within the beam footprint (with the likelihood of emission at a given point being a function of the beam intensity at that point) and that any photoelectrons produced at a given point would have a probability of being emitted in a given direction that was dependent upon the polarisation of the X-ray beam. That is, for a linearly polarised beam, the probability of emission in a given direction would go as  $\cos^2(\theta)$ , where  $\theta$  is the angle made to the polarisation vector. The former requirement is addressed below in Sections 4.4 and 4.5. The MATLAB script used to model the latter of these requirements proceeded in the following basic sequence:

- (1) Read CASINO energy-by-position data into an array, centred on the emission location of the PE.
- (2) In steps of  $1^\circ$ , rotate the array, scale it by  $\cos^2(\theta)$ , where  $\theta$  was the angle made to the x-axis (as we were simulating a horizontal linear polarised X-ray beam), and sum the results for  $\theta = [0, 360]$ .
- (3) Scale the array to sum to  $E_{e^-} = E_X - E_K$ , where  $E_{e^-}$  was the average energy of a PE produced in the sample,  $E_X$  was the X-ray beam energy, and  $E_K$  was the average K-shell energy in the sample weighted by chemical composition and elemental cross-section. That is, the array (*Pt Src E Map* in Figure 6) represents the average distribution of energy (about a given point) due to a single photoionisation event.

### 4.4. Model Beam

The beam footprint was generated as a 2D Gaussian array which was scaled to sum to the total beam flux of  $1 \times 10^{10}$  photons  $s^{-1}$ . This array and the relevant interaction cross-sections (see Section 4.1) were then used to generate arrays whose elements were the rate of PE emission, Compton scattering, and coherent scattering (collectively *Interaction Rates* in Figure 6).

### 4.5. Model Energy Deposition

An array representing the energy deposited in the entire beam footprint (*2D Dose Map* in Figure 6) was produced by summing arrays representing the footprint due to photoelectric absorption and that due to Compton scattering. Two versions of the former were used: one modelled energy deposition due to photoelectric absorption in the absence of PE escape (the control case), which was obtained by multiplying the PE emission rate array by the average PE energy; and the other accounted for PE escape and was obtained by convolving the PE emission rate array with *Pt Src E Map*. The array representing energy deposition due to Compton scattering was modelled by multiplying the Compton scattering rate array by the average Compton recoil energy, which we took to be half the maximum recoil energy (see note below in Section 4.9). The resulting 2D array represented a slice through the beam orthogonal to its axis; energy deposition was assumed to be isotropic in the z-direction (parallel to the X-ray beam). Energy values were converted to dose based on the density of the crystal and the volume of a single voxel. One dose map was produced per set of beam/sample configurations. Finally, the array was cropped to the size of the crystal, assumed to be centred in the beam.

### 4.6. RADDPOSE-3D

RADDPOSE-3D was used as a comparative tool to produce 3D dose maps with sample- and beam-parameters matched, as closely as possible, to those used for the 2D dose maps. RADDPOSE-3D is in continuous development, with updates being made available to users via their GitHub repository (<https://github.com/GarmanGroup/RADDPOSE-3D>). We downloaded what was then the current

stable version of RADDPOSE-3D [9] on 22 June 2017, and used this version and the output it produced in all work presented in this study. The version we used did not include either Compton scattering or PE escape; recently, a new version has been made available that includes both these effects.

#### 4.7. Model Spot Intensity

The damage state of the crystal was represented by an array whose elements were values between 0 and 1, with 1 representing a completely undamaged region of the crystal, which would contribute strongly to the diffracted intensity, and 0 representing a completely disordered region. Each element of this array was linked to the element of the dose rate array with the corresponding row- and column-indices. The global dose (when accounting for both PE escape and Compton scattering) was used to determine the time required for the crystal to absorb a total dose of 60 MGy and this time was split into 300 equal time steps. At each time step, Equation (1) was used to calculate the values of the damage array, with  $I_0$  normalised to 1.  $H$  was Howell's criterion ( $10 \text{ MGy } \text{\AA}^{-1}$ ), and  $d$  was  $3 \text{ \AA}$ , the resolution of the Bragg spots we were simulating. For example, if the element of the dose rate array with indices  $x_1, y_1$  held the value  $0.3 \text{ MGy s}^{-1}$ , at time step  $t = 30 \text{ s}$  the element of the damage array with indices  $x_1, y_1$  would be set to

$$\exp\left(-\ln(2)\frac{0.3 \text{ MGy s}^{-1} \cdot 30 \text{ s}}{10 \text{ MGy } \text{\AA}^{-1} \cdot 3 \text{ \AA}}\right) = 0.8123.$$

At each time step, the elements of the damage array were summed and the total divided by the summed value of the array at time  $t = 0 \text{ s}$ . For each combination of crystal size, beam FWHM, and beam energy, this spot-fading model was run using three dose maps: the 2D dose maps generated as described in Section 4.5, one accounted for both PE escape and Compton scattering, one did not include PE escape but did include Compton scattering, and the third set of dose maps was produced using RADDPOSE-3D (the script was altered slightly to account for the RADDPOSE-3D maps being 3D rather than 2D).

At steps of 15 MGy, a qualitative indication of which parts of the crystal were contributing most to diffracted intensity was obtained by taking a point-wise product of the coherent scatter and damage arrays to produce a spatially-resolved map of diffraction-contribution. Reference to the coherent scatter array accounted for the fact that, even if a region of the crystal were largely undamaged, if there was minimal flux through that region, it would not contribute strongly to the diffracted intensity.

#### 4.8. Calculations of Diffraction Efficiency

We calculated diffraction efficiency by two methods: (i) using cross-section data from the NIST XCOM database [25]; and (ii) comparing the total coherent scatter to the global dose for each of our simulated data-sets. For calculations based on cross-sections for coherent scatter,  $\sigma_C$ , incoherent scatter,  $\sigma_I$ , and photoelectric absorption,  $\sigma_P$ , the diffraction efficiency was calculated as

$$DE(E) = \frac{\sigma_C}{\sigma_I + \sigma_P}. \quad (3)$$

$DE(12 \text{ keV})$  was our reference value and we expressed all other values as a ratio  $DE(E)/DE(12 \text{ keV})$ . For each of our simulated data sets, we calculated the diffraction efficiency (per unit volume) as

$$DE(E) = \frac{N_C}{\sigma_X^3 D_G}, \quad (4)$$

where  $N_C$  is the number of elastically scattered photons,  $\sigma_X$  is the side-length of the crystal, and  $D_G$  is the global dose in MGy. The reference value used was the diffraction efficiency of a  $1 \mu\text{m}$  crystal in a  $1 \mu\text{m}$  beam at 12 keV, as calculated using the RADDPOSE-3D dose maps, and all other values were expressed as a ratio to this reference.

#### 4.9. Simplifications and Assumptions in Our Model

We simulated the loss of integrated intensity of a single 3 Å Bragg spot (or, equivalently, all 3 Å spots in the diffraction pattern). Three angstroms was chosen as a reasonable minimal desirable resolution, though ultimately the choice is arbitrary. While it is noted in the literature that loss of integrated intensity is commonly used as a metric for radiation damage and that this may be measured in a full-data set, a repeated wedge, or single peak [13], it is not clear that these observations are necessarily equivalent. It is expected that simulating a single resolution is sufficient to provide a qualitative measure of the effect of including or excluding PE escape in simulations of radiation damage, which is the focus of this work. In theory, the simulation could be expanded to account for the fading of all spots in a diffraction pattern, between given minimum and maximum resolution, though this would significantly increase execution time.

The RADDPOSE-3D documentation [28] for the version we used in this study stated that it was not accurate at higher beam energies (>20 keV), advising users to instead use RADDPOSE v2 via the online web server (<http://www.raddo.se/legacy/>). The cause and nature of this inaccuracy were not given. RADDPOSE v2 does not produce spatially-resolved dose maps and provides the maximum dose for the case of a Gaussian beam, rather than the average dose across the beam profile [11]. The former point precluded us from using it in our spot-fading models; we used RADDPOSE-3D at all energies. To determine what effect this stated lack of accuracy may have had on the legitimacy of comparison between our simulations and RADDPOSE-3D, we compared the dose predictions of RADDPOSE v2 with the maximum dose predicted by RADDPOSE-3D at energies above 20 keV, for all crystal and beam sizes used in our simulations. In all cases, the RADDPOSE v2 dose was higher than the RADDPOSE-3D maximum dose by approximately 13%. Accordingly, we would expect results obtained with RADDPOSE-3D to *underestimate* damage rates at higher energies, compared to RADDPOSE v2.

In post-processing CASINO energy-deposition data to generate a 2D map of energy deposition due to a single photoionisation event, we took the intersection of the beam and sample surface (in CASINO) to be the origin of the photoionisation. A more precise model would take this point as the first collision in which the photoelectron was involved. That is, this point should be one IMFP away from the origin of photoionisation, potentially leaving a small area of reduced damage at the centre of the pattern. At the resolutions we used in our simulations, this is unlikely to cause significant error as the energy deposition will be displaced by only 2–5 pixels, depending on beam energy.

We made the simplifying assumption that the average Compton recoil energy was half the maximum recoil energy. Assuming a uniform distribution of Compton scattering angles (between 0° and 180°), an average taken over recoil energies calculated in 0.5° steps differs from this value by only 1–3% in the energy range we simulated, so this simplification has minimal impact on the outcome, with our simulations slightly under-predicting the impact of Compton scattering on dose.

## 5. Conclusions

We performed spatially- and temporally-resolved simulations of radiation damage to protein crystals during micro-crystallography experiments that account for the escape of photoelectrons from the diffraction volume. Our simulations explored damage rates for 1 µm, 1.5 µm, 2.0 µm and 5 µm crystals in beams with energy from 9 keV to 30 keV and beam size that was either matched to or larger than the crystal size. By accounting for the escape of photoelectrons from the diffraction volume, our simulations predict longer lifetimes (when compared to control simulations and simulations performed using a version of RADDPOSE-3D that did not account for PE escape), and increased diffraction efficiency (by a factor of between 7 and 30), for smaller crystals (<5 µm), provided the beam focus is not much larger than the crystal itself, and when using higher beam energy (≥20 keV). These results suggest significant benefit may be realised by conducting MX experiments on microcrystals at around 20 keV or higher. A systematic experimental exploration of radiation damage under the range of parameters we have simulated is now needed to validate and refine our models.

**Author Contributions:** Conceptualization, C.D. and B.A.; Formal analysis, H.M.; Methodology, H.M.; Project administration, B.A.; Software, H.M.; Supervision, C.D. and B.A.; Writing—original draft, H.M.; Writing—review & editing, C.D. and B.A.

**Funding:** This work was supported by the Australian Research Council Centre of Excellence in Advanced Molecular Imaging (CE140100011) (<http://www.imagingcoe.org/>).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

BCDI	Bragg coherent diffractive imaging
FWHM	full-width at half-maximum
IMFP	inelastic mean free path
MC	Monte Carlo
MX	macromolecular crystallography
PE	photoelectron
RSM	reciprocal space map(ping)

## Appendix A. Diffraction Efficiency for Simulations with No Photoelectron Escape

Diffraction efficiency of all individual simulated data sets that did not incorporate PE escape are shown in Tables A1 and A2, with averages and standard deviation.

**Table A1.** Diffraction efficiency, relative to the 12 keV value for RADDOSE-3D simulations, of all simulated data sets based on our control dose maps that included Compton scattering but excluded PE escape. The data show minimal variation with beam and crystal size.

Energy [keV]:	9	12	15	20	25	30
$\sigma_X = 1 \mu\text{m}, \sigma_B = 1 \mu\text{m}$ :	0.81	0.99	1.15	1.37	1.51	1.63
$\sigma_X = 1 \mu\text{m}, \sigma_B = 1.5 \mu\text{m}$ :	0.81	0.99	1.15	1.37	1.51	1.63
$\sigma_X = 1 \mu\text{m}, \sigma_B = 2 \mu\text{m}$ :	0.81	0.99	1.15	1.37	1.51	1.63
$\sigma_X = 1 \mu\text{m}, \sigma_B = 5 \mu\text{m}$ :	0.81	0.99	1.15	1.37	1.51	1.63
$\sigma_X = 1.5 \mu\text{m}, \sigma_B = 1.5 \mu\text{m}$ :	0.81	0.99	1.15	1.40	1.51	1.56
$\sigma_X = 2 \mu\text{m}, \sigma_B = 2 \mu\text{m}$ :	0.81	0.99	1.15	1.37	1.51	1.53
$\sigma_X = 5 \mu\text{m}, \sigma_B = 5 \mu\text{m}$ :	0.81	0.99	1.15	1.37	1.51	1.55
<b>Average:</b>	0.81	0.99	1.15	1.37	1.51	1.59
<b>Standard Deviation (%):</b>	(0.00%)	(0.00%)	(0.00%)	(1.01%)	(0.00%)	(2.67%)

**Table A2.** Diffraction efficiency, relative to 12 keV, of all simulated data sets based on RADDOSE-3D [9] dose maps that did not account for either Compton scattering or PE escape. The data show minimal variation with beam and crystal size.

Energy [keV]:	9	12	15	20	25	30
$\sigma_X = 1 \mu\text{m}, \sigma_B = 1 \mu\text{m}$ :	0.79	1.00	1.19	1.46	1.69	1.88
$\sigma_X = 1 \mu\text{m}, \sigma_B = 1.5 \mu\text{m}$ :	0.79	1.00	1.19	1.46	1.70	1.87
$\sigma_X = 1 \mu\text{m}, \sigma_B = 2 \mu\text{m}$ :	0.79	1.00	1.19	1.46	1.70	1.88
$\sigma_X = 1 \mu\text{m}, \sigma_B = 5 \mu\text{m}$ :	0.79	1.00	1.19	1.46	1.70	1.88
$\sigma_X = 1.5 \mu\text{m}, \sigma_B = 1.5 \mu\text{m}$ :	0.79	1.00	1.19	1.48	1.69	2.03
$\sigma_X = 2 \mu\text{m}, \sigma_B = 2 \mu\text{m}$ :	0.79	1.00	1.20	1.47	1.69	1.94
$\sigma_X = 5 \mu\text{m}, \sigma_B = 5 \mu\text{m}$ :	0.79	1.00	1.19	1.47	1.70	1.90
<b>Average:</b>	0.79	1.00	1.19	1.47	1.69	1.91
<b>Standard Deviation (%):</b>	(0.05%)	(0.13%)	(0.12%)	(0.5%)	(0.18%)	(3.06%)

## Appendix B. Global Dose Comparisons

The RADDOSE-3D documentation (for the version we used in this study) states that dose predictions are known to be inaccurate above 20 keV and recommends the use of RADDOSE v2,

which is not spatially-resolved, in this energy regime. As we required spatial-resolution for our comparisons, we had to use RADDOSE-3D at all energies. To determine what effect this would have on our comparisons, we repeated global dose predictions for energies above 20 keV with RADDOSE v2 as well as RADDOSE-3D. RADDOSE v2 gives a maximum dose when using Gaussian beams, rather than an average dose. We compared it to the maximum dose as calculated using RADDOSE-3D. As Table A3 shows, RADDOSE-3D predictions are lower than RADDOSE v2 predictions.

**Table A3.** Global dose, as calculated by our simulations, RADDOSE-3D [9], and (for beam energies  $\geq 20$  keV) RADDOSE v2, for a  $1\ \mu\text{m}$  crystal in a  $1\ \mu\text{m}$  (FWHM) beam with flux of  $1 \times 10^{10}$  photons  $\text{s}^{-1}$  and energy ranging from 9 keV to 30 keV. The total dose calculated by RADDOSE v2 for Gaussian beams is the maximum dose, not the average, hence, we compared it to the maximum dose reported by RADDOSE-3D.

Energy [keV]	Our Simulation (Average)	Dose [MGy] (Average)	RADDOSE-3D (Maximum)	RADDOSE v2 (Maximum)
9	2.70	5.26	8.02	-
12	0.78	2.83	4.32	-
15	0.29	1.74	2.65	-
20	0.09	0.92	1.40	1.59
25	0.04	0.56	0.85	0.97
30	0.03	0.37	0.57	0.64

## References

- Hettel, R. DLSR design and plans: An international overview. *J. Synchrotron Radiat.* **2014**, *21*, 843–855. [[CrossRef](#)] [[PubMed](#)]
- Schroer, C.G.; Falkenberg, G. Hard X-ray nanofocusing at low-emittance synchrotron radiation sources. *J. Synchrotron Radiat.* **2014**, *21*, 996–1005. [[CrossRef](#)] [[PubMed](#)]
- Garman, E.F.; Weik, M. Radiation Damage in Macromolecular Crystallography. *Synchrotron Radiat. News* **2015**, *28*, 15–19. [[CrossRef](#)]
- Garman, E.F.; Weik, M. Radiation damage to macromolecules: Kill or cure? *J. Synchrotron Radiat.* **2015**, *22*, 195–200. [[CrossRef](#)] [[PubMed](#)]
- Garman, E.F.; Weik, M. Radiation damage to biological macromolecules: Some answers and more questions. *J. Synchrotron Radiat.* **2013**, *20*, 1–6. [[CrossRef](#)] [[PubMed](#)]
- Garman, E.F. Radiation damage in macromolecular crystallography: What is it and why should we care? *Acta Cryst. D* **2010**, *66*, 339–351. [[CrossRef](#)] [[PubMed](#)]
- Holton, J.M. A beginner's guide to radiation damage. *J. Synchrotron Radiat.* **2009**, *16*, 133–142. [[CrossRef](#)] [[PubMed](#)]
- Howells, M.R.; Beetz, T.; Chapman, H.N.; Cui, C.; Jacobsen, C.J.; Kirz, J.; Lima, E.; Marchesini, S.; Miao, H.; Sayre, D.; et al. An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. *J. Electron Spectrosc. Relat. Phenom.* **2009**, *170*, 4–12. [[CrossRef](#)] [[PubMed](#)]
- Zeldin, O.B.; Gerstel, M.; Garman, E.F. RADDOSE-3D: Time- and space-resolved modelling of dose in macromolecular crystallography. *J. Appl. Crystallogr.* **2013**, *46*, 1225–1230. [[CrossRef](#)]
- Paithankar, K.S.; Garman, E.F. Know your dose: RADDOSE. *Acta Cryst. D* **2010**, *66*, 381–388. [[CrossRef](#)] [[PubMed](#)]
- Paithankar, K.S.; Owen, R.L.; Garman, E.F. Absorbed dose calculations for macromolecular crystals: Improvements to RADDOSE. *J. Synchrotron Radiat.* **2009**, *16*, 152–162. [[CrossRef](#)] [[PubMed](#)]
- Murray, J.W.; Garman, E.F.; Ravelli, R.B.G. X-ray absorption by macromolecular crystals: The effects of wavelength and crystal composition on absorbed dose. *J. Appl. Crystallogr.* **2004**, *37*, 513–522. [[CrossRef](#)]
- Garman, E.F.; Weik, M. X-ray radiation damage to biological macromolecules: Further insights. *J. Synchrotron Radiat.* **2017**, *24*, 1–6. [[CrossRef](#)] [[PubMed](#)]

14. Coughlan, H.D.; Darmanin, C.; Kirkwood, H.J.; Phillips, N.W.; Hoxley, D.; Clark, J.N.; Vine, D.J.; Hofmann, F.; Harder, R.J.; Maxey, E.; et al. Bragg coherent diffraction imaging and metrics for radiation damage in protein micro-crystallography. *J. Synchrotron Radiat.* **2017**, *24*, 83–94. [[CrossRef](#)] [[PubMed](#)]
15. Coughlan, H.D.; Darmanin, C.; Phillips, N.W.; Hofmann, F.; Clark, J.N.; Harder, R.J.; Vine, D.J.; Abbey, B. Radiation damage in a micron-sized protein crystal studied via reciprocal space mapping and coherent diffractive imaging. *Struct. Dyn.* **2015**, *2*, 041704. [[CrossRef](#)] [[PubMed](#)]
16. Finfrook, Y.Z.; Stern, E.A.; Alkire, R.W.; Kas, J.J.; Evans-Lutterodt, K.; Stein, A.; Duke, N.; Lazarski, K.; Joachimiak, A. Mitigation of X-ray damage in macromolecular crystallography by submicrometre line focusing. *Acta Cryst. D* **2013**, *69*, 1463–1469. [[CrossRef](#)] [[PubMed](#)]
17. 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. Radiation damage in protein crystals is reduced with a micron-sized X-ray beam. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6127–6132. [[CrossRef](#)] [[PubMed](#)]
18. Finfrook, Y.Z.; Stern, E.A.; Yacoby, Y.; Alkire, R.W.; Evans-Lutterodt, K.; Stein, A.; Isakovic, A.F.; Kas, J.J.; Joachimiak, A. Spatial dependence and mitigation of radiation damage by a line-focus mini-beam. *Acta Cryst. D* **2010**, *66*, 1287–1294. [[CrossRef](#)] [[PubMed](#)]
19. Bury, C.S.; Brooks-Bartlett, J.C.; Walsh, S.P.; Garman, E.F. Estimate your dose: RADDPOSE-3D. *Protein Sci.* **2018**, *27*, 217–228. [[CrossRef](#)] [[PubMed](#)]
20. Cowan, J.A.; Nave, C. The optimum conditions to collect X-ray data from very small samples. *J. Synchrotron Radiat.* **2008**, *15*, 458–462. [[CrossRef](#)] [[PubMed](#)]
21. Nave, C.; Hill, M.A. Will reduced radiation damage occur with very small crystals? *J. Synchrotron Radiat.* **2005**, *12*, 299–303. [[CrossRef](#)] [[PubMed](#)]
22. Stern, E.A.; Yacoby, Y.; Seidler, G.T.; Nagle, K.P.; Prange, M.P.; Sorini, A.P.; Rehr, J.J.; Joachimiak, A. Reducing radiation damage in macromolecular crystals at synchrotron sources. *Acta Cryst. D* **2009**, *65*, 366–374. [[CrossRef](#)] [[PubMed](#)]
23. Drouin, D.; Couture, A.R.; Joly, D.; Tastet, X.; Aimez, V.; Gauvin, R. CASINO V2.42—A Fast and Easy-to-use Modeling Tool for Scanning Electron Microscopy and Microanalysis Users. *Scanning* **2007**, *29*, 92–101. [[CrossRef](#)] [[PubMed](#)]
24. Drouin, D.; Hovington, P.; Gauvin, R. CASINO: A New Monte Carlo Code in C Language for Electron Beam Interaction—Part II: Tabulated Values of the Mott Cross Section. *Scanning* **1997**, *19*, 20–28. [[CrossRef](#)]
25. Berger, M.; Hubbell, J.; Seltzer, S.; Chang, J.; Coursey, J.; Sukumar, R.; Zucker, D.; Olsen, K. XCOM: Photon Cross Sections Database. Available online: <http://physics.nist.gov/xcom> (accessed on 22 September 2017).
26. Warkentin, M.A.; Atakisi, H.; Hopkins, J.B.; Walko, D.; Thorne, R.E. Lifetimes and spatio-temporal response of protein crystals in intense X-ray microbeams. *IUCr* **2017**, *4*, 785–794 [[CrossRef](#)] [[PubMed](#)]
27. Holton, J.M.; Frankel, K.A. The minimum crystal size needed for a complete diffraction data set. *Acta Cryst.* **2010**, *D66*, 393–408. [[CrossRef](#)] [[PubMed](#)]
28. RADDPOSE-3D Readme File. Available online: <https://github.com/GarmanGroup/RADDPOSE-3D/blob/master/README.md> (accessed on 21 June 2018).



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 3.5 Summary of Results & Discussion

We outline here the key components of the results and discussion presented in the paper.

Our simulation used three parallel streams in the spatially-resolved dose map module to test the potential for PE escape to mitigate damage in MX with micro-crystals. RADDPOSE-3D [66] was used as a control, based on its widespread use in providing realistic dose estimates in traditional MX. Of the two sets of dose maps we simulated, one did not include PE escape, the other did. We examined the effect of variation in crystal size, beam size, and beam energy. Across all results, our simulations that excluded PE escape were very closely aligned with RADDPOSE-3D, with a small divergence at higher energies due to our inclusion of Compton scattering.

### 3.5.1 Effect of Beam/Crystal Size

We observed a significant improvement in crystal lifetime for small crystals with small beams (see figs. 1 and 2 in the paper (pp. 4–5)). When the crystal and beam were both 1  $\mu\text{m}$ , the simulated lifetime was 4 times longer when PE escape was taken into account. This effect was reduced with an increase in beam size, with no difference apparent at 5  $\mu\text{m}$ . The beam size was the limiting factor rather than the crystal; the results for a 2  $\mu\text{m}$  crystal in a 2  $\mu\text{m}$  beam were not significantly different to those for a 1  $\mu\text{m}$  crystal in a 2  $\mu\text{m}$  beam. This is because PEs are produced in the solvent as well as the crystal, whence they scatter back into the crystal, offsetting the effect of PEs escaping it. The balance of this effect will be impacted by the heavy-atom content of the protein and solvent, for example a high salt concentration in the solvent (not unusual) could result in a cross section for photoionisation in the solvent that is higher than that of the crystal, potentially increasing, rather than mitigating, damage.

### 3.5.2 Effect of Beam Energy

Our examination of the effect of increasing the beam energy was restricted to a 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  beam. We observed a large improvement in crystal lifetime between 12 keV and 20 keV and continuing, though reduced, improvement between 20 keV and 30 keV (see fig. 3 in the paper (p. 6)). This reduced benefit above 20 keV is possibly due to the majority of PE energy already leaving the crystal at this point, with further increase in energy merely serving to scatter the PEs further into the surrounding solvent. Cowan and Nave [12] predicted a continued improvement in crystal lifetime up to and beyond 30 keV for larger crystals (up to 20  $\mu\text{m}$ ).

### 3.5.3 Diffraction Efficiency at Higher Beam Energies

Diffraction efficiency is the ratio of the number of coherently scattered photons per unit volume,  $N_C/\sigma_X^3$  for a cubic volume, to the average global dose,  $D_G$ . Even in the absence of PE escape diffraction efficiency improves with increased beam energy; cross sections for elastic scattering and photoionisation are both decreased at higher energy, but the photoionisation cross section falls more steeply than the elastic scattering cross section. Based on cross sections from the NIST XCOM database [114], diffraction efficiency should increase by a factor of 2 between 12 keV and 30 keV, without accounting for PE escape.

Using the diffraction efficiency at 12 keV, in the absence of PE escape, as a reference, we calculated an improvement by a factor of 3.6 at 12 keV, due to PE escape, and by a factor of 30 at 30 keV, due to the combination of PE escape and the relative change in cross sections. The improvement curve is steepest between 15 keV and 25 keV, and the overall benefit is reduced with an increase in crystal and/or beam size. Results are shown in fig. 4 of the paper (p. 7)

### 3.5.4 Evolution of the Diffraction Volume

At the beginning of an exposure, the diffracted signal is strongest from sample within the central part of the (Gaussian) beam, in proportion to the increased flux density. However, this region of the crystal absorbs a higher dose, relative to the edges, damaging more quickly. After significant exposure, the total diffracted intensity is reduced, and a higher portion of it comes from the less damaged regions of the crystal, in the tails of the beam. The effect is illustrated in fig. 5 of the paper (pp. 8–9) Our observations of the change in diffraction contribution are consistent with calculations of Warkentin *et al.* [76], and an experiment by Coughlan *et al.* [26, 28] who used Bragg coherent diffractive imaging (BCDI) to observe the change in real space of micro-crystals subject to high total doses (100s of MGy). The impact of non-uniform illumination of crystals is discussed in more depth in Chapter 5.

## 3.6 Additional Discussion

Any simulation is subject to simplifications and limitations. Ours is more theoretical than some, prioritising simulated detail over an attempt to represent real world conditions. For example, noise is not considered in our simulation, although it is omnipresent in diffraction experiments and places limitations on performance beyond the theoretical maxima. The theoretical benefits predicted in our simulated results cannot be expected to be replicated exactly in actual experiments, especially early experiments that must contend with the technical challenges of operating

beamlines outside their typical parameters. The simulation is also tightly constrained in *scope*: it provides a prediction of the rate at which a Bragg spot of a given resolution (selected by the user) will fade. It has nothing to say about whether spots of a given resolution will actually be present in a diffraction pattern, as it does not model the degree of crystalline disorder of the sample, which places a limit on observable diffraction resolution.

It is simple to simulate a beam of a given energy, with an arbitrary focus that permits illumination of the crystal but not the surrounding solvent. This is much more difficult in practice. Many MX beamlines are capable of operating at higher energies (with reduced flux) but this capability is seldom used, and the operation of beamlines beyond their nominal energy (somewhere around 12 keV) may not be well tested. Achieving very small foci is also something that may be within the theoretical operational capacity of the beamline but require significant optimisation time in practice. Simultaneously optimising both of these parameters is far from routine. An alternative path to minimising beam size is to minimise the amount of solvent around the crystal, which presents its own technical challenges. A larger beam is also more likely to illuminate multiple crystals if using a chip-style sample delivery system to obtain high throughput for serial crystallography.

One of the most significant experimental factors not accounted for in our simulation is the efficiency of the detector. X-rays interact weakly with matter, an effect that is proportional to energy. This weak interaction is critical to the mathematical interpretation of diffraction data, as it allows phenomena such as multiple scattering to be ignored, however it is problematic for actually detecting x-rays, as they must interact with the detector face to be observed. The majority of detectors are silicon based, giving them a much higher interaction cross section than organic matter; by increasing the thickness of the active layer (up to a point) the likelihood of x-rays being successfully registered by the detector is improved, normally at the cost of spatial resolution. However, detector efficiency generally diminishes with increasing x-ray energy, offsetting gains in diffraction efficiency. Our results in this chapter suggest that higher energy is unquestionably better in terms of improved diffraction efficiency, and past studies [12] have suggested even higher energies, up to 40 keV as ideal. Shortly after the publication of our paper, Dickerson and Garman [19] published simulations that accounted for diffraction efficiency. Their simulations predicted that energies beyond 26 keV were not likely to yield real-world improvement in crystal lifetimes, and this was assuming the use of new cadmium-telluride (CdTe) detectors, with an improved efficiency at higher energies. Silicon based detectors are unlikely to be able to take advantage of the improved diffraction efficiency at higher energies.

Despite the real-world factors not being accounted for in our simulation, our results suggest that there is real potential for damage mitigation in micro-crystallography. Real-world data

are required for validation of our model's predictions, and to illuminate the extent to which experimental constraints will impact any damage mitigation that may occur. Subsequent to the publication of this paper we conducted two experiments designed to test the predictions made here; they are described in Chapter 4, which also includes further discussion of the practical considerations raised in this section.

### 3.7 Summary

We constructed a detailed spatially- and temporally-resolved simulation of the impact of radiation damage on Bragg spot intensity, incorporating Monte Carlo simulation of electron trajectories via CASINO [107], and used the predictions of RADDPOSE-3D [66] to validate our approach. We used our simulation to test the difference in spot fading rates for a 3 Å Bragg peak, for crystal and beam sizes between 1 µm and 5 µm, and beam energies between 9 keV and 30 keV, comparing rates when PE motion was taken into account or not.

The results of our simulation showed a significant improvement in diffraction efficiency (elastically scattered flux per unit dose) at higher x-ray energies, especially for the smallest crystal and beam sizes: when accounting for PE escape, the diffraction efficiency for a 1 µm crystal centred in a 1 µm, 30 keV beam was 30 times greater than the reference value (12 keV with no PE escape). This improvement was diminished when beam foci were larger than the crystal, resulting in illumination of surrounding solvent. Little to no benefit was seen when the size of the crystal or beam was greater than 5 µm, at standard x-ray energy (approx 12 keV).

The significant improvement in lifetime of small crystals seen in our simulations at high x-ray energies suggest there is potential utility in conducting MX experiments at higher energy than at the 12 keV typically used, especially for smaller crystals. There are several technical challenges to be addressed in order to realise such experiments. Although developments such as the diffraction limited storage rings of 4<sup>th</sup> generation synchrotrons are making the beam sizes used in this simulation possible, they are currently not routine. Operating an MX beamline at higher x-ray energy requires careful, and potentially time-consuming, adjustment and calibration, and will come at the expense of total flux and possibly reduced focusing capacity. Exclusion of solvent-illumination is critical to gaining an extension of crystal lifetime via PE escape; this may be accomplished by minimising solvent around the crystal or carefully matching the beam to the crystal, both approaches require finesse. Detector inefficiency at higher energies can mitigate any improvement in diffraction efficiency of the crystal; newer cadmium-telluride detectors have a sufficiently flat response curve up to around 26 keV and should be used in preference to silicon detectors. Early experiments should focus on experimental verification of the lifetime

improvements that might be realised under realistic conditions. Such experiments could serve to highlight further challenges in high energy micro-crystallography in advance of experiments to actually determine new protein structures under these novel conditions.

In order to explore the validity of the model introduced in this chapter, we performed two experiments designed to test our results, these are described in Chapter 4.

## Chapter 4

# Radiation Damage Experiments Using Protein Micro-Crystals

In this chapter we describe two iterations of an experiment designed to test the theoretical predictions discussed in Chapter 3. The experiments took place at the P11 beamline at PETRA-III at DESY (the Deutsches Elektronen-Synchrotron) in Hamburg, 13–16 September 2018; and the 17-ID-2 FMX beamline at NSLS-II at Brookhaven National Laboratory, New York, 11–13 August 2019. Both experiments had the same overarching objective, with adjustments to the experimental plan made based on available apparatus. An attempt was made to optimise the experimental conditions for the second experiment based on the results of the first experiment.

Section 4.1 contains a review of relevant background. Materials and methods are covered in section 4.2, which is divided into three parts: 4.2.1 describes crystallisation, this protocol is identical for both experiments; 4.2.2 and 4.2.3 describe the PETRA-III and NSLS-II experiments, respectively. Results are presented in section 4.3 and discussed in section 4.4.

### 4.1 Introduction

Radiation damage has been a limiting factor in structure determination via MX since its inception. Researchers continue to push the boundaries in a quest to obtain structural information from smaller and less-perfect crystal samples. For over a decade, computational simulations have suggested that radiation damage might be significantly mitigated when working with crystals smaller than a few microns in size, provided that the beam can be focused to the same size as the crystal [11, 12, 21, 104]. If these predictions can be realised in practice, it would be of significant benefit for structural biologists, who are increasingly working with protein targets that do not readily form large crystals. Reliably focusing an x-ray beam to such a small spot-size has not been routinely attainable; something that is beginning to change with the arrival of 4<sup>th</sup>

generation synchrotrons with incredibly low-emittance storage rings [39, 40, 115, 116]. Although not yet common-place by any means, the ability of MX beamlines to provide spot-sizes down to 1  $\mu\text{m}$  (or very close) should become increasingly common, providing the possibility for experimenters to achieve the conditions under which radiation damage is predicted to be greatly reduced. As such, there is great interest in testing the predictions of radiation damage using computational simulations to gauge what kind of benefit might be expected under real-world conditions.

As discussed in Chapter 2, x-rays can interact with matter in several ways. Under the conditions relevant to this chapter (x-ray energies on the order of 10 keV interacting with biological samples, composed primarily of light atoms), there are four processes that may occur to x-rays when encountering the sample:

1. They will pass through the sample without interacting at all; for thin biological samples, this is by far the most likely occurrence;
2. They will scatter elastically; it is this interaction that provides the 'signal' in a macromolecular crystallography experiment, with elastically scattered photons contributing to the diffraction patterns created at the detector;
3. They will scatter inelastically, a process also known as Compton Scattering, in doing so they will impart some energy to the electrons of the samples, essentially heating it. A small number of inelastically scattered x-rays may also reach the detector, contributing to background noise, though most will scatter at large enough angles to miss the detector;
4. They will be absorbed by the sample, resulting in ejection of inner-shell electrons (most typically *K*-shell), known as photoelectrons (PEs). These PEs will be ejected with kinetic energy close to that of the incident beam energy and will therefore move through the material, colliding (inelastically) with other atomic electrons, gradually depositing their energy in a small volume surrounding the point of the initial x-ray absorption. Other than transmission, this is most likely result, occurring around 10 times as often as elastic scattering.

These interactions are summarised in fig. 2.1.

Photoionisation predominantly involves *K*-shell electrons. Elastic and inelastic scattering may result from interactions with any atomic electrons. Both Compton scattering and photoabsorption result in the deposition of energy in the sample, damaging it. Photoionisation is the primary driver of radiation damage in MX for two reasons: it has a higher cross-section than Compton scattering (at the relevant energies), and PEs have, individually, much higher energy than

Compton recoil electrons. Damage manifests in many ways, broadly categorised as *specific damage*, which is observed in real space; and *global damage*, observed in reciprocal space [74]. Perhaps the simplest damage metric to observe (especially in real time) is the loss of intensity in diffraction patterns [71], a marker of global damage; it is this metric that we observed as our figure of merit in the experiments described in this chapter.

For samples at cryogenic temperatures, there is evidence that the rate at which damage accumulates is proportional to the total dose (absorbed energy per unit mass), independent of the rate at which this dose accrues up to a flux density of  $1 \times 10^{15}$  photons  $\text{s}^{-1} \text{mm}^{-2}$  [117–119]. While for room temperature samples, the dose-rate does affect the damage rate [120, 121]. Dose cannot be directly measured in experiments; it is estimated based on beam- and sample-parameters. Damage metrics are typically tracked against dose, rather than time, as this allows comparison between beamlines—x-ray beams with different photon flux-densities and beam size may deliver vastly different doses to the same crystal in a given time—and samples—a denser sample will absorb a lower dose than a less dense sample when exposed to the same flux.

The lack of correlation between damage rate and dose rate is an important result for MX as the flux-densities available at MX beamlines have increased significantly in recent years, with a concomitant increase in dose rate. This trend continues with the arrival of a fourth generation of synchrotrons [39, 40, 115, 116], which are associated with an increase in available flux. Coupled with improvements in beamline optics and detectors [45], this increased flux has been a boon for MX, allowing the use of smaller and less-perfect crystals in structure solution. At the same time, continuing to push the boundaries of crystal-suitability has meant that the issue of radiation damage has remained a limiting factor in experiments, resulting in a sustained interest in the scientific literature in the pathways, effects, and potential mitigation of radiation damage [11, 12, 27, 28, 104, 109, 110, 122].

One potential pathway for mitigating radiation damage to small crystals is the dispersal of absorbed energy through PE escape. As noted above, the energy transferred to a sample by PEs is spread over a finite volume; as early as 2005 [11], it was proposed that this mechanism may lead to a reduction in dose when using small crystals, small beam sizes, and higher beam energies. Over the last fifteen years, several studies [12, 104, 111] have simulated radiation damage with varying degrees of sophistication, in order to explore the specific conditions under which PE escape might benefit MX experiments. At the same time, a handful of experiments have investigated the potential for damage reduction when using small beams [27, 104], line focus beams [108] and/or crystals with particular morphologies [108] such as platelettes and needles. To a limited extent, the effect of beam energy has also been investigated [109, 110]. At the time of our experiment, these investigations had yet to be put together into a systematic, experimentally

validated, model of the precise relationship between experimental parameters and the potential for damage reduction due to PE escape.

In our simulations (see Chapter 3) we examined the potential impact of crystal size, beam size, and beam energy on the damage rate in MX experiments, as indicated by loss of intensity in Bragg spots. We observed that increasing the beam energy greatly reduced the rate at which spots faded, and that this trend continued up to 30 keV, although the improvement began to plateau due to the increased contribution of Compton scattering to radiation damage, over photoionisation. One real-world limitation which was not accounted for in our simulation is the ability of an x-ray detector to accurately record the intensity of photon flux striking it, within an arbitrary energy range. Donath *et al.* [123] measured the quantum efficiency (QE) of detectors with silicon (Si) panels of varying depth, noting that QE falls rapidly above 15 keV, even for the thickest detector panels on the market (the peak energy is lower than 10 keV for some panels), and that QE is below 50% above 22.1 keV. Any potential gains expected from a reduction in the rate of radiation damage would have to exceed the loss of measured flux at higher energies, based on the QE of the detector. Recently, high quality CdTe detectors have begun to become available, although they are yet to replace or augment Si detectors on MX beamlines. Zambon *et al.* [124] measured the QE of a CdTe detector with a 750  $\mu\text{m}$  panel, observing a near-flat response up to just below 26.7 keV, the *K*-edge of cadmium. A comparison of the findings of both groups is shown in fig. 4.1a.

Dickerson and Garman [19] incorporated the QE of the detector into their recent model of radiation damage. This thorough model employs dose estimates made using RADDOSE-3D [66, 67] version 4, and also includes the effects of changes in interaction cross-sections, and PE escape to predict changes in diffraction efficiency (DE, the intensity of elastically scattered x-rays per unit dose) for crystals of various sizes, with both Si and CdTe detectors. They observed little-to-no improvement in DE for crystals larger than 2  $\mu\text{m}$ , when using a Si detector with a 1000  $\mu\text{m}$  thick sensor panel, and recommended the use of 26.00 keV as this energy leverages the maximal QE of a CdTe detector. A selection of their results are summarised in fig. 4.1b.

There is now a strong body of (simulated) evidence to suggest that PE escape might present an avenue for partial mitigation of radiation damage when using small beams, small crystals, and high energy x-rays (coupled with a CdTe detector). The main question, then, for our experiments is whether it is feasible, given currently available technology, to anticipate actually achieving a reduction in damage rate under these conditions. And if not, what areas would need to be further researched and developed.

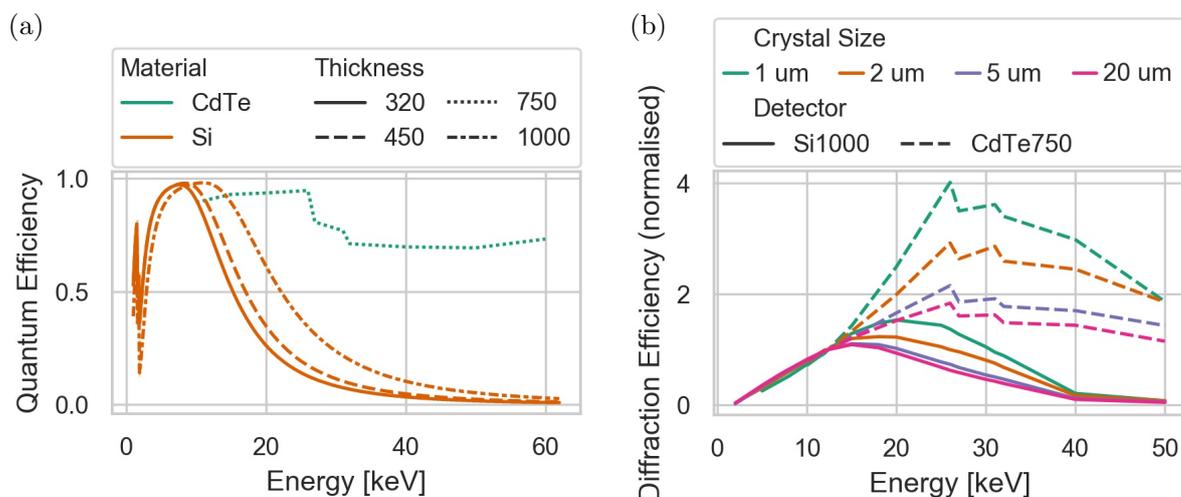


Figure 4.1: **Quantum efficiency & diffraction efficiency of silicon (Si) and cadmium-telluride (CdTe) detectors.** (a) Comparison of the theoretical quantum efficiency of Si detectors (orange) with three panel thicknesses —320  $\mu\text{m}$ , 450  $\mu\text{m}$  and 1000  $\mu\text{m}$ — and a CdTe detector (teal) with a 750  $\mu\text{m}$  thick panel. (b) Theoretical diffraction efficiency (elastically scattered photons per unit dose), normalised to 12.4 keV, of a 1000  $\mu\text{m}$  Si detector (solid lines) and a 750  $\mu\text{m}$  CdTe detector (dashed lines) for four crystal sizes: 1  $\mu\text{m}$  (teal), 2  $\mu\text{m}$  (orange), 5  $\mu\text{m}$  (lavender), and 20  $\mu\text{m}$  (pink). Plots in (b) account for variation in photoionisation and elastic scattering cross sections, and photoelectron escape. Plot data: (a) Donath *et al.* [123], Zambon *et al.* [124]; (b) Dickerson and Garman [19].

## 4.2 Materials and Methods

### 4.2.1 Sample Preparation & Characterisation

#### Materials and Crystallisation

Lysozyme was purchased from Sigma-Aldrich. Our crystalliation process was based on a batch-mode protocol for micro-crystals provided by Dr Dominik Oberthür [125] from the Centre for Free-Electron Laser Science (CFEL) at DESY. We conducted an optimisation trial to obtain a variety of crystal sizes.

The protein buffer was 50 mM acetate buffer. This was prepared by adding 40 mL of milliQ to a 50 mL tube, then adding 10 mg of sodium acetate and 143 mg of glacial acetic acid. We used hydrogen chloride (HCl) and sodium hydroxide (NaOH) to adjust the pH to 3.5.

4 mL of 130 mg mL<sup>-1</sup> lysozyme solution was prepared by adding 520 mg of lysozyme to a 10 mL tube and adding our buffer to bring the total volume to 4 mL.

10 mL each of 27 precipitant solutions were prepared, with varying concentrations of sodium chloride (1 M – 1.25 M), ethylene glycol (35% – 40%, w/v), and poly-ethylene glycol 3350 (PEG3350) (15% – 18%, w/v), as shown in table 4.1.

Table 4.1: **Precipitant solution constituents for crystallisation trials.** Crystallisation trials were performed by mixing 130 mg mL<sup>-1</sup> lysozyme in 50 mM acetate buffer (pH 3.5) at a 1:3 ratio with 27 precipitant solutions containing different concentrations of ethylene glycol (EG), salt (NaCl) and polyethylene glycol (PEG) 3350.

Trial				Trial			
		Concentration				Concentration	
Trial No.	EG (%)	NaCl (M)	PEG 3350 (%)	Trial No.	EG (%)	NaCl (M)	PEG 3350 (%)
1	35	1	15	15	38	1.2	20
2	35	1	18	16	38	1.25	15
3	35	1	20	17	38	1.25	18
4	35	1.2	15	18	38	1.25	20
5	35	1.2	18	19	40	1	15
6	35	1.2	20	20	40	1	18
7	35	1.25	15	21	40	1	20
8	35	1.25	18	22	40	1.2	15
9	35	1.25	20	23	40	1.2	18
10	38	1	15	24	40	1.2	20
11	38	1	18	25	40	1.25	15
12	38	1	20	26	40	1.25	18
13	38	1.2	15	27	40	1.25	20
14	38	1.2	18				

Crystallisation was performed in a cold-room, with an ambient temperature of 4 °C. Prior to crystallisation, protein buffer and precipitant solutions were cooled to 0 °C by storing their containers in an ice bath. We used a steel Eppendorf tray, also immersed in an ice bath to maintain batches at a low temperature ( $\leq 2.5$  °C) during crystallisation.

We prepared and labelled 27 1.5 mL Eppendorf tubes, which were stored in the cooled metal tray. 125  $\mu$ L of protein solution was added to each of these tubes. For each of the 27 samples, the following crystallisation processes was performed:

1. 375  $\mu$ L of precipitant was added to the protein solution,
2. Sample was vortexed for 30 s,

3. Sample was left to incubate (in the cooled steel tray) for 1 min to 2 min
4. Steps 2 and 3 were repeated three more times, the last two incubation periods were extended to 5 min each,
5. Sample was vortexed once more for 30 s.

After all batches were complete, they were left in the ice bath for 30 min before being stored at 4 °C.

This crystallisation trial was conducted prior to our PETRA-III experiment. For our NSLS-II experiment, we repeated the procedure described here, but used only 4 of the original 27 precipitant solutions: trial numbers 5, 7, 20, and 22 from table 4.1.

Crystals were stored at 4 °C at all times, except when samples were taken for optical imaging, and during transport to the beamlines.

### **Optical Characterisation**

Optical characterisation was performed at room temperature with an Olympus BX61 microscope, UIS2 objective lenses (10x and 40x), a Colorview II digital camera, and the *analySIS*<sup>®</sup> imaging software. Crystals from the initial trial (27 batches) were imaged ten days after crystallisation. Crystals from the second round of crystallisation (4 batches) were imaged three days after crystallisation. For each crystal batch, the sample was first vortexed to evenly distribute crystals, then a 2  $\mu$ L aliquot of sample was transferred to a microscope slide and the droplet was covered with a glass cover slip. We inspected the samples using the 10x and 40x objectives to locate regions on the slide that contained:

1. A reasonable concentration of crystals,
2. A representative sample of the crystal size distribution within the batch.

A single wide-field image was then taken with the 40x objective for each batch. Examples of optical microscopy images from the initial 27 trial batches listed in table 4.1 are shown in fig. 4.2.

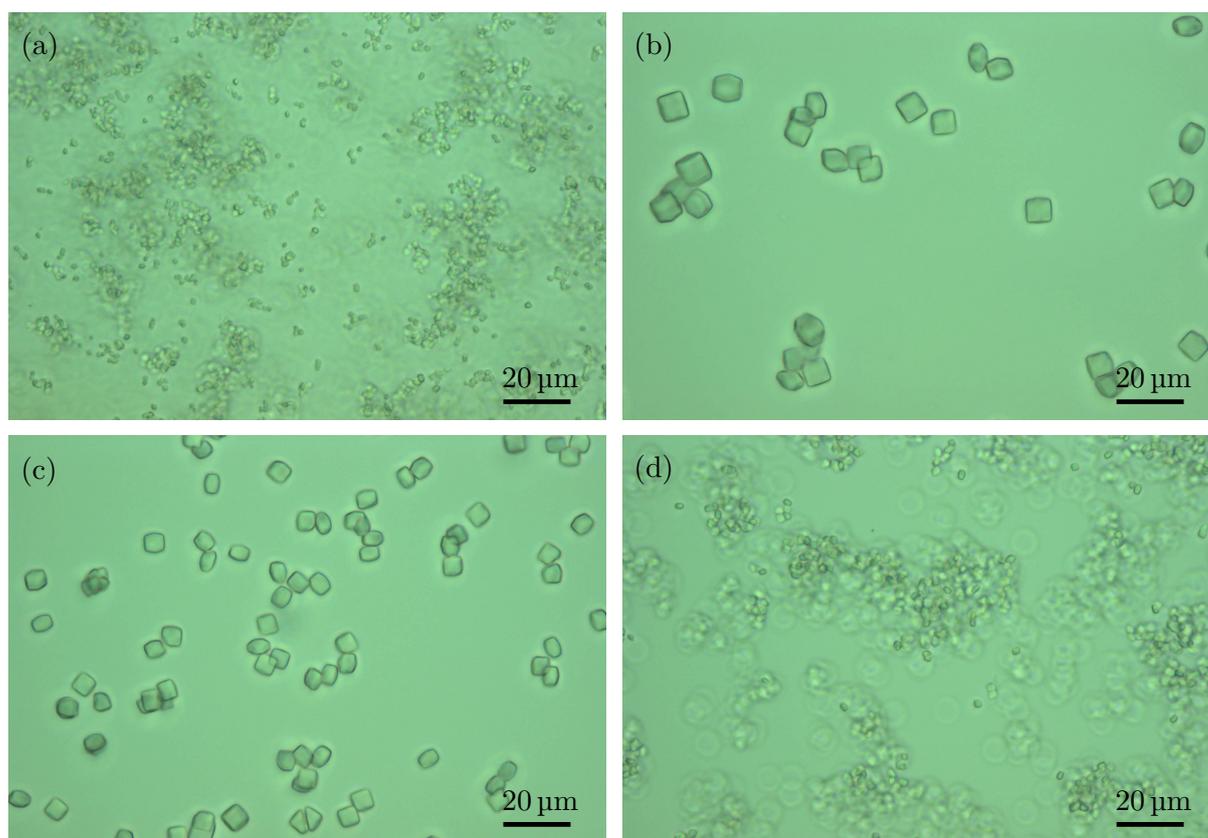


Figure 4.2: **Optical microscopy images of lysozyme crystals from first round of crystallisation.** Four of 27 batches of crystals are shown, highlighting the range of crystal sizes produced. The smaller crystals tended to sit on top of each other in the solution, resulting in the blurred patches seen in images (a) and (d) as the microscope could be focused on one layer or the other. The aspect ratio of the crystals, at all sizes, varied between 1.1 and 1.6.

### Size Analysis from Optical Characterisation

To determine the size distribution of the crystals in the sample we used the FIJI distribution of ImageJ [126–128] to measure as many crystals as possible in each image. A maximum of 50 crystals were measured in a given image. The smallest number of crystals measured in a single trial was nine. Of the initial 27 batches, 21 had at least 20 crystals measured. Initial size estimates are shown in fig. 4.3. The majority of the crystallisation conditions yielded crystals between 1 µm and 2 µm, with a handful of conditions resulting in larger crystals, up to 8 µm. In all batches, the crystals produced had an average aspect ratio of between 1:1.1 and 1:1.6.

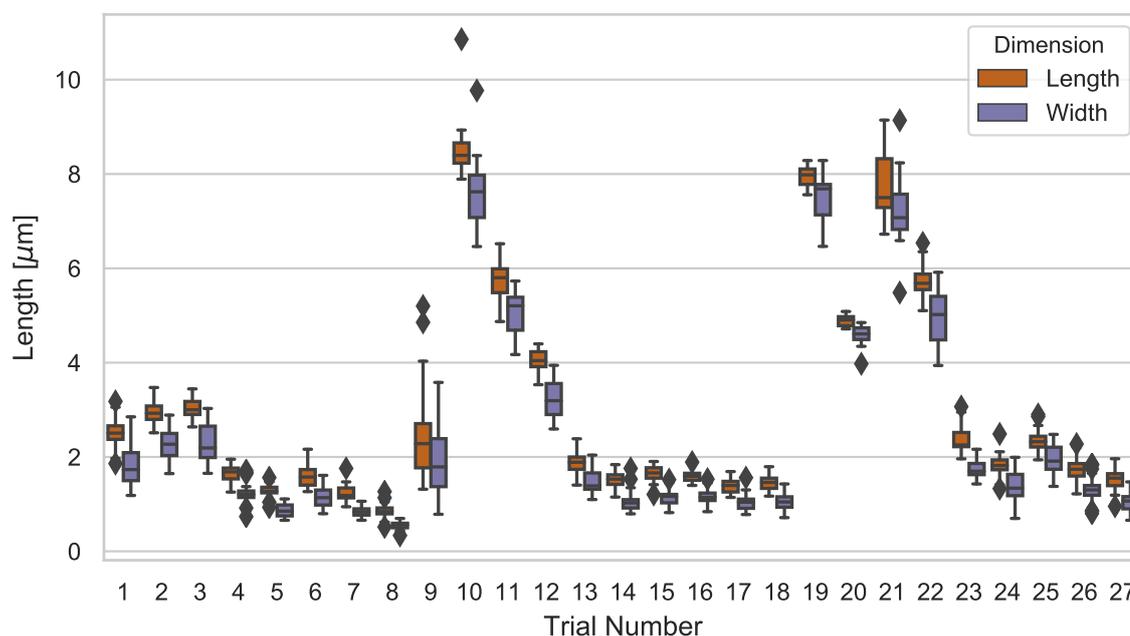


Figure 4.3: **Box plot showing size distribution of lysozyme crystals from 27 initial trials.** 27 batches of crystals were produced; between batches, the precipitant had varying concentrations of NaCl (1 M to 1.25 M), ethylene glycol (35% to 40% w/v), and PEG3500 (15% to 20% w/v). Crystal size varied from roughly 1 μm to 8 μm; aspect ratio varied from 1.1 to 1.6.

### TEM Characterisation

Characterisation using transmission electron microscopy (TEM) was performed for the crystals used in the second experiment (NLSL-II). To prepare the sample for TEM, we made a 1 in 10 dilution of the original stock. A 2 μL aliquot of our diluted crystal sample was applied to a carbon-coated TEM grid and allowed to sit for two minutes so that crystals settled on the grids. Blotter paper was then used to remove the solvent. The crystals were then immediately negative stained with the addition of 2 μL of 1% aqueous solution of uranyl acetate ( $\text{UO}_2(\text{CH}_3\text{COO})_2 \cdot 2\text{H}_2\text{O}$ ). Excess stain was immediately removed and after standing for several minutes to dry, the samples were stored at room temperature for 24 h before imaging.

Imaging was performed, with the assistance of Dr Julian Ratcliffe (La Trobe University), using a JEOL JEM2100 desktop transmission electron microscope. 26 images were collected, which were subsequently analysed using an offline version of Gatan's Digital Micrograph imaging software [129]. 44 crystals were measured across all images. Examples of TEM images can be seen in fig. 4.4.

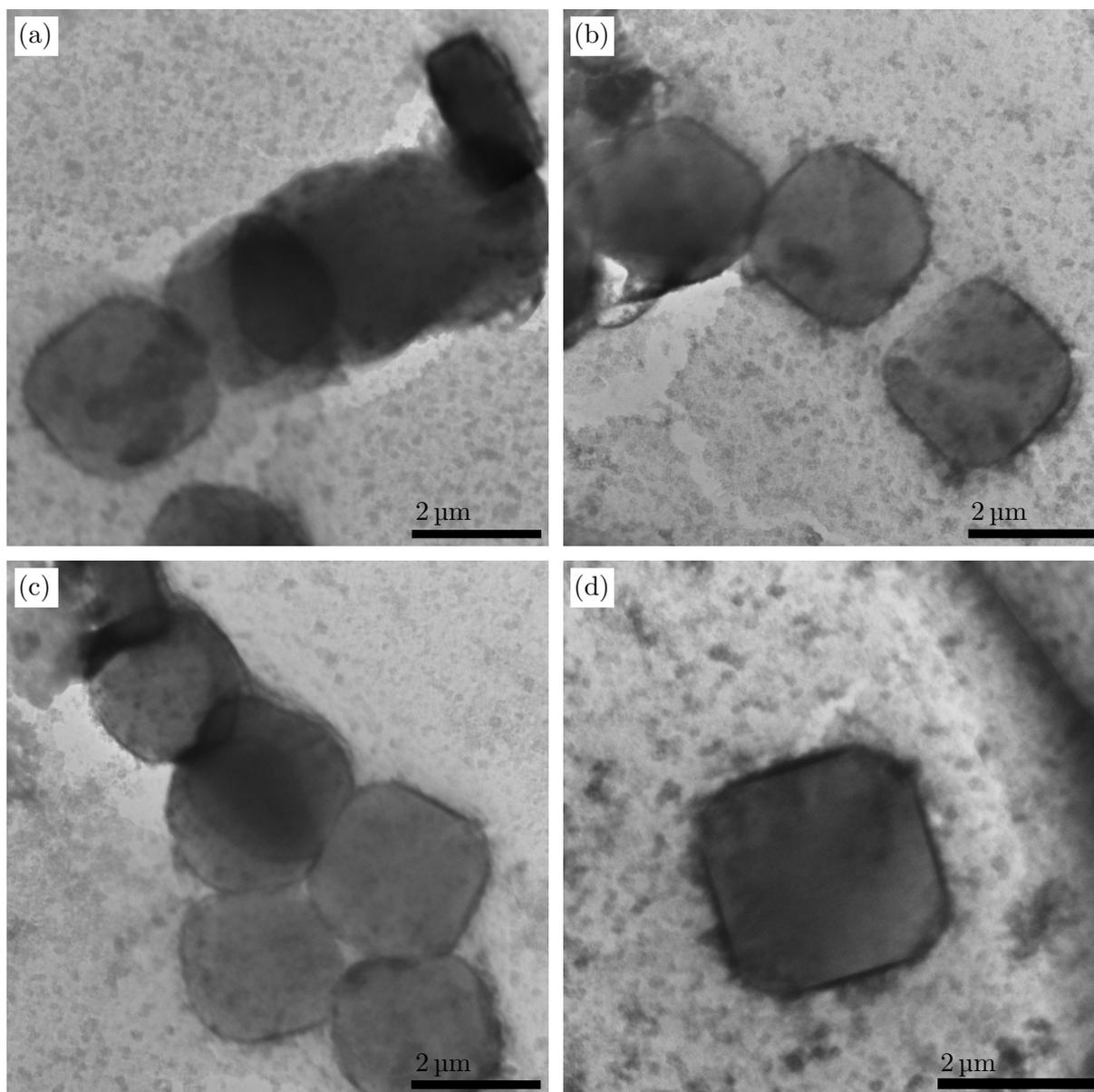


Figure 4.4: **TEM images of lysozyme crystals used in the NSLS-II experiment.** Crystals were negative stained with uranyl acetate 1%. The majority of the crystals imaged with TEM were roughly square, this is due to preferential alignment in the TEM grids. A rectangular crystal can be seen in (a); this crystal likely has the same three-dimensional aspect ratio as the others, but is aligned with the shorter axis visible. All images are from the same batch of crystals.

In contrast to the optical microscopy images, the TEM images primarily contain crystals with an aspect ratio close to 1:1. This indicates that the crystals have two dimensions that are roughly

equal, while being shorter in the third; in optical microscopy they seem to have been preferentially aligned such that the shorter dimension was visible, while in TEM grids they have preferentially aligned such that it is not. An example of a crystal aligned with the shorter dimension viewable can be seen in fig. 4.4a. An alternative is that there are two distinct shapes in the crystal population but given we saw overwhelmingly one orientation in optical microscopy and overwhelmingly the other in TEM, this is not deemed likely.

44 crystal measurements from TEM were sorted by aspect ratio in order to group them into two categories: those that were 'lying flat' and those that were 'on their side'. 35 crystals exhibited a close-to-square aspect ratio (between 1:1 and 1:1.4); 9 showed a markedly shorter axis, with an aspect ratio between 1:1.7 and 1:2.5. Based on this division, we estimated the crystal size to be  $(2.45 \pm 0.22) \mu\text{m} \times (2.34 \pm 0.26) \mu\text{m} \times (1.24 \pm 0.17) \mu\text{m}$ , with the uncertainty being taken as the standard deviation in the measurements.

## 4.2.2 X-ray Data Collection at PETRA-III

The first experiment was conducted at the P11 [130] Bio-Imaging and Diffraction beamline of PETRA-III at DESY in Hamburg, Germany.

### Sample Preparation

Batch 20 sample was used at the PETRA experiment (see table 4.1 and fig. 4.3). The crystal size was  $5 \mu\text{m} \times 4.5 \mu\text{m}$ .

For sample delivery, silicon chips were used. The chip design was a modified version of a design by Roedig *et al.* [54]. They were manufactured at the Melbourne Centre for Nanofabrication (MC<sup>N</sup>). The chips consisted of a  $2.5 \text{ mm} \times 4.0 \text{ mm}$  Si housing surrounding a  $1.5 \text{ mm} \times 1.5 \text{ mm}$  silicon nitride ( $\text{Si}_3\text{N}_4$ ) window (fig. 4.5a), patterned with a hexagonal lattice of apertures of varying size and shape. The outer housing was  $200 \mu\text{m}$  to  $300 \mu\text{m}$  thick; the silicon nitride windows were  $10 \mu\text{m}$  to  $20 \mu\text{m}$  thick. 50% of the window surface area was patterned with square apertures, with the remaining 50% having triangular apertures (fig. 4.5b). Inter-aperture spacing was  $20 \mu\text{m}$ . The aperture dimensions were  $5 \mu\text{m} \times 5 \mu\text{m}$  for the square apertures and  $7 \mu\text{m} \times 7 \mu\text{m}$  for the triangular apertures. The chips were affixed by an acrylic post to a magnetic base (fig. 4.5c) that could be directly attached to a standard macrocrystallography goniometer.

Prior to adding the sample to our chips, the crystal stocks were diluted to a 1:3 ratio with additional precipitant solution. No additional cryoprotectant was added as the concentration of PEG3350 in the solution was sufficient for this purpose. In order to maintain the hydration of the crystals while removing as much excess solvent as possible prior to cryogenic fixing, the

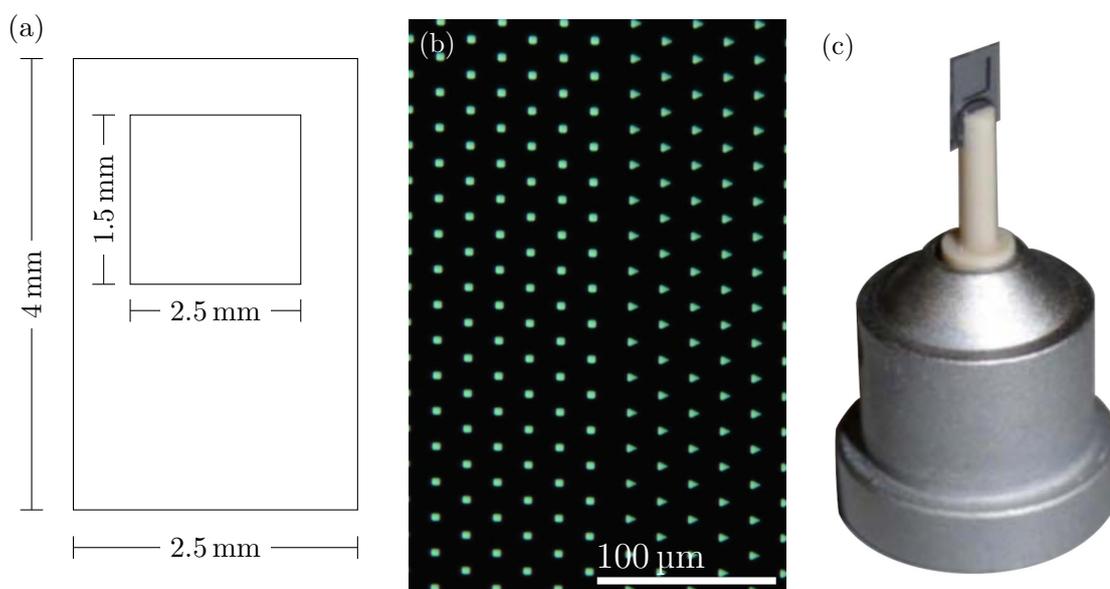


Figure 4.5: **Silicon chips used for sample delivery at the PETRA-III experiment.** (a) A schematic shows the dimensions of the chip; the thickness of the housing was  $200\ \mu\text{m}$  to  $300\ \mu\text{m}$ , the thickness of the window was  $10\ \mu\text{m}$  to  $20\ \mu\text{m}$ . (b) An optical microscope image of the patterning of the chip window ( $5\ \mu\text{m} \times 5\ \mu\text{m}$  squares,  $7\ \mu\text{m} \times 7\ \mu\text{m}$  triangles). (c) A photograph of the chip mounted, via an acrylic post, to a standard magnetic macrocrystallography goniometer cap.

samples were mounted —under an optical microscope— in a gentle jet of humidified air, as shown in fig. 4.6a. Crystals were evenly distributed in the solvent by vortexing the sample for 20 s before a  $2\ \mu\text{L}$  aliquot was taken and deposited on the chip window. Blotter paper was then gently applied to the reverse side of the window, drawing away excess solvent; this process was observed under the microscope and as soon as the excess solvent was removed, the chip was plunged into liquid nitrogen, wherein it was stored until it was mounted at the beamline.

### Beamline Configuration and Data Collection

Samples were mounted on the beamline goniometer as shown in fig. 4.6b. The sample was continually cooled by a jet of gaseous nitrogen to 100 K. The beam was focused in the sample plane. The detector used was a Dectris Pilatus 6M-F area detector. Sample–detector distance was chosen to provide a maximum resolution of  $2\ \text{\AA}$  at the edge of the detector. Distance for each energy used is shown in table 4.2. A capillary beamstop was mounted between the sample and the detector. An on-axis optical microscope was used to position the chips.

Diffraction data was collected at 12 keV, 18 keV, 20 keV, 22 keV, 24 keV and 26 keV, with the

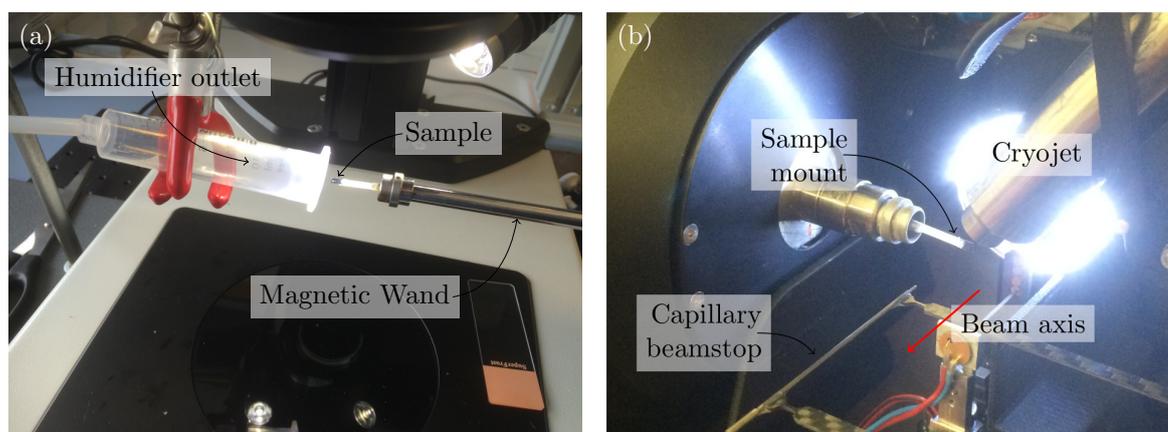


Figure 4.6: **Sample loading and mounting at the P11 beamline at PETRA-III.** (a) The chip-loading apparatus: a humidifier and sample holder were set up under an optical microscope. The humidifier maintained the sample in a hydrated environment while excess solvent was removed prior to flash-cooling in liquid nitrogen. (b) The configuration of the beamline with the sample mounted on the goniometer and the cryojet aligned. The red arrow indicates the direction of the x-ray beam.

beam parameters shown in table 4.2. Beam profiles were measured with a YAG (Cesium(III)-doped Yttrium Aluminium Garnet crystal). Flux was measured with a  $100\ \mu\text{m}$  thick silicon diode. At each energy, one or more chips were used during testing and troubleshooting, e.g. confirming appropriate exposure times and scan protocols, and dealing with any technical issues that arose when the beam energy was changed. Once the data-collection protocol was optimised for that energy, a fresh chip was loaded and one or more automated raster-scans were set up as detailed below.

With the optical microscope it was possible to identify areas of the chip that were bare vs areas that contained sample. During testing the chips were positioned so that the beam illuminated these sections while taking exposures of varying lengths. For data collection, the goniometer was programmed to translate the chip so that the beam was raster-scanned across large areas of the chip that contained sample. The spacing of the exposure positions in the raster scan was  $20\ \mu\text{m}$  for measurements at 12 keV. This was increased to  $25\ \mu\text{m}$  for exposures at 18 keV and 20 keV, and to  $30\ \mu\text{m}$  for measurements at 22 keV and above. This spacing of exposure positions was chosen to ensure that PEs produced in earlier exposures did not result in damage to crystals exposed later in the scan.

At each point in the raster scan, a number (between 50 and 250) of successive exposures was taken without rotating the sample between exposures. For beam energies above 12 keV, the

Table 4.2: **Beam parameters for the PETRA-III experiment.** Beam profiles were measured with a YAG (Cesium(III)-doped Yttrium Aluminium Garnet crystal). Flux was measured with a 100  $\mu\text{m}$  thick silicon diode at the sample position.

Energy [keV]	Beam Size [ $\mu\text{m}$ ] (H $\times$ V)	Beam Flux [photons $\text{s}^{-1}$ ]	Detector Distance [mm]	Exposure Time [s]	Exposures / Position
12	$21.3 \times 5.7$	$7.70 \times 10^{12}$	368.1	0.04	60
18	$9.8 \times 5.4$	$3.70 \times 10^{12}$	587.8	0.14	100
20	$9.5 \times 6.1$	$1.63 \times 10^{12}$	659.1	0.40	100
22	$15.9 \times 12.8$	$1.24 \times 10^{12}$	729.7	0.60	200
24	$9.1 \times 6.4$	$1.70 \times 10^{11}$	800.1	10.00	50
26	$9.8 \times 5.8$	$3.50 \times 10^{11}$	870.1	2.40	250

scaling of diffracted intensity and energy deposition was estimated based on the change in cross-sections for elastic scattering, inelastic scattering, and photoionisation. The exposure times were selected to obtain roughly equal elastically scattered flux *per frame*. The number of frames was chosen to approximately balance the total energy absorbed by the crystal over the course of *the whole time series*. Exposure times and the number of exposures per position at each energy are shown in table 4.2. The dose estimate did not account for PE escape. One exception to this approach was data collected at 24 keV; here the low flux necessitated a much longer exposure per frame and the number of frames collected was reduced to keep the total scan time reasonable.

The number of positions at which we recorded diffraction data varied with energy, with a minimum of 32 positions at 24 keV and a maximum of 3068 positions at 18 keV. The main factor affecting how many positions were measured at a given energy was the collection time per position, which varied between 2.4 s at 12 keV and 10 min at 26 keV.

### Data Analysis

Analysis of our data was performed in part on local hardware and in part by remote access to The DESY MAXWELL computing cluster. The former was used for manual graphical analysis, the latter for batch processing of the data. The following protocol was used:

1. On a local machine, the Dectris image viewing software ALBULA [131] was used to manually inspect diffraction patterns collected at 12 keV, in order to identify chip positions that contained well-diffracting crystals.

2. On MAXWELL, peak-finding trials were performed using *indexmajig*, part of the CrystFEL [132] suite for serial crystallography, using the same 12 keV diffraction data that we had manually reviewed in step 1.
3. The peak counts obtained with *indexmajig* for the 12 keV patterns were compared with the list of 'good' positions identified with manual analysis, to see if positions containing well-diffracting crystals (at any energy) could be inferred from the peak count data.

The data were structured such that each chip position corresponded to a single folder, containing a number of .cbf files, each of which was a single frame/exposure at that position. In order to manually inspect a range of positions in the 12 keV data we used the BASH shell to copy the .cbf of the first exposure at each position to another folder and rename the files to be sequentially numbered by position. This allowed the images to be viewed sequentially in ALBULA.

For our *indexmajig* trials, BASH scripting was used to schedule 526 MAXWELL jobs, one for each position at which data was collected at 12 keV. Each job ran 12 iterations of *indexmajig* peak finding using the PeakFinder8 algorithm and the default integration radii: (3,4,5). Across the trials we varied the signal-to-noise ratio (2.5, 5) and the threshold (25, 50, 100, 200, 400, 800). The output of *indexmajig* (.stream) contains, among other things, a list of all coordinates in an image identified as peaks (based on the threshold, SNR, etc) and their intensities. A BASH script was used to count the peaks in each image for each set of *indexmajig* parameters and tabulate the results in a spreadsheet. In the spreadsheet, positions were marked that had well-diffracting crystals, as determined by manual analysis. Peak counts were analysed to determine if they could be used to reliably indicate the presence of well-diffracting crystals, without the need for manual viewing of diffraction patterns.

Following this process with the 12 keV data, a subset of time series was selected from among the 12 keV, 24 keV and 26 keV data sets for further manual review and analysis. The CrystFEL image viewer *hdfsee* was used to review the peak positions identified in these images. Stream files with the most accurate matches were then processed using Python scripts to sum intensities within 0.1 Å and 1 Å resolution shells. These intensities were plotted as a function of time (exposure/frame number) to check spot fading trends.

### 4.2.3 X-ray Data Collection at NSLS-II

A follow up experiment was conducted at the 17-ID-2 FMX beamline [133, 134] of the NSLS-II [135] synchrotron at Brookhaven National Laboratory (BNL) in Upton, New York. This experiment was undertaken as part of a block allocation group (BAG) proposal, in collaboration with

the beamline team, the Garman group from Oxford University, and the Perry group from UMass Amherst.

### Sample Preparation

Crystallisation trial number 20 (table 4.1) was used at NSLS-II. The crystals were from our second round of crystallisation. The crystal size was  $(2.45 \pm 0.22) \mu\text{m} \times (2.34 \pm 0.26) \mu\text{m} \times (1.24 \pm 0.17) \mu\text{m}$ .

Transparent polymer chips were used at the NSLS-II experiment, provided by the Perry lab at UMass Amherst. The chips were 102  $\mu\text{m}$  thick SU-8, with exterior dimensions of 18 mm  $\times$  2.7 mm. They were patterned with a grid of 100  $\mu\text{m}$  square wells; the base of each well had 10  $\mu\text{m}$  apertures (for drawing off solvent), spaced 30  $\mu\text{m}$  apart. The design of the chips is shown in fig. 4.7. The aspect ratio of the chips made it necessary to re-enforce them to keep them rigid within the cryostream; this was accomplished by epoxying wire supports to the back side of the chips, as seen in fig. 4.8. Chips were seated in magnetic bases that could be attached to a standard macrocrystallography goniometer (fig. 4.7a).

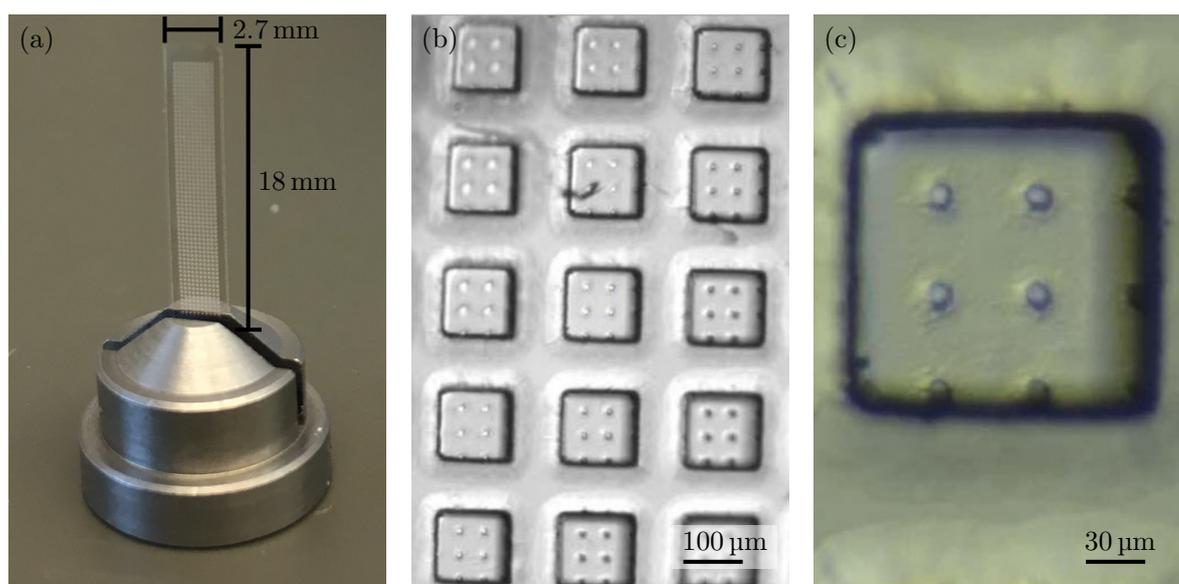


Figure 4.7: **Polymer chips for sample delivery at NSLS-II experiment.** The chips were 102  $\mu\text{m}$  thick SU-8 with overall dimensions of 18 mm  $\times$  2.7 mm. They could be fixed to standard goniometer mounts as shown in (a). (b) and (c) show optical microscope images of the chips, highlighting the sample wells and solvent drainage channels. The well profile is square with a nominal width of 100  $\mu\text{m}$ ; drainage channels have a nominal diameter of 10  $\mu\text{m}$  and spacing of 30  $\mu\text{m}$ .

The crystals were loaded onto the chips in the same manner as the silicon chips used at PETRA-III: the sample was diluted at a 1:3 ratio with the original precipitant solution and vortexed to evenly distribute the crystals; a 2  $\mu\text{L}$  aliquot was deposited on the chip; blotter paper was applied to the reverse side to draw off the solvent; and the sample was flash-cooled in liquid nitrogen. The mounting process was accomplished with the aid of a humidifier and an optical microscope.

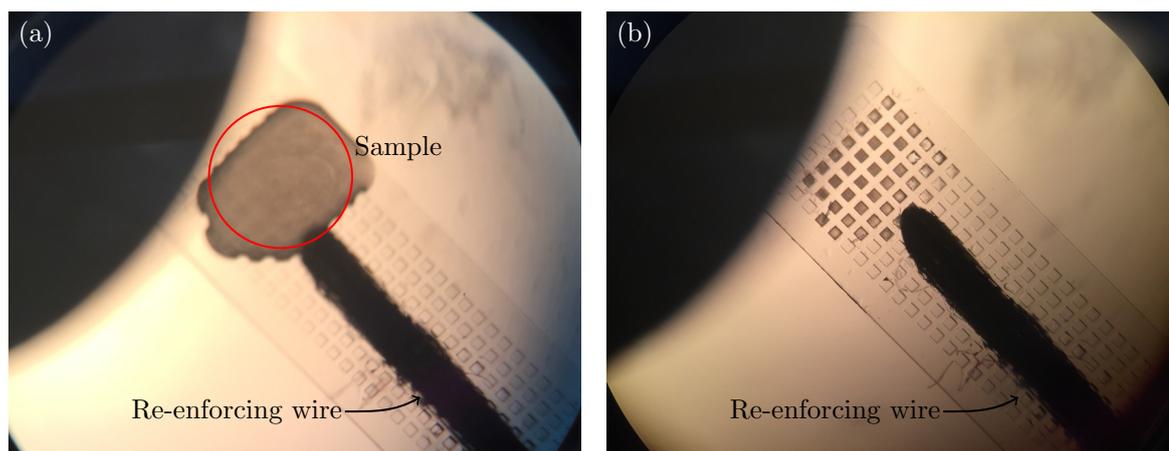


Figure 4.8: **Sample loading at NSLS-II using the Perry lab's polymer chips.** (a) shows the chip immediately after 2  $\mu\text{L}$  of sample had been applied. (b) The dark squares show the crystals sitting within the wells after excess solvent had been drained away. The dark line in the centre of the chip is a wire that was epoxied to the rear of the chip to increase rigidity.

### Beamline Configuration and Data Collection

The samples were mounted on the beamline goniometer as shown in fig. 4.9, where they were continually cooled to 100 K by a stream of gaseous nitrogen. The beam was focused in the sample plane. For the duration of the experiment, the FMX beamline had an Eiger 1M CdTe on loan from Dectris installed especially for our use. An on-axis optical camera was used for sample positioning.

Diffraction was measured at 12.66 keV and 26.00 keV with the beam parameters shown in table 4.3. The flux was measured with a calibrated PIN silicon diode provided by the Garman group [136]. Beam profiles were characterised by scanning a 30 nm chromium wire on a silicon substrate across the beam and measuring fluorescence. At 12.66 keV, the beam was attenuated with aluminium plates. Exposure times and attenuation were chosen to roughly match the expected elastic yield per frame at the two energies. The initial diffraction measurements were made at

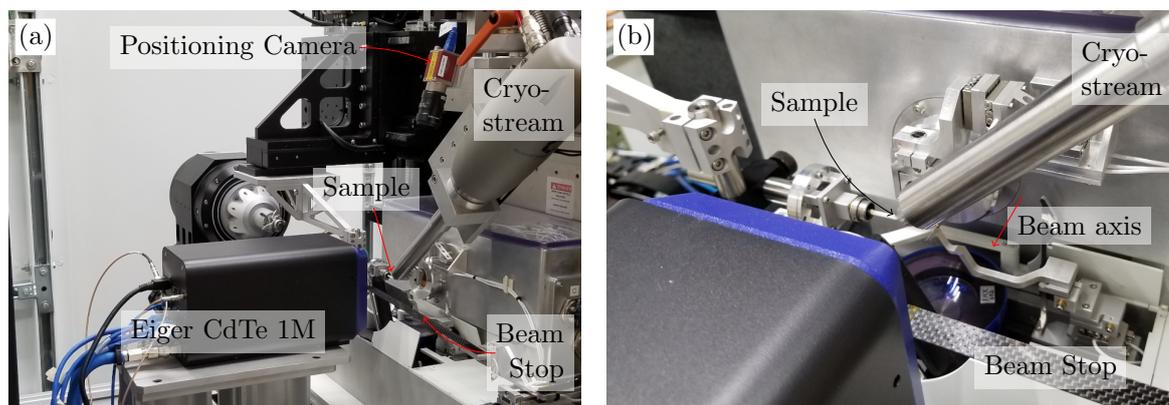


Figure 4.9: **Sample mounting at the NSLS-II FMX beamline.** A wide view of the beamline set up is shown in (a), highlighting all key components of the apparatus. (b) provides a closer view of the sample position. The red arrow indicates the direction of the x-ray beam.

12.66 keV to establish a workflow. Data were then collected at 26.00 keV before changing energy back to 12.66 keV for further data collection.

Table 4.3: **Beamline parameters for data collection at NSLS-II 17-ID-2 FMX.** Beam profiles were measured by scanning a 30 nm chromium wire, on a silicon substrate, across the beam and measuring fluorescence. Flux was measured with a calibrated silicon PIN diode provided by the Garman group [136]. The degree of attenuation of the 12.66 keV beam, and the exposure times, were selected to roughly equalise the expected elastic scattering per frame at each energy.

Beam Energy	12.66 keV	26.00 keV
Wavelength	0.979 Å	0.477 Å
Beam Size	1.35 μm × 2.9 μm	1.3 μm × 3.3 μm
Beam Flux	$1.96 \times 10^{12} \text{ s}^{-1}$	$8.80 \times 10^{10} \text{ s}^{-1}$
Transmission	0.3	1.0
Flux at Sample	$5.88 \times 10^{11} \text{ s}^{-1}$	$8.80 \times 10^{10} \text{ s}^{-1}$
Exposure Time	0.04 s	0.4 s

Low-intensity raster scans were performed across small regions of the chips to identify positions most likely to yield strong diffraction. At each position selected, a time series of 250 exposures was recorded, with a rotation of  $0.002^\circ$  between each frame (the minimum rotation permitted by the automation software). We then moved to the next chosen position. Only a few positions were selected from each raster scan. The scan-and-measure process was then repeated on a new region of the chip. If few positions with good diffraction were detected in a scan region,

this area was excluded from the data collection and a new region immediately scanned. The Dectris image-viewing application ALBULA was used to observe diffraction data as they were available, in order to estimate which time series would be suitable for analysis. We collected 21 time series at 26.00 keV and 14 time series at 12.66 keV.

### Data Analysis

In contrast to the experiment at PETRA-III P11, we had a small number of data sets to deal with after the experiment at NSLS-II 17-ID-2 FMX, and all analysis was conducted on a single workstation. `indexamajig` from the CrystFEL suite was used to perform peak finding and integration, and ALBULA and `hdfsee` to manually inspect diffraction images. Additional processing was performed using scripts written with Python. Our data were contained in HDF5 files, with two files per time series: a data file containing all 250 sequential images and a master file containing header information and pointers to the data file. The analysis protocol was as follows:

1. The peak positions were identified in the first frame of each time series.
2. Intensities were measured at these positions in all the frames in the time series.
3. The dose rate per frame was calculated using RADDPOSE-3D [66, 67].
4. The intensity was plotted against dose.

The peak positions for each time series were identified first and fixed so that the same peaks were being measured in each frame, even when they fell below the original threshold for peak detection. Two new HDF5 data files were created, each of which contained the first frame of each time series at one energy, with corresponding master files. These files were used to trial peak finding parameters with `indexamajig`. The parameters varied were peak finding algorithm (peakfinder8, zaef), SNR (2.5, 4, 6), threshold (10, 25, 50, 75, 100), and gradient (75, 100, 125). Peak positions were compared manually to diffraction images and selected peak positions were written to time series master files.

A Python script, using the `h5py` library, was used to write the peak positions identified in the first frame of each time series to the master file for the corresponding series. `Indexamajig` was then run on each time series. In these runs, peak-finding was skipped, instead the peak positions identified in the master file were used. i.e. this process measured *in each frame of the time series* the intensity of diffraction at the peak positions identified in the first frame of the series. In this process peaks were re-centred prior to integration, so small displacement of the peaks (a few pixels) did not affect the measurements adversely.

Post-processing of the stream files produced by `indexmajig` was carried out with scripts written in Python. The scripts read the peak lists from the various stream files into a single pandas [137, 138] dataframe (a 2-dimensional, indexable array which permits heterogeneous data and complex arithmetic and logical operations upon the data), with each peak tagged by various parameters—a dataset identifier, the beam energy at which that dataset was captured, a frame number, resolution, etc—that would facilitate easy sorting and visualisation. The intensity data in this dataframe was quantised, in the time dimension, by frame number. However, the exposure conditions at the two energies produced a different dose per frame so the intensity data had to be interpolated onto a dose axis rather than a frame axis. This was also accomplished by Python scripting, using dose rates calculated with RADDPOSE-3D [66, 67], version 4. This version of RADDPOSE-3D accepts PDB references to define sample parameters; the PDB entry 6FTR [125] was used as these crystals were prepared using the same protocol as was used in this experiment. The effects of PE escape were ignored and it was assumed that the crystal was centred in the beam. In order to account for uncertainty in crystal size, calculations were repeated with three different crystal sizes: the mean and one standard deviation either side, based on the distribution of sizes observed with the TEM (see section 4.2.1). For each crystal size, three orientations of the crystal in the beam were modelled where one side of the crystal was normal to the beam axis. Each execution of RADDPOSE-3D included 10 runs of the Monte Carlo simulation of photon-sample interactions [139], each of which modelled  $1 \times 10^6$  photons. i.e. at each energy 90 individual runs were performed. The average was calculated across all 90 runs at each energy to obtain dose estimates of  $(37.15 \pm 6.63) \text{ MGy s}^{-1}$  (or  $(1.49 \pm 0.27) \text{ MGy per frame}$ ) at 12.66 keV and  $(1.1 \pm 0.3) \text{ MGy s}^{-1}$  (or  $(0.43 \pm 0.10) \text{ MGy per frame}$ ) at 26.00 keV, with the error being the standard deviation.

The measured peak intensities were normalised by fitting an unconstrained exponential curve with the equation  $I = a \exp(-bD)$ , where  $I$  was the measured intensity,  $D$  the calculated dose and  $a$  and  $b$  were fitting coefficients; and dividing all intensity measurements by  $a$ . This fitting process was performed for each peak position.

Based on a visual inspection of peak fading, 10 peaks were selected from each energy that showed the least deviations from exponential fading, with resolutions between 3.1 Å and 4 Å. The intensity loss of these peaks was then compared by plotting against dose.

## 4.3 Results

### 4.3.1 PETRA-III Results

We collected diffraction data at 12 keV, 18 keV, 20 keV, 22 keV, 24 keV and 26 keV. Automated raster scans were used to collect 4016 time series across these six energies. Intensity was plotted against time for resolution shells with widths of 0.1 Å and 1 Å, and some time series were manually inspected. Intensity plots showed very unstable changes in intensity, with no exponential trends present. Sudden changes were detected in diffraction patterns from one exposure to the next, such as the disappearance of previously bright peaks and/or the appearance of new peaks. This body of results was inconsistent with the assumption that the same crystals were illuminated, at the same angle, for the duration of our time series.

An example of loss of diffracted energy over time is shown in fig. 4.10. Summed intensity of three resolution shells is shown for data collected at 24 keV. Other resolution shells were excluded for clarity but the shells displayed are indicative of trends at all resolutions. There is no indication of exponential intensity loss. The majority of the data shows a sharp drop to below 40% of the initial intensity, followed by oscillation and/or a subsequent rise in intensity.

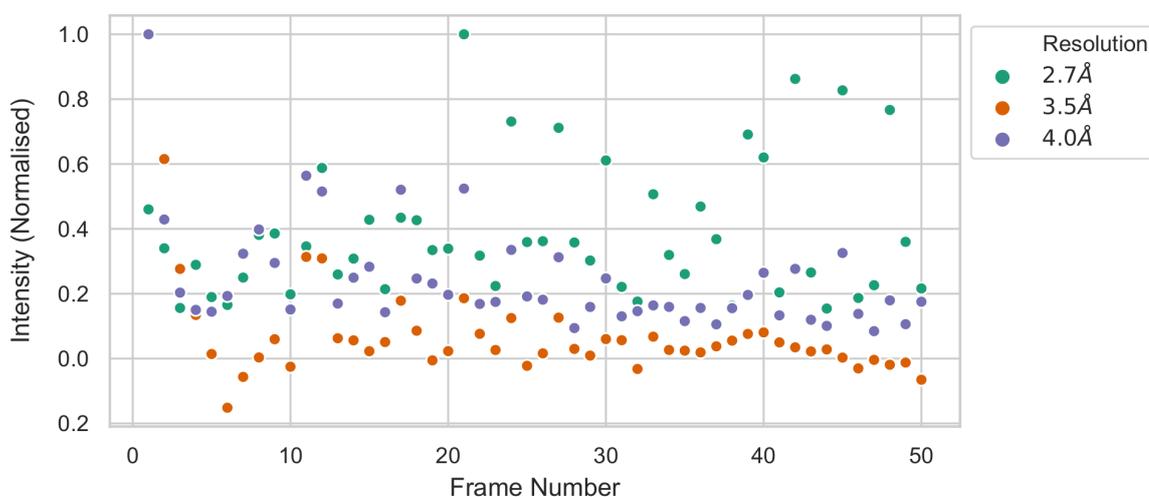


Figure 4.10: **Plot of peak intensities as a function of time.** Diffraction data were collected at 24 keV at PETRA-III. Each frame/exposure was 10 s. For clarity, only three resolution shells are shown. Intensity values have been normalised by dividing by the maximum intensity value for the shell. No exponential trend is visible in the data at any resolution.

Figure 4.11 shows an example of the diffraction patterns collected during the experiment. A sudden change in the distribution of Bragg peaks was observed between consecutive images.

Two consecutive exposures from a time series taken at 26 keV are shown. Some of the same peaks are visible in both exposures, for example in the upper-middle panel. However many of the peaks are different from one exposure to the next, which is most visible in the lower two middle panels. The different peak positions seen in the two images indicate a difference in the Bragg condition from one frame to the next.

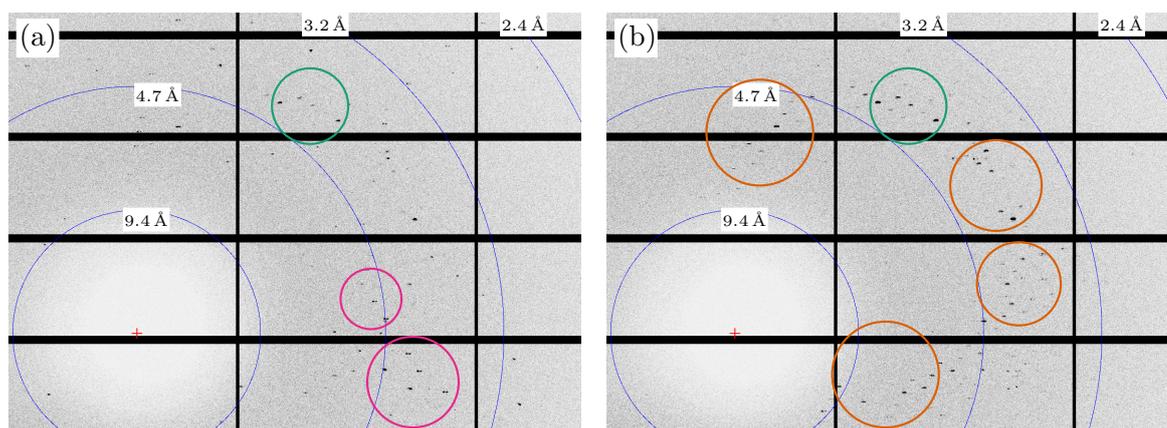


Figure 4.11: **PETRA-III diffraction images showing a variation in peak positions in two consecutive images.** The two images are successive exposures (images 20 and 21 of 250) at 26 keV. Some peaks are visible in both exposures (circled in teal). Several peaks visible in the first image are not visible in the second (circled in pink), and more peaks, not visible in the first image, have appeared in the second (circled in orange). The presence of very different peaks in the two images suggests an instability in either the beam or the sample, or both.

### 4.3.2 NSLS-II Results

Each data set was a time series of 250 exposures with a rotation of  $0.002^\circ$  between each exposure. 14 time series were collected at 12.66 keV and 21 time series were collected at 26.00 keV. Examples of the diffraction data collected are shown in fig. 4.12. These images have been cropped to improve the visibility of the Bragg peaks. Diffraction was stronger at 12.66 keV compared to 26.00 keV by a factor 1.72. Estimates of elastic scatter made with RADDOSSE-3D [66, 67] were  $2.68 \times 10^6$  photons per frame at 12.66 keV and  $1.56 \times 10^6$  photons per frame at 26.00 keV. Dose rate (also estimated with RADDOSSE-3D) per frame was greater at 12.66 keV by a factor of 3.41 (1.76 MGy per frame at 12.66 keV compared to 0.52 MGy per frame at 26.00 keV). 10 time series at 12.66 keV and 15 time series at 26.00 keV were used for analysis. The other time series did not contain strong diffraction and therefore were excluded from the analysis. The unit cell was not determined experimentally (at either energy) as there was insufficient data for reliable indexing.

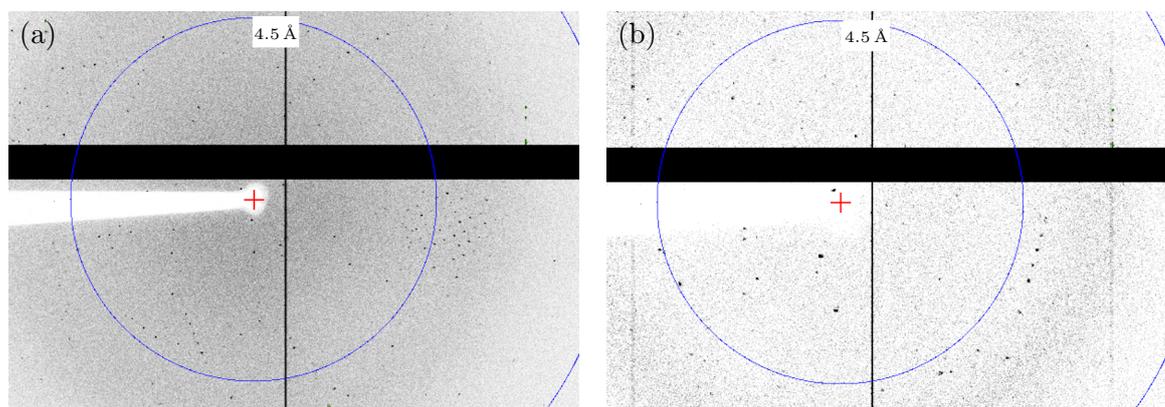


Figure 4.12: **An example of the diffraction data collected at NSLS-II at two different beam energies.** (a) Diffraction image from 12.66 keV and (b) diffraction image from 26.00 keV. Both images are the first exposure of their respective time series. Bragg peaks at 26.00 keV are larger and sparser compared to peaks detected at 12.66 keV.

From the time series analysed, the 10 best individual peaks at each energy were isolated. The criteria for selecting peaks were:

1. Spot intensity.
2. Exponential intensity decay.
3. Higher resolution peaks were prioritised.

The resolutions (between 3.1 Å and 4.1 Å) and spot half doses of the the peaks are shown in table 4.4. The loss of peak intensity vs dose is plotted in fig. 4.13, which shows the mean, with 95% confidence interval, of the fading of the 20 peaks, whose intensities have been normalised. The 26.00 keV intensities faded more slowly compared to the 12.66 keV intensities. Also shown are three trend lines showing predicted intensity loss of a 3.5 Å Bragg peak, based on an empirical spot fading model (incorporating components of the model presented in Chapter 3 and RADDOSE-3D v4). The predictions include one that ignores PE escape (the same at both energies) and two that include PE escape, one for each energy. All three trend lines show slower fading than observed in our data. The two trend lines that account for PE escape show slower fading by an order of magnitude than what was observed in our data. The difference between the two energies in the predictions that include PE escape is similar to the difference between the two energies in our experimental data.

The half dose for each of the 20 peaks was extracted from the exponential decay curves fitted to

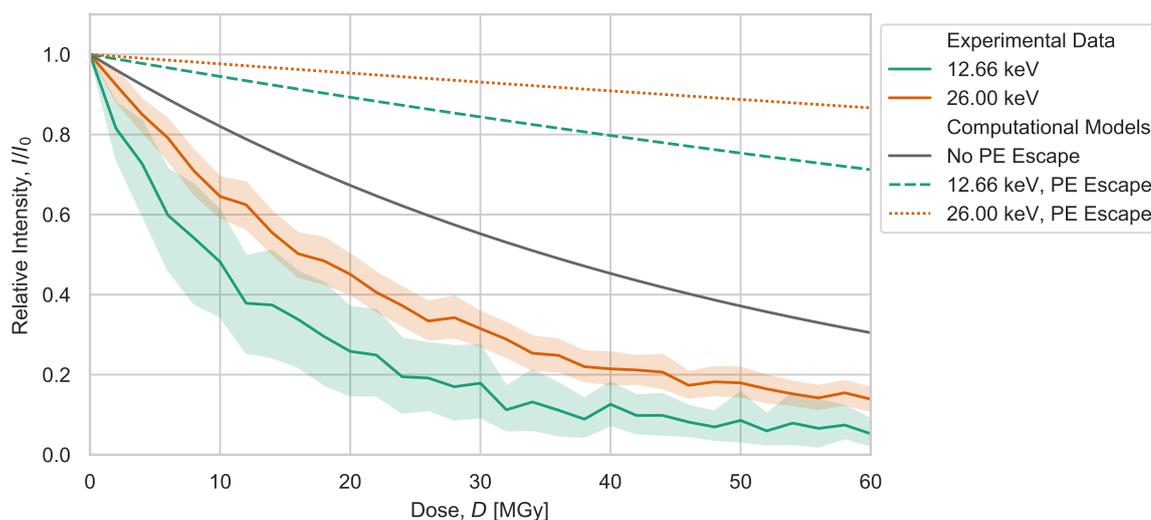


Figure 4.13: **Relative intensity loss of 10 Bragg peaks at 12.66 keV and 10 Bragg peaks at 26.00 keV compared to predictions based on radiation damage models.** Plots of experimental results show the mean of the 10 peaks (dark teal (12.66 keV) and orange (26.00 keV) lines) with a 95% confidence interval (shaded areas). Slower fading of Bragg peaks is evident at 26.00 keV relative to 12.66 keV. Three trend lines show the predictions of our spot fading model under different conditions: the solid grey line is for both energies with PE escape not included in dose calculation, dashed teal line is for 12.66 keV with PE escape included in dose calculations, and the dotted orange line is for 26.00 keV with PE escape included in dose calculations. All trend lines predict slower fading of spots than observed in the experimental data. The relative difference between 12.66 keV and 26.00 keV is similar for the experimental data and the computational model.

the peak intensities. These values are shown in table 4.4. We determined half dose using dose rate estimates that both included and excluded PE escape. The numbers in parentheses in table 4.4 are the spot half dose divided by the spot resolution (in Å). These values are provided to facilitate a comparison between the peaks of different resolutions as spot-fading is resolution-dependent and therefore it is not meaningful to compare raw spot half doses for peaks with different resolutions. Among the 12.66 keV spots there is no discernible trend in the relationship between spot resolution and half dose, and there is significant noise, e.g. the four 3.6 Å peaks have half doses (excluding PE escape) of 15 MGy, 18 MGy, 4 MGy and 22 MGy. Among the 26.00 keV peaks there is also noise, but there exists some correlation between peak resolution and half dose, with lower resolution peaks tending to have higher half doses.

We took averages of the half dose weighted by resolution (i.e. the half doses are given in  $\text{MGy} \text{Å}^{-1}$ ) as the half dose is expected to be resolution dependent and therefore averaging half doses of different resolution peaks without weighting is not meaningful. When PE escape was excluded from dose rate estimates, the average half dose (per Å) of the 26.00 keV peaks was 1.7

Table 4.4: **Half doses ( $D_{1/2}$ ) for 10 peaks at 12.66 keV & 10 peaks at 26.00 keV.** For each peak, the dose at which spot intensity had faded to half its initial value was determined by fitting the experimental data to an exponential decay curve. The half dose per angstrom of the spot resolution is shown in brackets as a way to standardise values across peaks at different resolutions. For each energy, the half dose was determined based on two dose rates calculated by RADDOSE-3D, one of which accounts for PE escape, the other of which does not.

12.66 keV Peaks			26.00 keV Peaks		
$d$ [Å]	$D_{1/2}$ [MGy] (Å <sup>-1</sup> )		$d$ [Å]	$D_{1/2}$ [MGy] (Å <sup>-1</sup> )	
	No PE	PE		No PE	PE
3.2	8 (2.5)	2.3 (0.7)	3.1	17 (5.5)	2.0 (0.7)
3.2	10 (3.1)	2.9 (0.9)	3.2	13 (4.1)	1.6 (0.5)
3.4	6 (1.8)	1.7 (0.5)	3.3	17 (5.2)	2.0 (0.6)
3.5	9 (2.6)	2.6 (0.7)	3.4	14 (4.1)	1.7 (0.5)
3.6	15 (4.2)	4.3 (1.2)	3.5	29 (8.3)	3.5 (1.0)
3.6	18 (5.0)	5.1 (1.4)	3.5	13 (3.7)	1.6 (0.4)
3.6	4 (1.1)	1.1 (0.3)	3.6	15 (4.2)	1.8 (0.5)
3.6	22 (6.1)	6.3 (1.7)	3.9	24 (6.2)	2.9 (0.7)
3.9	2 (0.5)	0.6 (0.1)	4.0	21 (5.3)	2.5 (0.6)
3.9	16 (4.1)	4.6 (1.2)	4.1	22 (5.4)	2.7 (0.6)
<b>Av.</b>	<b>(3.1)</b>	<b>(0.9)</b>	<b>Av.</b>	<b>(5.2)</b>	<b>(0.6)</b>

times the average of the 12.66 keV peaks. When PE escape was included in the dose estimates, the average of the 12.66 keV peaks was slightly higher than the 26.00 keV peaks.

## 4.4 Discussion

In this section, we initially provide some context for the discussion (4.4.1) and describe the conception and planning of the experiments (4.4.2). The outcomes of the PETRA-III experiment (4.4.3) are then explored, focusing on the key issues encountered and their implication for future similar experiments. We highlight the conclusions of the analysis of our NSLS-II data (4.4.4) before exploring the discrepancy between the results of our NSLS-II experiment and the predictions of the spot fading model used (4.4.5), and the difference observed in spot fading at the two energies (4.4.6). The discussion is concluded by highlighting the key challenges facing future experiments that seek to mitigate radiation damage via PE escape (4.4.7).

#### 4.4.1 Context

Theoretical studies over the past 15 years [11, 12, 104] have suggested that radiation damage during MX experiments may be partially mitigated (relative to the diffraction signal) if the crystals and beam are small enough (on the order of 1  $\mu\text{m}$ ), and if the the beam energy is increased beyond the 12 keV typical of such experiments. The mechanism proposed for this mitigation is the escape of PEs from the diffraction volume before they have deposited all their energy, leading to a reduction in dose for a given intensity of elastically scattered photons (as discussed in Chapter 3).

If such a reduction in dose could be realised experimentally, it could prove helpful for potential targets that do not grow large, well-diffracting crystals and where sample is limited. Small crystals are also of interest due to the extra information that can be obtained from their larger Bragg spots using, for example, Bragg coherent diffractive imaging (BCDI) [26–28]. A deeper understanding of the impact of radiation damage on diffraction data as a function of dose could permit the improvement of structure retrieval algorithms to account for dose, thereby extending the usable life-dose of the crystals.

Until recently, the kind of experimental parameters required to realise a reduction in damage due to PE escape have been beyond the routine operational parameters of most MX beamlines or, indeed, most beamlines of any kind. Now, 4<sup>th</sup> generation synchrotrons, with diffraction-limited storage rings (DLSRs) are beginning to come on line [39, 40, 115, 116], offering the potential to meet the beam size limitations required. Coupled with new cadmium-telluride (CdTe) detectors that maintain sensitivity to much higher energies than the silicon (Si) detectors in common use [19, 123], improvements in synchrotron storage rings mean that many MX beamlines around the world may soon be able to routinely operate in the regime necessary to take advantage of PE escape for damage mitigation. Accordingly, developing a detailed understanding of the experimental requirements and limitations, and the kind of results that can be expected is now a very practical concern.

Computational simulations, such as the one presented in Chapter 3 and RADDPOSE-3D [66, 67], have attempted to estimate the reduction in dose under a range of parameters, in order to inform experimental design. The predictions made by these simulations can be helpful but one should be cautious about expecting too much from them in quantifiable terms. Given the complexity of the interlocking systems that make up a modern MX experiment, and the necessary (relative) simplicity of computational models, it is perhaps not reasonable to expect precision in their predictions, no matter how much they are refined. However, having a systematically

collected body of experimental results to compare to predictions certainly has utility in terms of understanding *how much* we expect actual results to differ from the models.

#### 4.4.2 Experimental Design

The goal of a typical MX experiment is to obtain sufficient diffraction data from the crystal(s) available in order to solve the structure of the protein in question. This involves recording diffraction from the crystal (or many micro-crystals) at multiple angles and combining the data using computational techniques. The fundamental datum of an MX experiment then is a single diffraction image; the combination of many of these images, with varying crystal orientation, comprises a complete data set. Our experiment shares with MX a fundamental datum, but has a different data set, as we are not attempting to solve the structure, but to track the damage state of the crystal—via the decaying intensity of Bragg peaks—over time. As such, specific crystal orientation is not critical. What is important is that we capture multiple exposures of *the same Bragg peaks* of the *same crystal*, i.e. the angle should remain the same for a given series of exposures of a given crystal, but any relative difference in angle between crystals is immaterial.

Previously, Coughlan *et al.* [26] tracked Bragg peak fading as a metric for radiation damage using Bragg coherent diffractive imaging (BCDI) to examine a single over-sampled Bragg peak. They used the additional data obtained by oversampling to reconstruct an image of the crystal in real space. As our approach in simulation was to track the loss in Bragg peak intensity, and not to consider the real-space evolution of the crystal, our experiment could be simplified and we could examine multiple peaks within a given diffraction pattern.

In Chapter 3, we presented a computational simulation of radiation damage that modeled the loss in Bragg peak intensity as a metric for tracking the impact of radiation damage on a crystal. Our simulation predicted a much slower rate of intensity loss for small crystals in small beams (both crystal and beam FWHM  $\leq 2 \mu\text{m}$ ), with the effect being more pronounced for higher energy beams. Similar conclusions have been drawn, in varying degrees of detail, by previous models [11, 12]. Our experiment was therefore designed to produce data that can be compared to our model: sequential diffraction patterns taken under variation of at least one of the three key parameters we examined in simulation: crystal size, beam size, and beam energy. We elected to focus on beam energy, as the increased penetration depth of higher energy PEs offered the best potential for observing damage reduction due to PE escape.

#### Beamline Requirements & Configuration

In order to realise our experiment, we required a beamline with several capabilities:

- An energy range that included 12 keV and extended to at least 26 keV, with the ability to easily change energy;
- The ability to focus down to 1  $\mu\text{m}$ –2  $\mu\text{m}$ ; a small beam is required to minimise illumination of the solvent surrounding the crystal, as PEs generated in the solvent can be scattered back into the crystal, off-setting the reduction in dose due to PE escape;
- High flux that could be maintained, as much as is possible, at higher x-ray energies;
- Facilities for the preparation of biological samples;
- An area detector capable of maintaining efficiency at higher beam energies.

Although the majority of these requirements are met individually by many MX beamlines at facilities around the world, beamlines that meet *all* these requirements are still rather rare. In particular, the combination of energy range and spot-size requirements seemed particularly uncommon in our initial survey. Perhaps the hardest requirement to meet was access to a suitably efficient detector. After a wide survey, we identified the P11 beamline at PETRA-III at DESY in Hamburg, Germany, as a good candidate. Later the FMX beamline at NSLS-II, BNL, New York, USA, was chosen for a follow-up experiment. One of the criteria that made FMX an appealing option was the availability of a CdTe area detector, providing improved sensitivity at higher beam energies.

### Data Collection Strategy

Given the data structure required and the samples available, the basic outline of our experiment was to:

1. Measure diffraction with an area-detector, without rotating the sample (as a rotation of any significance would bring new peaks into the Bragg condition, preventing us from seeing the behaviour of given peaks over time);
2. Obtain data from many individual crystals (i.e. employing a serial MX approach) in order to build a statistically robust set of results;
3. Mount our crystals on chips so they were statically positioned and multiple exposures could be taken of a single crystal (at a given position on the chip);
4. Repeat the experiment, with crystals from the same batch (i.e. they should be as similar as possible), at various beam energies including 12 keV and as many other energies as feasible up to around 30 keV.

As detailed in section 4.2, precisely how this general strategy was realised varied based on available tools and expertise, but the basic approach was the same for both iterations of the experiment.

#### 4.4.3 PETRA-III Results: Overview

The analysis of results for the PETRA-III experiment did not progress beyond the exploratory stages. There were two reasons for this:

1. It became apparent that we were unlikely to see logical spot fading in our data, let alone any energy-dependent trends.
2. The opportunity arose to participate in the BAG at NSLS-II providing the opportunity to generate better data.

The results presented in section 4.3.1 present a snapshot of the PETRA-III analysis to that point and illustrate the impact of some of the issues faced in the data-collection and analysis process. Here we explore these issues briefly before discussing the results of the NSLS-II experiment in greater detail.

##### Beam Size

The primary issue encountered during the PETRA-III experiment was the lack of availability during the experiment of a suitably small focal spot. In our simulations (see Chapter 3), 5  $\mu\text{m}$  was an approximate cut-off point at which PE escape ceased to be beneficial (at 12 keV). This size limit applied to the crystal, if the beam was a similar size or larger, but it also applied to the beam; i.e. even a 1  $\mu\text{m}$  crystal saw no mitigation of radiation damage at 12 keV when illuminated by a 5  $\mu\text{m}$  beam. This is because PEs are also generated in the solvent around the crystal and can scatter back into the crystal, off-setting the benefit of the PEs that escape it. Our simulations presented an ideal theoretical limit, hence the real-world cut-offs were expected to be more constrained. The P11 beamline at PETRA-III was selected based on its advertised minimum beam spot of 1  $\mu\text{m} \times 1 \mu\text{m}$ . However, due to technical difficulties during beamtime, a focus spot of this size could not be achieved. As shown in table 4.2, the smallest beam size produced at PETRA-III during the experiment was approximately 9  $\mu\text{m} \times 6 \mu\text{m}$ , larger than the maximum size at which PE escape was expected to lead to a reduction in damage at 12 keV. It is possible that at higher energies some reduction in damage, attributable to PE escape, might be seen, but at higher energies the PEs produced in the solvent also have a higher chance of spreading more of their energy into the crystal and the volume illuminated by our beam made it likely this effect would be strong.

Several crystal sizes were available for use at the PETRA-III experiment, with a minimum crystal size of  $2\ \mu\text{m} \times 1.5\ \mu\text{m}$ . These would have been appropriate for use with a beam of approximately the same size. Due to the potential for scattering of PEs from outside the crystal, the largest sample available at beam time was used. Due to the available beam size, no significant difference in damage rate was expected between the larger and smaller crystals, and the larger crystals were deemed more likely to provide stronger diffraction and better SNR.

### **Beam Stability**

The different peaks present in the two consecutive exposures shown in fig. 4.11 indicate that the Bragg condition was different for the two images. Some of the Bragg peaks are in a similar position in both images but there were many peaks in the second image that were not visible in the first, and several peaks visible in the first image were missing in the second. Possible explanations for these observations are: a single crystal was illuminated in both images but the angle of incidence between the crystal and the beam was different from one image to the next; or multiple crystals were illuminated in some combination across the two images. In the first case, an unintended rotation may have brought some Bragg peaks into at least partial intersection with the Ewald sphere, while others were shifted out of such intersection. This would require movement of the crystal on the chip, or movement of the chip/goniometer itself. A rotation of the cryofixed crystal on the chip is unlikely, which leaves the possibility of the sample holder itself moving slightly. By default, the software that controls sample rotation requires that the goniometer be rotated between exposures. In order to prevent this, the technical support staff at PETRA-III implemented a custom script to force zero rotation. This was written during beam-time and untested. It is possible that some rotational instability in the sample mount resulted from this. In the second case, if multiple crystals were exposed due to a slight displacement of the beam, then overlapping Bragg peaks should have been evident in the diffraction patterns. Visual analysis of our images did not indicate that multiple-crystal diffraction was likely. During the experimental set up, when the beam energy was changed and the beam refocused, some issues were observed with beam stability. Hence, vibration or defocusing of the beam, and/or rotational vibration of the sample holder are deemed the most likely explanation for the effect observed in fig. 4.11.

### **Uncertainty in Beam Parameters**

Changing energies at the PETRA III MX beamline proved to be quite challenging. Loss of flux at higher energies was greater than predicted based on previous beamline tests, and the beam was observed to be unstable. In addition, the obtainable flux at the experiment was observed to

be lower than expected and at one point dropped to just 10% of the value given in the beamline parameters. Another issue which was noted during the experiment was that the YAG measurements were impacted by the cryostream meaning that the beam profiles could not be recorded accurately. This had implications for the accuracy of dose calculations and by extension the accuracy of the dose axis on which intensity data would be plotted.

### **Detector Efficiency**

The quantum efficiency (QE) of silicon detectors deteriorates as a function of increasing x-ray energy, within the range we were investigating. As discussed in section 4.1, the QE of even the thickest available silicon detector is below 50% above 22.1 keV; this presents additional challenges in terms of achieving sufficient SNR to characterise the influence of PE escape. A CdTe detector was not available for use during the PETRA-III experiment.

Ultimately the technical issues with the beamline and lack of access to a CdTe detector prevented the collection of high-quality radiation damage data. However, the protocols and experience gained during this experiment guided the design of the follow up experiment at NSLS-II.

#### **4.4.4 NSLS-II Results: Overview**

In contrast to the PETRA III experiment, the results from the NSLS-II experiment, as shown in fig. 4.13, are consistent with a reduction in damage at higher beam energy, due to PE escape. This supports the idea that it is useful to continue investigating the feasibility of performing MX with micro-crystals at higher than normal beam energies, ideally around 26 keV. Figure 4.13 also showed a large difference between our results and spot fading predictions based on the simulation presented in Chapter 3. This is not due the Bragg peaks in our diffraction patterns fading uncharacteristically quickly, as a comparison to the literature shows that our results are within normal bounds. Rather, the dose tolerance assumed in our simulation, based on previous estimates, was found not to be realistic. In section 4.4.5 we explore the discrepancy between our experimental results and simulations, with reference to the literature. We conclude that the simulations provide a useful guide, but that the lack of inclusion of specific experimental parameters in the model, limit its accuracy in reproducing the NSLS-II results. In section 4.4.6, we compare the metrics that were observed to similar results and finally to a similar study conducted around the same time as this experiment.

#### 4.4.5 NSLS-II Results: Discrepancy Between Experimental Intensities and Computational Predictions

The Bragg spots in our diffraction data faded more quickly than the predictions generated by computer modelling (see fig. 4.13). Simulations were run that both included and excluded the effects of PE escape. The *relative* difference between experimental results at 12.66 keV and 26.00 keV, and simulations that include PE escape were found to be the same. If the dose were somehow underestimated by an order of magnitude, the experimental results and the simulated results would match each other very well. Below we discuss whether such a margin of error for the calculated dose is realistic and what the possible explanations for this discrepancy could be.

##### Potential For Miscalculation of Dose

RADDOSE-3D v4.0 (the most recent java version as of August 2020), was used to estimate the dose absorbed by the sample. By default, RADDOSE-3D provides estimates both with and without PE escape taken into account. Dose estimates that accounted for PE escape were lower than those that didn't by a factor of 3.5 at 12.66 keV and a factor of 8.3 at 26.00 keV. By comparison to the simulations presented in Chapter 3, this represents a greater than anticipated dose reduction at 12.66 keV, and a less than anticipated dose reduction at 26.00 keV, though in both cases the discrepancy is within a factor of 2, indicating general agreement between the two models. Accounting for PE escape is a relatively new feature only present in the more recent versions of RADDOSE-3D and it is possible that the effect has been over-estimated in the simulations. This step uses MC techniques to simulate the motion of PEs through the sample and the resulting spread of their energy [139]. Given the stochastic nature of PE emission and subsequent motion, it is reasonable to assume that the MC approach is a valid approximation of the physical process of PE mediated damage.  $1 \times 10^6$  photons were simulated in each iteration of the MC component of the simulation, so statistical anomalies are unlikely to be the cause of substantial error. Given that the simulations that do not account for PE escape already show slower fading than observed in the experiment, any inaccuracy in simulation of PE escape can, at any rate, not account for the *entire* discrepancy shown in fig. 4.13.

RADDOSE-3D takes a large number of inputs in order to characterise the beam, the sample, and the interaction between the two. Inaccuracies in these inputs will lead to inaccurate predictions. Here, we consider several key parameters which include crystal size, sample composition, beam flux, and beam size.

Crystal size was determined using TEM, as discussed in section 4.2.1. We ran multiple iterations of RADDOSE-3D to account for both the distribution in size of our crystals and the randomness

of their orientation relative to the beam. For crystal size, we used the mean dimensions with an assumed error of plus-or-minus one standard deviation. We simulated three orientations for each crystal size; one with each crystal dimension aligned to the beam. Each iteration of RADDPOSE-3D included 10 duplicate runs, so the dose estimate at each energy was averaged over 90 individual runs representing, with strong confidence, the bounds of the sample parameters.

The chemical compositions of both the protein and the solvent will also strongly impact the photoionisation cross-section. The solvent fraction is also important in this determination. RADDPOSE-3D allows use of PDB files to determine the chemical composition of the protein and the solvent fraction. Molar quantities of heavy atoms in the solvent are provided separately. We used the 6FTR [125] PDB file, which was obtained with lysozyme crystals grown using the same protocol. Given the accuracy with which both crystal size and sample composition were known, it is unlikely that mischaracterisation of the sample in RADDPOSE-3D is a source of any significant error.

Beam flux was measured with a calibrated Si PIN diode; 12.66 keV is the standard operating energy for the beam line and any significant error in calibration would be easy to detect. Any small, undetected, errors in beam flux would not lead to large uncertainties in the dose calculation. The beam profile was obtained by scanning a 30 nm chromium (Cr) wire (on a Si substrate) across the beam and measuring the fluorescence. If the actual beam deviated from the size measured, it would likely broaden, due to either true defocussing or vibration. Broadening would result in a reduction in flux through the crystal, and is therefore incompatible with the hypothesis that the calculations have *underestimated* the flux. The measured beam size was close to the minimum achievable on the beamline and therefore it is considered unlikely that the beam was actually smaller.

For measurements at 12.66 keV, the beam was attenuated to 30% transmission. It is possible, though not likely, that there is a mismatch between the recorded attenuation amount and the actual attenuation. However, this would not affect the 26.00 keV series, which were collected at full flux, and the difference we observe between the model and the experimental data appears similar at both energies.

Given the precision with which the sample and beam parameters relevant to a RADDPOSE-3D dose estimate were known, a factor-of-two discrepancy in simulated vs actual dose is possible, but an order of magnitude error (consistent with a hypothesised underestimation of the dose) is highly unlikely. As such, uncertainty in the beam parameters was ruled out as an explanation for the difference observed between the measured rate of spot fading, and the simulated prediction.

It has been claimed [86] that the average doses (to both the whole crystal and the exposed region) produced by RADDPOSE-3D can be problematically small when dealing with non-uniform illumination, and that the maximum dose should be used as a better estimate in this case. Maximum dose is not among the metrics currently output by RADDPOSE-3D v4.0, so we repeated our dose calculations using RADDPOSE-3D v1.3. This version does not include energy deposition due to Compton scattering, however this exclusion should be negligible at 12.66 keV, as the Compton scattering cross section and average recoil energy are both very low, and is expected to result in an underestimation of dose by around 3-4% at 26.00 keV. The average doses calculated with v1.3 were within 6% of the doses calculated by v4.0. The maximum dose was higher (at both energies) by a factor of 1.6. This result suggests that, assuming the maximum dose, we may have underestimated the dose in the calculations —though, again, not to a degree that would account for the observed discrepancy between simulation and experiment.

In summary, whilst it is possible that the combination of factors discussed above could account for a portion of the discrepancy between simulation and experiment they cannot account for the whole discrepancy observed. This leaves the possibility that, due to factors as yet uncharacterised, the dose tolerance assumed in our models does not accurately describe the samples used at NSLS-II.

### Resolution-Dependent Dose Tolerance Assumed in Our Model

The spot fading model plotted in fig. 4.13 is a simplified version of the model presented in Chapter 3, incorporating the faster dose calculations of RADDPOSE-3D. It was included in our analysis in order to test the predictions made by our model against experimental data, providing the opportunity for quantitative comparison between theoretical predictions and real world experiments. It is based on an empirical spot fading model suggested by Holton [71] (eq. (2.18)). As such it makes the assumption that spot half dose is related to resolution,  $d$  by the *Howells Criterion*, i.e.  $D_{1/2} = Hd$ . This model is not universally accepted; Atakisi *et al.* [86] proposed an alternative model, which is proportional to  $d^\alpha$  rather than  $d$ , where  $\alpha$  is determined by fitting. In their meta-analysis of previous results, they determined a best fit value of  $\alpha = 1.86$ . The requirement for fitting (not only for  $\alpha$  but for an additional parameter outside the exponent) precludes an absolute prediction, hence the Atakisi model has been excluded from the comparison here. Further analysis of this experimental data, with reference to both the Howells-Holton model and the Atakisi model, is presented in Chapter 5.

### The Howells Criterion & Previous Estimates of dose tolerance

Several dose limits, based on modelling and experiments in MX and related disciplines such as cryogenic electron-microscopy (cryo EM) and x-ray diffraction microscopy (XDM), have been proposed, beyond which further data collection is unlikely to have much utility for structure solution. The figure of merit commonly used in discussion of dose tolerance is the so-called half dose ( $D_{1/2}$ ) of a diffraction pattern, which is the dose required to reduce the integrated intensity to half its initial value. It can also be instructive to look at the fading of Bragg spots in a given resolution band, or individual peaks, as we have done here, though it is more common to quote a single half dose for diffraction up to an initial resolution limit.

In 1990 Henderson [16] suggested a dose limit for protein crystals at cryogenic temperatures ( $\sim 40$  K to 100 K) corresponding to a half dose of 20 MGy. A decade later, Teng and Moffat [78, 79] observed a half dose of approximately 17 MGy for lysozyme crystals with an initial resolution limit of 1.6 Å, but recommended a dose limit of 10 MGy, based on completeness statistics for their data sets. Around the same time, Burmeister [72] measured a half dose of  $\sim 21$  MGy for myrosinase crystals diffracting to 2.0 Å. Owen *et al.* [119] measured a half dose of 43 MGy for apoferritin and holoferritin diffracting to  $\sim 2.3$  Å; they suggested a dose limit of 30 MGy, the commonly referenced *Garman limit*. More recently, Leibschner *et al.* [75] measured half doses for thaumatin crystals diffracting to various initial resolutions; for data sets with resolution limits of 1.22 Å, 2.08 Å, and 2.45 Å, they observed half doses of  $\sim 16$  MGy,  $\sim 18.5$  MGy, and  $\sim 20$  Å, respectively. The majority of these past results yield a half dose that is approximated by  $D_{1/2} = 10d_{MAX}(0)$ , where  $d_{MAX}(0)$  is the resolution limit of the undamaged crystal. In other words, these past studies suggest an alternative interpretation of the Howells-Holton model (eq. (2.18)), where the exponential fading should describe the *entire diffraction pattern* rather than a single Bragg spot. The results of Owen *et al.* [119] are a notable exception, with a much higher half dose. Based on this, the difference in fig. 4.13 between our experimental data and the model we used, is to be expected as the model should describe the fading of an entire diffraction pattern, rather than individual Bragg peaks.

Several other researchers have looked at the half doses for *individual resolution shells*, which allow a better comparison to individual peaks rather than looking at the whole diffraction pattern. A caveat is that individual peaks do not necessarily display the exponential behaviour seen in whole diffraction patterns and resolution shells [75, 76]. We have looked specifically at those peaks whose individual decays were most well matched to the exponential trend expected but, at a minimum, we would expect our results to be much noisier than results based on resolution shells. Teng and Moffat [78, 79] observed shell half doses of 12 MGy at 1.6 Å, 14 MGy at 1.9 Å, and 16 MGy at 2.2 Å. Atakisi *et al.* [86] gave individual shell half doses of 2 MGy to 3 MGy at 1 Å,

8 MGy at 2 Å, 16 MGy at 3 Å, and 30 MGy at 4 Å. The half doses shown in table 4.4 vary significantly, which is to be expected when examining individual peaks. Considering the average spot half dose weighted by spot resolution, our results of 3.1 MGy Å<sup>-1</sup> at 12.66 keV and 5.2 MGy Å<sup>-1</sup> at 26.00 keV are lower than the ~5 MGy Å<sup>-1</sup> to 7.5 MGy Å<sup>-1</sup> seen at comparable resolution ranges by Teng & Moffat, and Atakisi. If we base our estimates on maximum dose rather than average does (see section 4.4.5), our half dose values increase to ~5 MGy Å<sup>-1</sup> at 12 keV and ~8 MGy Å<sup>-1</sup> at 26.00 keV, comparable to previously observed half doses for individual resolution bands.

Based on the alignment of our results with half doses measured for individual resolution shells, and the alignment of previously recorded half doses for entire diffraction patterns with the Howells Criterion used in our model, we conclude that the bulk of the discrepancy observed between our simulated spot fading and our experimental results is due to the simulations being suited to describing the intensity loss of entire diffraction patterns, not individual Bragg peaks, as was the original intent. The qualitative similarity we see to our data is to be expected, as is the lack of quantitative similarity, as the faster fading of the highest resolution peaks results in steeper curves when considering individual peaks (or resolution shells) compared to an entire diffraction pattern.

What the comparison to previous dose tolerance studies does not elucidate is the extent to which the difference between our results at 12.66 keV and 26.00 keV suggests damage mitigation due to PE escape, which we consider in the next section.

#### 4.4.6 NSLS-II Results: The Influence of Beam Energy on Bragg Spot Fading

The discrepancy seen between our experimental data and predicted spot fading, shows a large difference in the *absolute* half doses observed vs. those predicted. It is also informative to consider the *relative* change in half dose between the two energies observed. Recent modelling by Dickerson and Garman [19] incorporates the effects of the quantum efficiency of the detector, the change in interaction cross section, and PE escape. For a 2 μm crystal their model predicts an improvement in diffraction efficiency—a metric permitting a direct comparison to an increase in half dose—between 12.4 keV and 26.00 keV by a factor of 2.9. We observed a difference of slightly more than half that predicted: 1.7.

Another way to view the alignment between our results and the relative change predicted by this model is to scale the dose axis of fig. 4.13 to incorporate PE escape. In this case, the two curves should sit more-or-less on top of each other. Such a plot is made impractical by the significant difference in the domains of the two data series when PE escape is taken into account, but it is possible to gain an indication of how the curves would differ by looking at the half doses of the

peaks. Table 4.4 summarises this half dose data for the 20 peaks used to plot fig. 4.13. When accounting for PE escape, the average half dose (adjusted for spot resolution) is not the same at 12.66 keV and 26.00 keV, rather the 26.00 keV spots now appear to lose intensity more quickly as a function of dose. However, the half dose values are more similar at the two beam energies than when PE escape is not accounted for in dose estimates, and the difference is in line with that observed in reference to the Dickerson-Garman model. We conclude that the difference in spot fading at 12.6 keV and 26.00 keV seen in fig. 4.13 is consistent with a reduction in damage due to PE escape.

### Comparison of Our Results to a Similar Study

Storm *et al.* [140] have reported on a similar experiment to ours, most notably also employing a cadmium telluride detector, the PILATUS CdTe 2M. They compared similar beam energies — 13.5 keV and 20.1 keV—, though they used larger crystals ( $5\ \mu\text{m} \times 3\ \mu\text{m}$  and  $20\ \mu\text{m} \times 8\ \mu\text{m}$ ). Both large and small crystals initially diffracted to between  $2.2\ \text{\AA}$  and  $1.6\ \text{\AA}$ . The average half doses they reported were, for the small crystals,  $\sim 40\ \text{MGy}$  at 13.5 keV and  $\sim 60\ \text{MGy}$  at 20.1 keV; for the larger crystals  $\sim 10\ \text{MGy}$  at both energies. They point to both the difference between the larger and smaller crystals, and the difference, for the  $5\ \mu\text{m}$  crystals, between the two energies as evidence of PE escape.

The half doses they reported for their  $20\ \mu\text{m}$  crystals are roughly in line with previous reports of dose tolerance, as discussed in section 4.4.5, though they fall 40-50% below the Howells Criterion. The half dose reported for the  $5\ \mu\text{m}$  crystals at 13.5 keV is much higher than most dose tolerances reported in the literature, with the exception of the results of Owen *et al.* [119], with which it aligns well. However, the crystals used by Owen *et al.* were approximately  $200\ \mu\text{m}$ , so the conditions are not comparable. Based on our simulations of radiation damage and those of Dickerson and Garman [19], there should be little-to-no reduction in damage due to PE escape for  $5\ \mu\text{m}$  crystals at 12 keV and near this energy there should be little observed difference between the dose tolerance of a  $5\ \mu\text{m}$  and a  $20\ \mu\text{m}$  crystal. This raises the question of whether some other, unidentified factor was responsible for the difference observed between the two crystal sizes at 13.5 keV. The beam sizes used by Storm *et al.* were roughly  $8\ \mu\text{m} \times 7\ \mu\text{m}$  for the smaller crystals and  $22\ \mu\text{m} \times 19\ \mu\text{m}$  for the larger crystals. It is likely, therefore, that the smaller crystals were more uniformly illuminated than the larger. It is possible that by using the average dose estimates of RADDPOSE-3D, the dose to the larger crystals has been underestimated [86], which could account for some of the discrepancy between the two crystal sizes at 13.5 keV, and the difference observed between the  $20\ \mu\text{m}$  crystals and the bulk of previous reports in the literature [72, 75, 78, 79].

The Dickerson-Garman model predicts improvement in diffraction efficiency by a factor of 1.7 between 12.4 keV and 20 keV for a 5  $\mu\text{m}$  crystal; the improvement by a factor of 1.5 observed by Storm *et al.* between 13.5 keV and 20.1 keV is in line with this. However, the model also predicts an improvement by a factor of 1.5 for a 20  $\mu\text{m}$  crystal, while Storm *et al.* observed no improvement.

A direct, quantitative comparison of half dose values between our experiment and that of Storm *et al.* is not meaningful, as they measured half dose for the whole diffraction pattern, not for individual peaks or resolution shells. Qualitatively, the difference in half dose they observed for the 5  $\mu\text{m}$  crystal is consistent with our observations. Both their observations and ours are consistent with the Dickerson-Garman model. Given a significant variation in experimental parameters and the metrics observed, the similarities between our results support the notion that using higher energy x-rays in MX experiments may have utility in damage-mitigation.

#### 4.4.7 Considerations For Future High Energy Diffraction Experiments

The experiments reported here were among the first to attempt a systematic study of the effect of beam energy on radiation damage in MX with micro-crystals. The challenges we faced, and particularly the improvements we realised between the two experiments, have some utility to guide future experimental design. As does the comparison between our experiments and that reported by Storm *et al.* [140].

#### The Importance a Cadmium-Telluride Detector

Several key factors prevented the collection of high quality radiation damage data at PETRA-III. One critical factor was the availability of a suitable detector. Even assuming the other technical issues with the beamline were addressed, the characteristics of the detector available for the experiment at the PETRA III beamline are not well-suited to collecting high-energy MX data. An improvement in diffraction efficiency (elastically scattered photons per unit dose) is only useful if the elastically scattered photons are registered at the detector, but higher energy x-rays are more likely to pass through the detector panel without interacting with it. Donath *et al.* [123] characterised the *quantum efficiency* (QE) of silicon detectors with various panel thicknesses, and showed that even the thickest panels (1000  $\mu\text{m}$ ) suffer from greatly reduced efficiency above 12 keV (see fig. 4.1a). By contrast, newer cadmium-telluride (CdTe) panels maintain high QE up to 26.7 keV, the *K*-edge of cadmium [124]. In our model of radiation damage (see Chapter 3), the QE of the detector was not explicitly considered. Dickerson and Garman [19] incorporated detector QE into a model of diffraction efficiency and determined that a CdTe detector was required to observe an improvement in diffraction efficiency due to PE escape and that the ideal

energy for experiments was 26 keV. Although diffraction efficiency (irrespective of the detector) continues to improve above 26 keV, as shown in our simulation, the drop in the QE of the detector above this energy off-sets that gain (see fig. 4.1b).

### **The Advantage of Using Polymer Chips For Sample Delivery**

We used two different chip designs for sample mounting at our PETRA-III and NSLS-II experiments: silicon chips based on the design of Roedig *et al.* [54], and polymer chips designed at the Perry lab, UMass Amherst. Both chips were attached to standard MX goniometer caps allowing them to be quickly and easily mounted at the beamline. The sample-loading process was also the same for both designs. The transparency of the polymer chips made it easier (compared to the silicon chips) to observe the sample on the chips under the microscope. This was helpful both during sample loading and when viewing the mounted chips with the on-axis camera. Draining excess solvent was much easier with the polymer chips though this may have been due simply to the drainage holes being larger. The polymer chips had a greater surface area, allowing more sample to be mounted on a single chip. This meant that fewer changes of sample were required, which reduced the need to enter the hutch, which we had to do frequently at the PETRA-III experiment. The main advantage of the silicon chips over the polymer chips was their rigidity. The longer aspect ratio, and more pliable material of the polymer chips meant that they vibrated under the pressure of the cryostream. This issue was resolved by epoxying wire supports to the chips. Of the two designs we used, the polymer chips are recommended for future experiments.

Storm *et al.* [140] used the cryo-EM grids that were originally suggested by Nave [11] as a tool that may allow minimisation of solvent around the crystals. However, they noted that the cryo-EM grids were challenging to align in the cryostream and to handle reproducibly at the beamline, which potentially resulted in different grids being exposed to different temperature and humidity, compromising the quality, and homogeneity, of their samples. Although we did not use cryo-EM grids, it seems likely that both styles of chip we used provided for easier handling of samples. Use of the cryo-EM grids would then only be recommended if it was clear that they provided better reduction of solvent compared to the chips. However, as we note below, beam focusing probably represents a better path to minimising solvent-illumination than physically reducing the solvent surrounding the crystal.

### **Beam Focusing Versus Removal of Solvent in the Sample Holder**

A fundamental issue with minimising radiation damage by exploiting PE escape is the scattering of PEs produced in the solvent into the crystal. This was discussed in an earlier paper by Nave [11]. The solution is to minimise the volume of illuminated solvent surrounding the

crystal, which can be achieved by matching the beam size (and position) to the crystal, by physically minimising the amount of solvent surrounding the sample, or by a combination of the two. Storm *et al.* focused on minimising solvent, though their analysis revealed that they still had 6.5  $\mu\text{m}$  to 11  $\mu\text{m}$  of solvent around their crystals. In our experiment, excess solvent surrounding the crystals was removed but the focus was on matching the beam size as closely as possible to the crystal size. The beam was not a perfect match however, so some illumination of solvent remained unavoidable. However, given our beam size (1.3  $\mu\text{m}$   $\times$  3.3  $\mu\text{m}$ ) and crystal size (2.45  $\mu\text{m}$   $\times$  2.34  $\mu\text{m}$   $\times$  1.24  $\mu\text{m}$ ) at NSLS-II, the amount of solvent illuminated was almost certainly less than that in the Storm experiment. Consequently, we would recommend focusing on matching beam size to crystal size, rather than relying exclusively on reducing solvent surrounding the crystals.

One thing that neither of these approaches address is the solvent channels that make up a significant portion of a protein crystal's volume (approximately 40% for lysozyme). It is impossible to avoid producing PEs in these solvent channels and they will inevitably scatter through the protein lattice. By minimising the crystal size, we can maximise the amount of energy they carry through the lattice and out into the solvent beyond. Ideally, the photoionisation cross section of the solvent would be reduced relative to the cross section of the protein itself, however this may not be feasible, given the sensitivity of protein crystals to, among other things, the ionic content of their solvents. We attempted to reduce the photoionisation cross section of the solvent by reducing the salt concentration (this being responsible for the heaviest atoms in our solvent) but our crystals rapidly dissolved. A systematic study of crystallisation conditions may suggest a protocol to reduce the heavy atom concentration of a given protein solvent somewhat, but photoionisation in the solvent channels is likely to remain an unavoidable source of damage.

## 4.5 Summary

Two experiments were conducted to examine the potential impact of PE escape on radiation damage in protein micro-crystals. The primary goals of the experiments were:

1. A qualitative comparison of the effect of beam energy on radiation damage, as measured by the metric of Bragg spot fading;
2. A quantitative comparison of experimental data to computational simulations of radiation damage that include PE escape, such as the model presented in Chapter 3, and the most recent version of RADDOSE-3D [139].

The analysis of data was discussed in the context of the relevant literature, including a similar study which was conducted more-or-less concurrently with this one. The main observations were that:

1. The data suggests the rate of Bragg spot fading is reduced at 26.00 keV relative to 12.66 keV, in keeping with theoretical predictions of PE escape.
2. Due to differences in the metrics used in modelling and in analysis of our results, a direct quantitative comparison is not meaningful. However, by comparison to other studies of dose tolerance, the results are in-line with the predictions.

In addition to these primary goals, the results provided several insights into the challenges of setting up and conducting experiments to take advantage of PE escape. The primary requirements for a successful experiment include:

1. The availability of a detector with high QE at higher energies, i.e. one with a cadmium telluride panel, rather than silicon;
2. A beamline capable of achieving a focal spot size close to 1  $\mu\text{m}$ , at high energy (e.g. > 20 keV);
3. An easy-to-use sample mounting system, such as the polymer chips used in our NSLS-II experiment.

The refinements and changes we realised between our PETRA-III experiment and our NSLS-II experiment resulted in a robust setup, well suited to the needs of examining the impact of PE escape on radiation damage. Future work should seek to repeat and extend upon the NSLS-II experiment. Initially, the collection of larger data sets at both standard MX (approx. 12.4 keV) and high (approx. 26 keV) x-ray energy for crystals of similar size to ours. Subsequently, examination of the impact of varying crystal size, different beam energies, and different protein crystals, with varying radiation tolerance.

## Chapter 5

# Comparison of Spot Fading Models

In this chapter we examine two resolution-dependent models of spot fading using diffraction data collected from lysozyme crystals at standard MX (12.66 keV) and high (26.00 keV) x-ray energy, and a discussion and comparison of spot fading research published over the last two decades. In section 5.1 we propose a framework for assessing and comparing models of spot fading. Section 5.2 contains an overview of existing spot fading models. Our modelling approach and methods for data collection and processing are discussed in section 5.3. We present our key results in sections 5.4.2 and 5.4.3, and those of our meta-analysis in section 5.4.4. All results are discussed together, with reference to the literature, in section 5.5, and our conclusions and recommendations for future study are presented in section 5.6.

### 5.1 Introduction

Theoretical and empirical modelling of radiation damage and its effects have been a component of the study of MX more-or-less since the introduction of protein crystallography in the middle of the 20th century. Radiation damage is manifest in a variety of observable changes, and individual models must typically be narrowly focused on a particular damage metric in order to be of practical value. The loss of intensity in diffraction patterns as a function of radiation dose is a commonly cited damage metric. This change may be readily observed in close-to-real-time at experiments, and serves as a coarse, yet useful, indicator of the quality of diffraction data. In this chapter we examine the progression of empirical models of spot fading and interrogate two current models as general predictors of damage rates in reference to a range of data from our own experiments and previously published results in the literature.

In making comparisons between multiple models and experimental data, it is necessary to qualify how we intend to judge them. A model might be judged to be 'good' or 'useful' in a variety

of ways. In the case of modelling loss of intensity as a metric for radiation damage, we present three criteria by which a model might be evaluated:

1. *Can the model accurately predict the rate at which Bragg spots will fade in an experiment?* The utility of a predictive model is in planning an experiment, ensuring that sufficient sample is available to obtain the required number of diffraction patterns, given the resolution requirements of the experiment. Ideally, the model will be accurate across a wide range of samples and beam-configurations. Accounting for the resolution-dependence of intensity loss is important.
2. *Can the model be used to retrospectively correct peak intensities in a manner that improves structure solution?* As noted in Chapter 2, the intensities of Bragg peaks *relative to each other* is a critical component in the Fourier analysis of diffraction patterns, revealing information about the contents of the unit cell (i.e. the composition of the protein). Global damage leads to a resolution-dependent reduction in peak intensity: high resolution peaks fade more quickly than lower resolution peaks. Site-specific damage can also lead to peaks of the same resolution fading at different rates. Both of these effects can reduce the utility of diffraction from damaged crystals to structure solution. A model used for the purpose of correcting peak intensities should at least accurately describe the resolution dependent loss of intensity. Site-specific effects would require the use of a more specialised model.
3. *Does the model offer insight into the physical mechanisms underlying radiation damage?* Although the immediate needs of experimenters may be met by models focusing on the first two criteria, an understanding of what is actually happening inside samples in order to produce the visible effects of radiation damage (such as spot fading) can have important implications for the planning of experiments, and for the understanding of their results. For example, cryogenic cooling was observed to radically extend the usable life-times of crystals with respect to radiation dose; how exactly this occurs may have implications for the interpretation of diffraction patterns taken at cryogenic temperature vs. room temperature.

A model with fixed coefficients will provide an unambiguous prediction. To be generally applicable, these coefficients should be based on empirical averages of previous, similar measurements. As there will be variation in radiation tolerance from sample to sample, these averages need not provide a precise fit to any given sample in order to be deemed useful. Conversely, if coefficients are optimised by fitting to a specific data set, the model will provide a more precise description of that specific data. This increases its utility for correcting peak intensities, but reduces its applicability as a general predictive tool. We discuss the models described in section

5.2.1 with reference to all three of the above criteria, but for the purposes of comparing them to experimental data, we focus on their capacity to make accurate predictions for a variety of samples and experimental configurations, with tightly constrained input parameters.

## 5.2 Background

### 5.2.1 Modelling the Impact of Radiation Damage on Diffracted Intensity

Deterioration of diffraction patterns during continuous x-ray exposure has been observed since the inception of x-ray protein crystallography, as has the increased rate at which the higher-resolution components of the pattern degrade. A reduction in diffracted intensity is far from the only issue arising from radiation damage, as has been discussed elsewhere in this thesis. However even considered in isolation an uneven reduction in intensity is problematic for protein structure solution as it is the *relative* intensities of the Bragg peaks that reveal the structure (via electron density) of the protein molecule. As noted above in point 2, rates of spot fading may differ as a function of resolution, or due to site-specific damage. In the former case, the rate of intensity loss will be independent of azimuthal angle, i.e. it will be the same for all spots of a given resolution. In the latter, the affected Bragg peak(s) will fade more quickly than other peaks of the same resolution. Both of these effects may compromise structure retrieval even while the overall intensity of a diffraction pattern is still strong. In principle, the former effect can be corrected for based on models of global damage like the ones explored in this chapter. Correction of the latter effect would rely on models of specific damage that are beyond the scope of this discussion.

#### Blake & Phillips: An Early Model of The Progression of Radiation Damage

Blake & Phillips [83] proposed an early comprehensive model of x-ray damage that was subsequently refined by Hendrickson [84], and later by Sygusch & Allaire [85]. The crystal was assumed to be composed of 3 parts: an undamaged fraction,  $A_1$ ; a disordered fraction,  $A_2$ , which would contribute significantly to diffraction only at lower scattering angles; and an amorphous fraction,  $A_3$ , which did not contribute at all to diffraction. Assuming the acquired disorder to be random, with a mean-square atomic displacement of  $X/8\pi$  (the original papers use  $D/8\pi$ , we substitute  $X$  for  $D$  to avoid confusion with dose), the intensity,  $I(t)/I(0)$ , of a diffraction pattern of a partially damaged crystal, relative to an undamaged one (i.e. after exposure time,  $t$ ), can be written as

$$\frac{I(t)}{I(0)} = A_1(t) + A_2(t) \exp(-Xs^2), \quad (5.1)$$

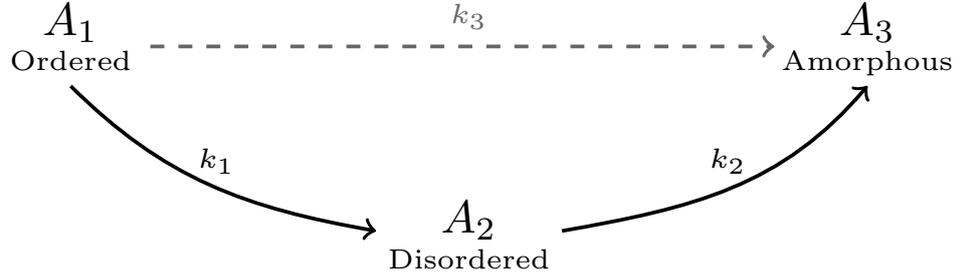


Figure 5.1: **Diagrammatic representation of the Blake & Phillips model of crystal damage.**  $A_1$ ,  $A_2$ , and  $A_3$  represent fractions of the crystal sample in three states: ordered, disordered, and amorphous, respectively. The  $A_2$  crystal fraction contributes to diffracted intensity only at low scattering angles (low resolution), the  $A_3$  fraction does not contribute to diffraction at all.  $k_1$ ,  $k_2$ , and  $k_3$  are the rates at which transitions occur between the three states; Hendrickson [84] obtained best fits to available data when  $k_3 = 0$ , suggesting that transition to the amorphous state happened exclusively via the disordered state (illustrated in this diagram by rendering the  $k_3$  transition as a greyed-out dashed-line).

where  $s = \sin(\theta)/\lambda$ . The evolution of  $A_i$  over time was governed by rates,  $k_i$ , as represented in fig. 5.1:  $k_1$  was the rate at which the crystal progressed from ordered to disordered ( $A_1 \rightarrow A_2$ );  $k_2$  the rate for disordered to amorphous ( $A_2 \rightarrow A_3$ ); and  $k_3$  the rate for ordered directly to amorphous ( $A_1 \rightarrow A_3$ ).

The rates,  $k_i$ , would be proportional to the pool of available sample in the appropriate state, producing the system of differential equations set out below:

$$\frac{dA_1}{dt} = -k_1A_1 - k_3A_1, \quad (5.2)$$

$$\frac{dA_2}{dt} = k_1A_1 - k_2A_2, \quad (5.3)$$

and

$$\frac{dA_3}{dt} = k_2A_2 + k_3A_1, \quad (5.4)$$

subject to the conditions  $A_1(t) + A_2(t) + A_3(t) = 1$  and  $A_1(0) = 1 \Rightarrow A_2(0) = A_3(0) = 0$ . Solving these equations and substituting the results into eq. (5.1) yields:

$$\frac{I(t)}{I(0)} = \exp[-(k_1 + k_3)t] + \frac{k_1}{k_1 + k_3 - k_2} \exp[-k_2t](1 - \exp[-(k_1 + k_3 - k_2)t]) \exp[-Xs^2]. \quad (5.5)$$

Hendrickson fit eq. (5.5) —subject to a variety of constraints, such as  $k_3 = 0$ ,  $k_1 = k_2$ — to Blake & Phillips' myoglobin data, getting the best fits when  $k_3 = 0$ , thus somewhat simplifying eq.

5.5 and concluding that the progression of damage was strictly sequential, with the native crystal always becoming partially disordered prior to becoming entirely amorphous. Hendrickson [84] proposed that the two processes,  $A_1 \rightarrow A_2$  and  $A_2 \rightarrow A_3$ , might be caused by two different free-radicals diffusing through the material, or a succession of interactions of the same type. The suggested utility of the model was in understanding the nature of radiation damage and in making corrections to intensity measurements to facilitate structure solution. The former point is no longer relevant as in the context of current understanding of radiation damage mechanisms (see Chapter 2), the coefficients of the Blake equations do not offer a useful physical interpretation, and the on-off nature of the crystal fractions as contributors to diffraction is too coarse to be compatible with a modern understanding of acquired disorder. However, with multiple parameters to optimise fitting, the model continues to have utility in correcting intensity measurements [86], but it does not offer any utility as a predictive tool, given that the coefficients are not, in general, known in advance. It may be possible to make use of previous fitting data for a given protein or protein-family to produce an empirical predictive model, however simpler models now exist, such as those due to Holton [71] and Atakisi [86], which show a strictly exponential relationship between dose and intensity; these models are explored in detail below.

#### **The Howells-Holton Model: Spot Fading as a Function of Dose and Resolution**

An empirical spot fading model (eq. (2.18)) was proposed by Holton [71], based on a comparison of recent data with the findings of Howells *et al.* [87], which included a meta-analysis of spot fading across a variety of electron and x-ray diffraction imaging techniques. Intensity loss was quantified based on the so-called *half dose*,  $D_{1/2}$ , which is the dose required to reduce diffracted intensity to half its initial value. Half dose is a commonly cited figure of merit in assessing the extent of global radiation damage. Given that studies of damage-induced spot fading have examined the intensity of individual spots, resolution shells, and entire diffraction patterns (with varying resolution limits) [74, 87, 119], quoted half doses can be ambiguous when it is not explicitly stated what measured intensity is halving. We will subsequently reserve the general term *half dose* to describe the dose required to reduce the intensity of *an entire diffraction pattern* to a half its initial intensity. Further this figure of merit will be accompanied by the initial maximum resolution of the diffraction pattern. When talking about the fading of an individual Bragg spot or resolution shell, we will use the terms *spot half dose* and *shell half dose*, noting that, on average, provided resolution shells are not particularly broad, these two labels should be equivalent.

Equation (2.18) doesn't suggest any information about the physical processes underlying the

reduction in spot intensity, rather it highlights an observed relationship between dose and intensity. Note that this strict dose-dependence, implying dose-rate-independence, is true for samples at cryogenic temperatures, but not necessarily true for samples at room temperature [74]. The primary utility of the Howell-Holton model is as a predictive tool. All the parameters of eq. (2.18) are knowable in advance, and it can be used to estimate maximum exposure times based on a desired limit to the ratio  $I_d(D)/I_d(0)$ . The form in which the equation is written provides an immediate, intuitive link to the spot half dose figure of merit:  $I_d(D)/I_d(0)$  will be 0.5 when  $D = Hd$ . As all the parameters in eq. (2.18) are fixed, it is not possible to optimise it to fit existing data. It is therefore not suitable to the task of correcting intensities for improved structure solution, which requires a less-constrained model such as the Blake-Phillips equation or simply an unconstrained exponential fit.

### **The Atakisi Model: A Non-Linear Relationship Between Resolution and Shell Half Dose**

Recently, Atakisi *et al.* [86] proposed a model of spot fading (eq. (2.19)) that combines some features of both the Blake-Phillips model and the Howells-Holton model. It has the same simple exponential form as the Howells-Holton model, but proposes a non-linear relationship between resolution and shell half dose, as in the Blake-Phillips model. Like the Howells-Holton model, the Atakisi model does not provide any insight into the physical nature of damage mechanisms, but rather tracks the empirical observation that spot intensity fades exponentially. It differs from the Howells-Holton model in assuming a non-linear relationship between resolution,  $q$  (or equivalently  $d$ ), and shell half dose; and by determining parameters by fitting to a specific data set rather than fixing them based on meta-analysis of previous data. Consequently, it has utility for correcting spot intensities, like the Blake-Phillips model, but is not, in its initial published form, suited to predicting spot fading, like the Howells-Holton model. It can, however, be adapted fairly simply for use as a predictive model, as described in section 5.3.1

### **5.2.2 MX at Different X-Ray Energies**

The vast majority of synchrotron MX experiments are conducted with an x-ray beam energy of around 12 keV to 13 keV, which can, in theory, provide an atomic-resolution structure. In practice, other experimental parameters place tighter constraints on the maximum-achievable resolution. Using lower energy x-rays fundamentally limits the attainable resolution, for example at 9 keV, an energy commonly accessible at MX beamlines, the maximum resolution would be reduced to 1.4 Å. Lower energy x-rays also potentially reduce experiment time as the cross-section for elastic scatter is higher at lower energies, however the exposure times required at 4<sup>th</sup>

generation synchrotrons are already quite short at 12 keV, and the cross-section for photoionisation also increases (more steeply than that for elastic scattering) at lower energies, so damage will accrue faster relative to diffracted intensity. The use of lower beam energies may be desirable for the purposes of phasing, as several potential targets for anomalous diffraction have absorption edges below 12 keV.

The primary arguments against conducting experiments at higher energies are that it doesn't necessarily lead to an increase in potential resolution; it extends experiment time due to the reduction in elastic scattering cross-section; and it adds complexity/difficulty to the experimental design, given that the beamline is likely optimised to run at around 12 keV for MX experiments. There is a relative gain in the ratio of elastic scattering to photoionisation cross-sections to be had at higher energies, but this is not particularly significant (a factor of 2, for example, at 30 keV relative to 12 keV). As discussed in Chapters 3 and 4, there is now a considerable body of theory, and some experimental evidence, that when using higher energy x-rays, the increased spread of PEs released in the crystal by photoionisation events may result in a reduction of damage, relative to diffraction, provided the crystals and beam spot sizes used are appropriately small (on the order of 1  $\mu\text{m}$ ). Under these conditions, it is expected that the relationship between dose and spot fading is unchanged, provided PE escape is taken into account in the dose calculation; if it is not, there would need to be a linear multiplier applied to account for the reduction in dose due to PE escape. However, at this time, it is not clear that the dose reduction predicted by our model and RADDPOSE-3D is quantitatively aligned with experimental results and further data would need to be collected to determine empirically what such a multiplier is expected to be. For the time being, it seems prudent to predict damage rates without accounting for PE escape, which will set a conservative limit on the lifetime of a crystal, rather than relying on a predicted damage reduction that may not be realised in practice.

In the following sections, we propose an adaptation of the Atakisi model that is designed to be predictive rather than descriptive, allowing it to be compared like-to-like with the Howells-Holton model. We then test the predictions of these two models against data sets collected during one of the experiments described in Chapter 4 and a meta-analysis of spot fading results from the literature, to assess the capacity of the two models to predict spot fading for different samples under various experimental conditions.

### 5.3 Methods

We reviewed spot fading and curve-fitting results from previous experiments [75, 78, 86, 117, 141, 142], in order to select fixed coefficients for a predictive spot fading model based on the

Atakisi equation (eq. (2.19)). The predictions of this model (Model A) and the Howells-Holton model (eq. (2.18), Model H) were then compared to spot fading rates from two sets of experimental diffraction data, collected at cryogenic temperature, using standard (12.66 keV) and high (26.00 keV) x-ray beam energies. We additionally compared the predictions of Model A and Model H to several previous spot fading studies, in order to obtain a sense of the variance in spot fading rates, and the limits of the spot fading models.

### 5.3.1 A Predictive Version of the Atakisi Model

As described in section 5.2.1, the selection of coefficients in the Howells-Holton and Atakisi models (eqs. (2.18) and (2.19)) are optimised for different tasks. In the Howells-Holton model, all the coefficients are set (once a resolution value,  $d$ , is selected by the user); the independent variable is dose,  $D$ , which can be estimated based on the configuration of the experiment. The primary utility of this model is then predictive, suggesting a likely spot half dose,  $D_{1/2} = Hd$ , as a figure of merit for the life-time of the sample at a desired resolution.

By contrast, the Atakisi model is optimised to fit existing experimental data by tuning its coefficients, thereby being more suitable for correcting spot intensities, but lacking in predictive utility. In order to facilitate a direct comparison of the two models, and to test them against experimental data, we express the Atakisi model in the form

$$I_d(D) = I_d(0) \exp\left(-\ln(2) \frac{D}{\kappa d^\alpha}\right), \quad (5.6)$$

where  $\kappa$  and  $\alpha$  are coefficients that give the relationship  $D_{1/2} = \kappa d^\alpha$ . In order to select values of  $\alpha$  and  $\kappa$ , we reanalysed the relationship between spot/shell resolution and shell half dose summarised by Atakisi *et al.* in fig. 7 of their paper, restricting  $d$  to a low-resolution limit of 5 Å. The spread in shell half dose values as a function of resolution means that including a wider range of resolutions will reduce the quality of fit within a given subset of that domain. Given that crystallographers are typically most interested in the highest resolution data in their diffraction patterns, we chose to limit the domain in order to provide a better fit at high resolution. The 5 Å limit we selected included results by Liebschner *et al.* [75], Bourenkov and Popov [141], Teng and Moffat [78], Warkentin and Thorne [142], and Sliz *et al.* [117], but excluded results from Coughlin *et al.* [26, 27], Howells *et al.* [87], Nave *et al.* [143], and Owen *et al.* [119], all of which investigated spots/shells at lower resolutions. We used WebPlotDigitizer [144] to extract values from the plot, compared these to the general shell half dose values proposed by Atakisi *et al.*, and tested values of  $\alpha$  and  $\kappa$  to fit the available data. Based on our analysis, we selected  $\kappa = \alpha = 2$ ,

expressing Model A as:

$$I_d(D) = I_d(0) \exp\left(-\ln(2)\frac{D}{2d^2}\right). \quad (5.7)$$

### 5.3.2 Collection & Analysis of Primary Data

Data collected at the 17-ID-2 (FMX) beamline of the NSLS-II synchrotron were used to test the models. The data-collection and primary analysis of diffraction images are detailed in section 4.2.3 of Chapter 4.

#### Processing Peak Intensities

Lists of peak intensities generated with CrystFEL [113] were read and processed using Python scripts. For each data set, we performed preliminary analysis to determine maximum and minimum peak resolutions, and to identify inconsistencies/errors that may exist due to noise, background not being accurately subtracted in peak-finding, or individual peaks strongly skewing average behaviour. Maximum resolution was 1.41 Å for 12.66 keV data sets and 2.42 Å for 26.00 keV data sets. For the 12.66 keV data sets there were very few peaks with resolution better than 2 Å, and for the 26.00 keV data sets there were very few peaks with resolution better than 3 Å; at both energies very few peaks were discernible below 10 Å, and those that were showed significant variation from expected exponential decay. Low peak numbers in these ranges resulted in poor averaging and accordingly, we elected to exclude peaks below 10 Å (for both beam energies) and above 2 Å (for 12.66 keV data sets) or 3 Å (for 26.00 keV data sets). Peaks were grouped into shells as shown in table 5.1. Both the total and average intensities were calculated per resolution shell for each frame. The total and average intensities of the entire diffraction patterns were also calculated for each frame (truncated at 10 Å), with upper resolution limits set, successively, at 2 Å, 2.5 Å, 3 Å, 3.5 Å, 4 Å, and 5 Å.

To normalise intensities we fit unconstrained exponential curves (of the form  $I = a \exp(-bD)$ ) to the summed/averaged data for each shell of each data set. We took  $a$  to be the undamaged intensity,  $I(0)$ , of that shell, and used it to scale subsequent intensity measurements. Integrated intensities were normalised in the same fashion, except that we only dealt with summed peak intensities, not average intensities.

For the data binned by resolution shell, we compared the results based on summed intensities to those based on average intensities and confirmed that after normalisation they were equivalent; we subsequently used only results based on summed intensities in our analysis.

Table 5.1: **Histogram bins for sorting of Bragg peaks into resolution shells.** Peaks in the diffraction patterns were grouped into the shells shown, according to the x-ray energy used for data collection.

Beam Energy	
12.66 keV	26.00 keV
(2.0 - 2.5)Å	(3.0 - 3.5)Å
(2.5 - 3.0)Å	(3.5 - 4.0)Å
(3.0 - 3.5)Å	(4.0 - 5.0)Å
(3.5 - 4.0)Å	(5.0 - 10.0)Å
(4.0 - 5.0)Å	
(5.0 - 10.0)Å	

### 5.3.3 Re-analysis of Previous Studies

WebPlotDigitizer [144] was used to extract spot fading data from results published by Liebschner *et al.* [75] (fig. 4 in their paper), Bourenkov and Popov [141] (figs. 3c, 4b, 6b in their paper), and Warkentin and Thorne [142] (fig. 4 in their paper). Across the three studies, loss of intensity within resolution bands ranging from 1.22 Å to 8.0 Å was observed for four different crystal samples (thaumatin [75, 142]; and P19-siRNA, feruloyl esterase (FAE), and the membrane protein FtsH [141]). Unconstrained exponential decay curves were fit to the data, using the results to normalise intensities (where the reported intensities were not already normalised) and to determine shell half doses, which were compared to the shell half doses extracted from the results of Atakisi *et al.*, with which they showed excellent agreement. All intensity data were then compared to predictions of Model A and Model H.

## 5.4 Results

### 5.4.1 Shell Half Doses & Model A

Figure 5.2 summarises the relationship between spot/shell resolution and *shell half dose* observed in five previous studies [75, 78, 117, 141, 142] for Bragg spots between the resolution of 1 Å and 5 Å. Shown on the plot are trend lines corresponding to Model A ( $D_{1/2} = 2d^2$ ), Model H ( $D_{1/2} = 10d$ ), and the best-fit ( $D_{1/2} = 3.5d^{1.6}$ ). The fit suggested by Atakisi *et al.* [86] ( $D \propto d^{1.86}$ ) is excluded to improve clarity of the graph; it would sit between the curves of best fit and Model A. Both Model A and the best-fit provide an excellent fit to the data. While Model H does not describe the data well in this resolution range.

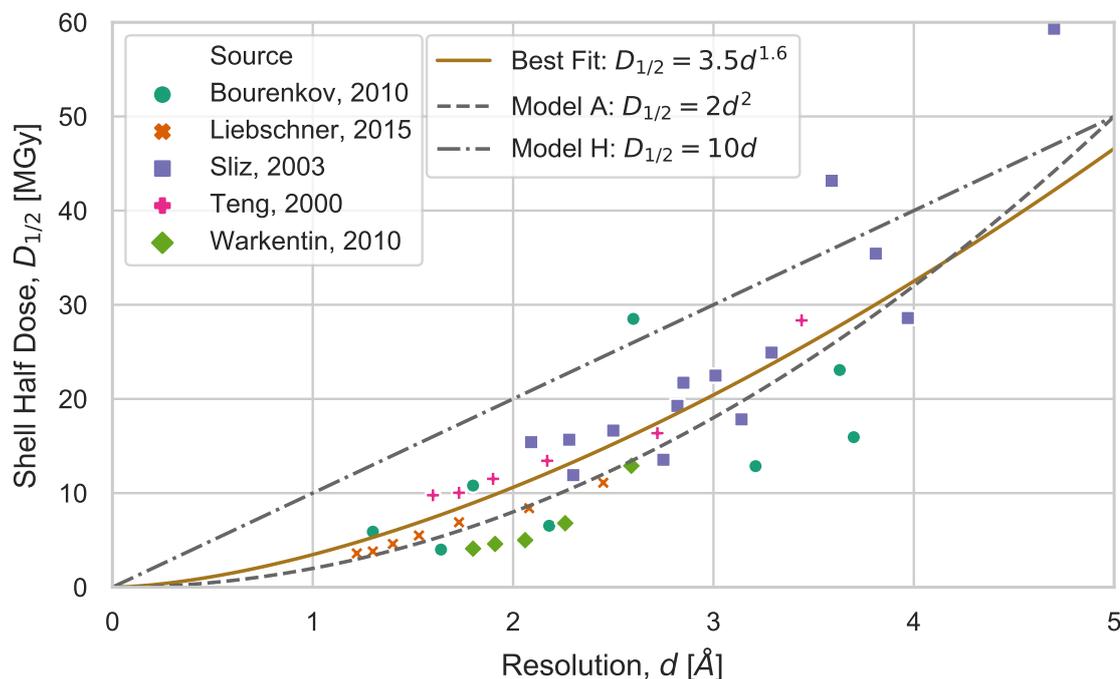


Figure 5.2: **Shell half dose measurements for resolutions between 1 Å and 5 Å, obtained from various studies.** A line of best-fit is shown, as are lines corresponding to Model A and Model H. There is minimal difference between the line of best fit and Model A; both curves model the data very well. Model H consistently predicts higher shell half doses than evident in the experimental data [75, 78, 117, 141, 142].

#### 5.4.2 Micro-Crystallography at Cryogenic Temperature and Standard MX X-Ray Energy

The two figures in this section are very similar in appearance. Both show normalised peak intensities for subsets of the six diffraction time series that we recorded with a beam energy of 12.66 keV. In fig. 5.3 each individual subplot includes peaks in a specific resolution band (e.g. all peaks with  $d \in [2.0, 2.5)\text{Å}$  in the teal plot), in fig. 5.4 the data are the summed intensities of all peaks in the diffraction pattern up to a given maximum resolution (e.g. all peaks up to  $d_{MAX} = 2\text{Å}$  in the teal plot). The lowest resolution peaks included in *any* of the plots is 10 Å. A consequence of this low-resolution cut-off is that the final (yellow) subplot of figs. 5.3 and 5.4 contain the same data, though the cut-off on the  $y$ -axis is slightly different, so they are not visually identical. The low resolution cut-off was used as our diffraction data contained very few Bragg spots with resolutions below 10 Å, and they tended to introduce significant noise to the data.

Figure 5.3 presents a comparison of the two resolution-dependent spot fading models. Figure 5.4

compares only Model H (the Howells-Holton model) to the summed intensities. The motivation for including the second figure was an observation that many half doses reported in the literature [72, 75, 76, 78, 79, 96, 119] for the integrated intensities of whole diffraction patterns seemed to align well to the prediction of the Howells-Holton model, even though that model was intended to predict spot fading of individual Bragg spots of a given resolution, and we wanted to analyse this further. This trend is explored in section 5.5.4, with reference to this experimental data and other, previously published, results.

Figure 5.3 shows normalised peak intensities for the six 12.66 keV experimental data sets, with Bragg peaks between 2 Å and 10 Å grouped into six resolution shells. Three trend lines are shown: our predictive version of the Atakisi model (Model A:  $I/I(0) = \exp(-\ln(2)D/2d^2)$ ), the Howells-Holton model (Model H:  $I/I(0) = \exp(-\ln(2)D/10d)$ ), and an unconstrained best-fit exponential curve ( $I/I(0) = \exp(-kD)$ , with  $k$  selected to optimise fit). For the purposes of fitting the models, the resolution was taken to be the upper edge of the shell (lower numeric value). i.e in the teal plot,  $d = 2$  Å.

In each subplot, there is significant variance in the experimental spot fading data. This variance is more apparent at higher resolutions. The best-fit exponential curve is a good description of the mean in each subplot. Both Model A and Model H fit the experimental data well. At higher resolutions, Model A is slightly steeper than the best fit, and Model H is slightly shallower. The largest discrepancy between either model and the best fit is a factor of 2. When  $d = 5$  Å, the two models overlap as  $10 \times 5 = 2 \times 5^2$ .

Figure 5.4 also shows normalised peak intensities for the six 12.66 keV experimental data sets. The data are the summed peak intensities from the low resolution cut-off of 10 Å up to a maximum resolution,  $d_{MAX}(0)$ , as labelled above each subplot. Two trend lines are shown: the Howells-Holton model and an unconstrained best-fit exponential curve. For the purposes of fitting Model H,  $d$  was interpreted as the maximum resolution of the undamaged diffraction pattern, i.e. for the  $d_{MAX}(0) = 2$  Å data set, the highest resolution in the first diffraction image was 2 Å, but by the time the crystal had absorbed 100 MGy, any 2 Å spots could well have faded below the background level.

In each subplot, there is significant variance in the experimental data. The best-fit line is a good description of the mean. Model H fits the data well for all maximum resolutions; it differs from the best-fit by up to a factor of 2 at the highest resolutions, coming closer to the best-fit at the lower resolutions.

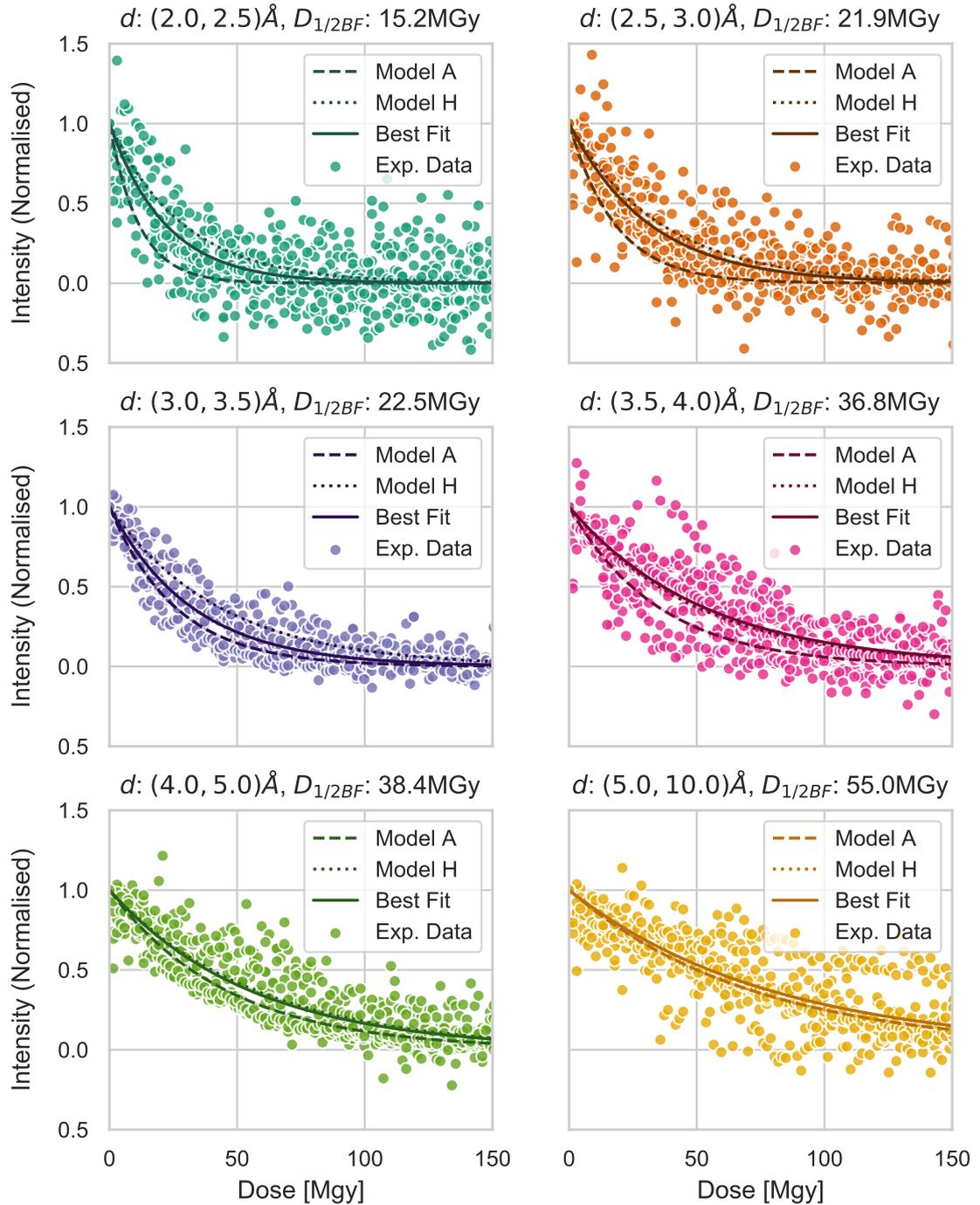


Figure 5.3: **Normalised Bragg spot intensities of six resolution shells as a function of dose at 12.66 keV.** Experimental data ( $\bullet$ ) are normalised peak intensities from six time series data sets, summed within six resolution shells. Variance in experimental data is high; it is most significant at higher resolutions, minimised around  $3\text{\AA}$  to  $4\text{\AA}$ . The (best fit) unconstrained exponential well describes the mean in each resolution band. Half doses for best fit lines are shown in subtitles. Both Model A (dashed lines) and Model H (dotted lines) describe the data well. At higher resolutions Model A predicts steeper fading than the best fit, while Model H predicts more gradual fading. At lower resolutions, both models and the best fit almost coincide. At  $d = 5\text{\AA}$ , Model A and Model H are equivalent.

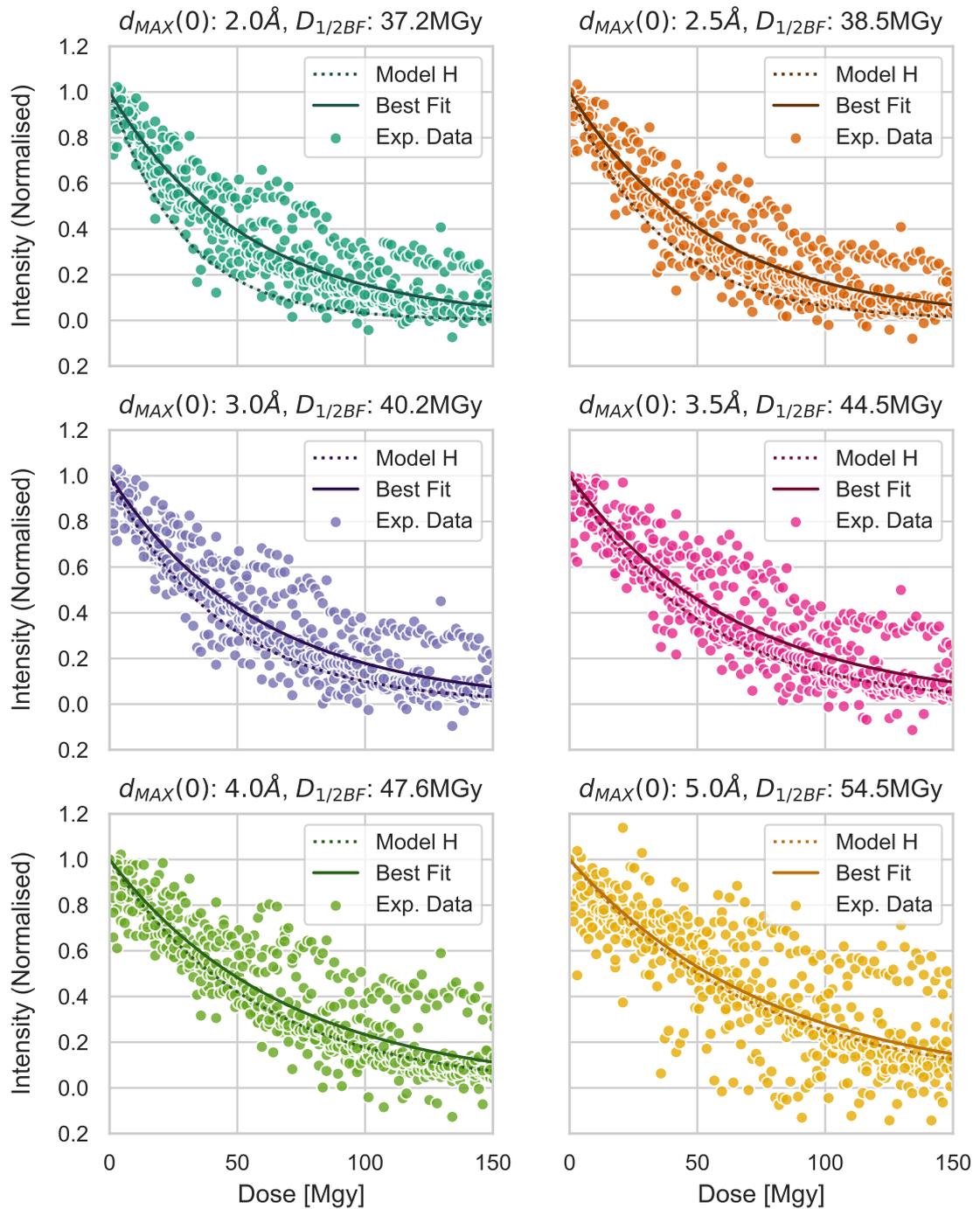


Figure 5.4: **Normalised intensities of entire diffraction patterns as a function of dose at 12.66 keV.** Experimental data ( $\bullet$ ) are normalised peak intensities from six time series data sets, summed over the entire diffraction pattern up to a given maximum resolution. All plots have a low resolution cut-off of  $10\text{\AA}$  and high-resolution cut-offs as labelled. Variance in experimental data is moderate, tending higher when high resolution peaks are excluded. The final graph (bottom right, for  $d_{max} = 5.0\text{\AA}$ ) is identical to the corresponding graph in fig. 5.3, however the  $y$ -axis has different limits. The (best fit) unconstrained exponential well describes the mean in each resolution band. Half doses for best fit lines are shown in subtitles. Model H (dotted lines) also tracks the data well, tending to predict slightly steeper fading than the best fit, but coming closer to it when the higher resolution are excluded.

### 5.4.3 Micro-Crystallography at Cryogenic Temperature and High X-Ray Energy

The two figures in this section have the same format as figs. 5.3 and 5.4 (see section 5.4.2). The data presented in figs. 5.5 and 5.6 are from 15 time series collected at 26.00 keV. They have the same low resolution cut-off (10 Å) as the 12.66 keV data, for the same reason explained above. The high resolution cut-off is 3 Å. There were a handful of Bragg peaks in the 26.00 keV data sets with resolution higher than this, but their SNR was too low to provide statistically meaningful results.

Figure 5.5 shows normalised peak intensities for the fifteen 26.00 keV experimental data sets, with Bragg peaks between 3 Å and 10 Å grouped into four resolution shells. As in fig. 5.3, three trend lines are shown: Model A, Model H, and an unconstrained exponential, with the upper edge of the resolution shell used for fitting.

In each resolution band, the 26.00 keV data has greater variance than the 12.66 keV data. This is due in part to the 26.00 keV data having a lower SNR, and in part to there being more data sets (in figs. 5.3 and 5.4 there are six data points —with independent  $y$ -coordinates— for each dose value on the  $x$ -axis, whereas in figs. 5.5 and 5.6 there are fifteen). Like the 12.66 keV data, there are a larger number of outliers at higher resolutions. The best-fit exponential curve is again a good description of the mean in each subplot. Both Model A and Model H fit the data reasonably well, deviating from the best fit by up to a factor of 2.5. Both models predict slightly slower fading than the best fit and the gap increases at lower resolutions, which is opposite to the trend seen in 5.3. As with the 12.66 keV plots, the two models are equivalent when  $d = 5$  Å, so their plots overlap.

Figure 5.6 shows normalised integrated intensities for the fifteen 26.00 keV experimental data sets. The data are the summed peak intensities from the low resolution cut-off of 10 Å up to a maximum resolution,  $d_{MAX}(0)$ , as labelled above each subplot. Two trend lines are shown: the Howells-Holton model and an unconstrained best-fit exponential curve. As with fig. 5.4, when fitting Model H,  $d$  was taken to be  $d_{MAX}(0)$

For the purposes of fitting Model H,  $d$  was interpreted as the maximum resolution of the undamaged diffraction pattern, i.e. for the  $d_{MAX}(0) = 3$  Å data set, the highest resolution in the first diffraction image was 3 Å, but by the time the crystal had absorbed 100 MGy, any 3 Å spots may have faded below the detectable limit.

In each subplot, there is significant variance in the experimental data, though for higher maximum resolutions there are fewer outliers than seen in fig. 5.5, where the peaks were grouped into resolution bands. The best-fit line is a good description of the mean. Model H provides an

adequate fit for the experimental data, though it consistently tracks the slowest fading spots. As with fig. 5.5, Model H is closer to the best fit at higher resolutions and diverges from it at lower resolutions, which is the opposite of the trend seen at 12.66 keV in figs. 5.3 and 5.4.

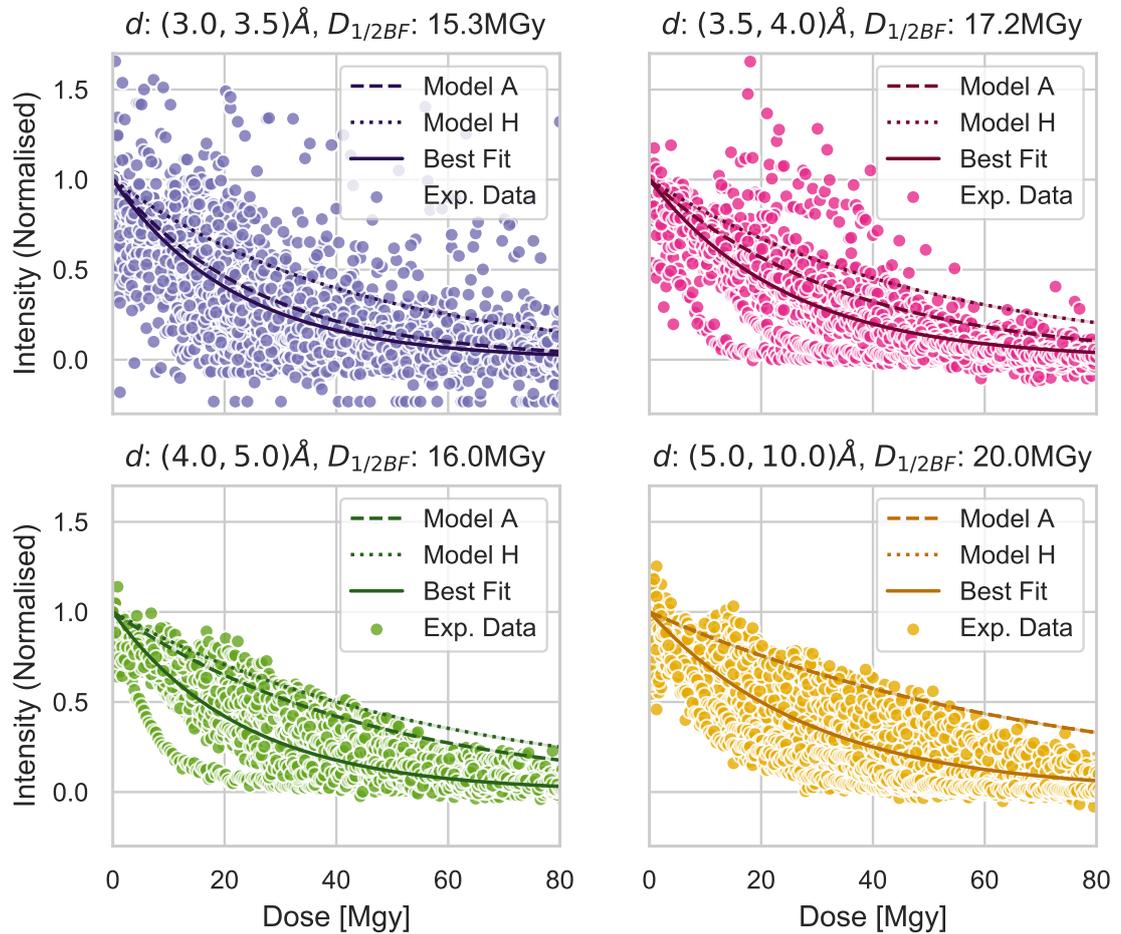


Figure 5.5: **Normalised Bragg spot intensities of four resolution shells as a function of dose at 26.00 keV.** Experimental data ( $\bullet$ ) are normalised peak intensities from fifteen time series data sets, summed within four resolution shells. Variance in experimental data is high, especially at 3  $\text{\AA}$  and 3.5  $\text{\AA}$ . The (best fit) unconstrained exponential well describes the mean in each resolution band. Half doses for best fit lines are shown in subtitles. Both Model A (dashed lines) and Model H (dotted lines) fit the data well, differing from the best fit by a factor of 2 at most. Both models predict slower spot-fading than the best fit. At  $d = 5 \text{\AA}$ , Model A and Model H are equivalent.

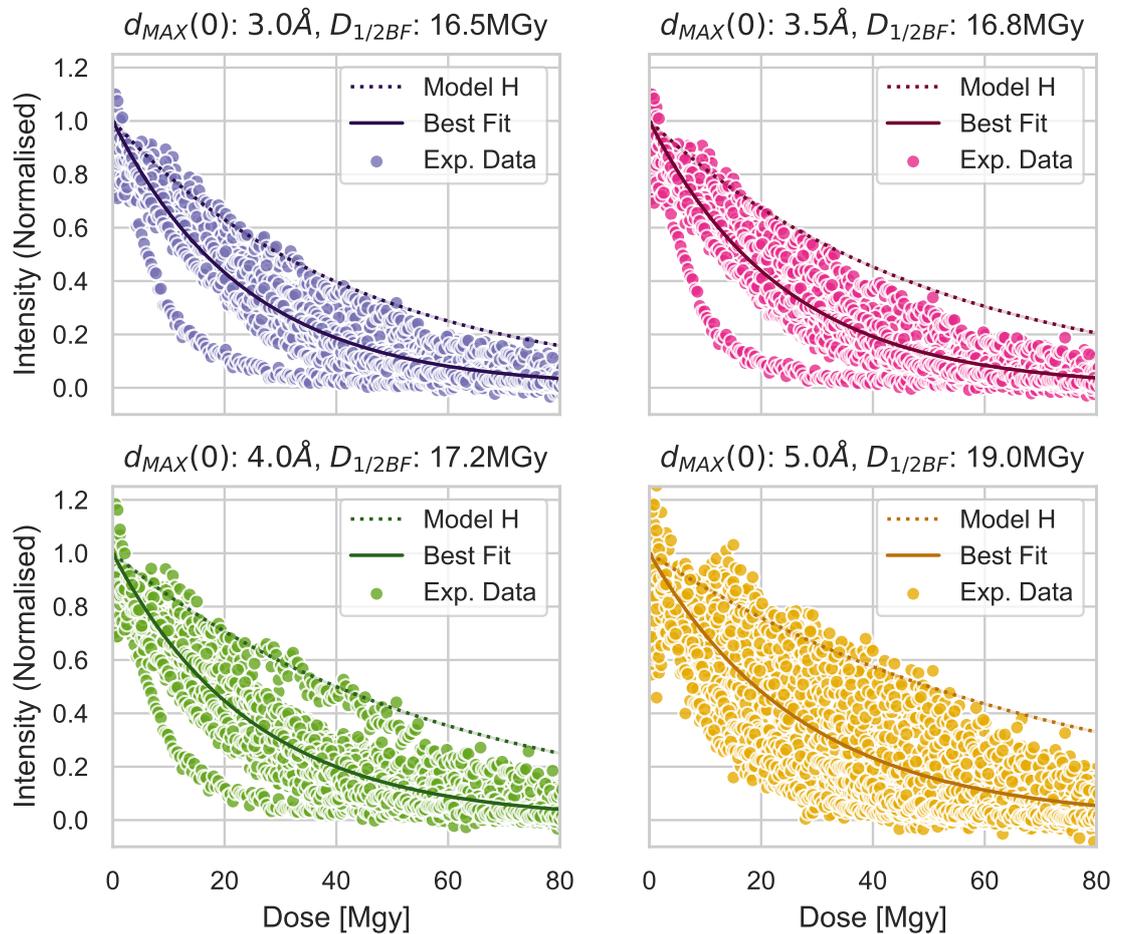


Figure 5.6: **Normalised intensities of entire diffraction patterns as a function of dose at 26.00 keV.** Experimental data ( $\bullet$ ) are normalised peak intensities from fifteen time series data sets, summed over the entire diffraction pattern up to a given maximum resolution. All plots have a low resolution cut-off of  $10\text{\AA}$  and high-resolution cut-offs as labelled. There is significant variance in the experimental data. Variance is highest when peaks with  $d < 5.0\text{\AA}$  are excluded; this graph (bottom right) is identical to the corresponding graph in fig 5.5, though the graph has a different cut-off on the  $y$ -axis. The (best fit) unconstrained exponential describes the mean well in each resolution band. Half doses for best fit lines are shown in subtitles. Model H (dotted lines) predicts fading at the slowest rate seen in the experimental data, for all resolution limits. This prediction is within a factor of 2 of the mean, except in the last (bottom right) graph, where it out by a factor of 2.5.

#### 5.4.4 Re-analysis of Previous Studies

Figure 5.7 shows plots for a selection of data that we included in our meta-analysis of spot fading and half dose results. The full spot fading meta-analysis included 22 data-sets of Bourenkov and Popov [141], Warkentin and Thorne [142], and Liebschner *et al.* [75], covering four different crystal samples, with shell resolutions between 1.22 Å and 8.0 Å. A list of resolutions for the studies is shown in table 5.2, along with key sample and beam parameters. Plots of the full data sets can be seen in appendix B.

Table 5.2: **Samples & resolution shells of three spot fading studies used in our meta-analysis.** All data were collected at cryogenic temperatures. Bourenkov & Popov [141] quoted resolutions, Warkentin & Thorne [142] and Liebschner *et al.* [75] quoted resolution shells. In comparing data to models we took  $d$  to be the upper edge of the resolution shell. Resolutions listed in **bold type** are included in the overview in fig. 5.7. Plots of spot fading for all resolution shells are shown in appendix B. Some crystal sizes were not specified. Multiple crystal sizes, and beam sizes and energies, were used in some experiments; we have included only the parameters that pertain specifically to the data sets used in our meta-analysis.

Sample	Beam	Resolution Shell [Å]
<b>Bourenkov &amp; Popov</b>		
FtsH (50 µm × 120 µm)	11.8 keV 40 µm × 30 µm	3.2, 3.7, 7.5
FAE (>> beam)	12.5 keV 40 µm × 30 µm	<b>1.3, 1.8, 2.6</b>
P19-siRNA-1 (200 µm × 25 µm)	12.5 keV 40 µm × 30 µm	<b>2.5, 3.6, 8.0</b>
<b>Warkentin &amp; Thorne</b>		
Thaumatococcus (> beam)	12.6 keV, 13.5 keV 100 µm × 100 µm	<b>(1.80, 1.91)</b> , (1.91, 2.06), (2.06, 2.26), (2.26, 2.59), <b>(2.59, 3.27)</b>
<b>Liebschner <i>et al.</i></b>		
Thaumatococcus (99 µm × 145 µm × 120 µm)	12.66 keV 51 µm × 29 µm	(1.22, 1.30), <b>(1.30, 1.40)</b> , (1.40, 1.53), (1.53, 1.73), (1.73, 2.08), (2.08, 2.45), <b>(2.45, 3.10)</b> , (3.10, 30.0)

In order to present an overview of the meta-analysis that allowed a visual comparison of sample type, study, and resolution, the ten data sets shown in bold in table 5.2 were plotted in fig. 5.7, alongside trend lines for Model A and Model H. The shells were chosen to provide a coverage of both high and low resolution ranges used across the experiments.

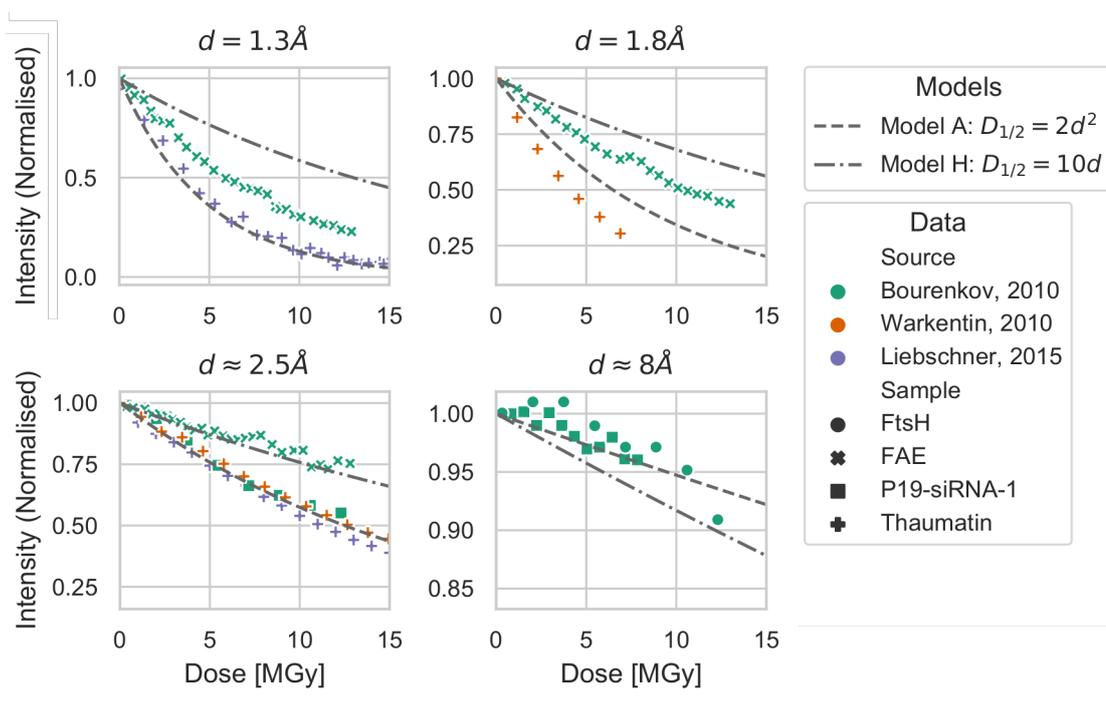


Figure 5.7: **Shell intensity data, at several resolutions, compared to Model A and Model H.** The plots include 10 data sets, selected from 22 included in our meta-analysis. Four protein samples (FtsH (circles), FAE (crosses), P19-siRNA-1 (squares), and thaumatin (plus-signs)) are included from three studies: Bourenkov & Popov (teal) [141], Warkentin & Thorne (orange) [142], and Liebschner *et al.* (lavender) [75]. The data sets used are grouped by resolution shell. Trend lines are for Model A and Model H. The two lower plots cover a small range of resolutions. In the lower left  $d \in [2.45, 2.6]\text{\AA}$ , in the lower right,  $d \in [7.5, 8.0]\text{\AA}$ . For clarity, only every 5<sup>th</sup> data point was plotted. 6 of the 10 data sets are well described by Model A, one is well described by Model H, the remaining 3 fall within a factor of 2 of Model A.

Across the resolution range shown in fig. 5.7, Model A shows very good agreement with the majority of the data sets, while Model H predicts slower fading at high resolution, and faster fading at 8 Å. The lowest resolution data (FtsH at 7.5 Å and P19-siRNA-1 at 8 Å, both from Bourenkov and Popov) are included as the trends across the entire meta-analysis—as well as the analysis of our own data sets—showed a degree of resolution-dependence in the accuracy of the models. However, it should be noted that the total doses received by these crystals were well below the anticipated shell half doses for such low resolution shells and, consequently, it is not possible to see a reliable trend in the data. There are a few data sets in fig. 5.7 that are not well described by Model A: Warkentin and Thorne, 1.8 Å data fade more quickly than both models predict; Bourenkov and Popov, FAE data at 1.3 Å and 1.8 Å sit between the two predictions; and FAE data at 2.6 Å is well-described by Model H. There are four different samples represented in the meta analysis. The variation in sample does not, among the limited data here, suggest a

sample-dependent trend in terms of matching one or other of the models used. The thaumatin data from Warkentin and Thorne, for example, does not match Model A at 1.8 Å, but it does at 2.59 Å; the thaumatin data of Liebschner *et al.* matches Model A at both 1.30 Å and 2.45 Å. The FAE data of Bourenkov and Popov sits between the two models at 1.3 Å and 1.8 Å, but matches Model H at 2.6 Å. The P19-siRNA-1 data of Bourenkov and Popov matches Model A well at 2.5 Å, and both samples (P19-siRNA-1 and FtsH) in the low resolution data seem to match Model A better than Model H.

## 5.5 Discussion

### 5.5.1 Setting $\kappa$ and $\alpha$ in Model A

The Howells-Holton model ( $D_{1/2}(d) \propto d$ ) [71, 87] has fixed coefficients and is suitable as a general predictive model of spot fading. The coefficients of the Atakisi model ( $D_{1/2}(d) \propto d^\alpha$ , with  $\alpha > 1$ ) [86] are determined by fitting intensities for a given data set, more suited for accurately describing a single case than predicting the general case. In order to compare the two models against experimental data, they should have the same form. To facilitate a comparison, we adapted the Atakisi model to produce a general prediction, as the Howells-Holton model does. This was accomplished by introducing fixed coefficients to the Atakisi model, as described in section 5.3.1. The fixed coefficient in the Howells-Holton model is the *Howells criterion*,  $H = 10 \text{ MGy } \text{Å}^{-1}$  [87], which was based on a meta-analysis of spot fading results. We took the same approach in order to fix the coefficients,  $\kappa$  and  $\alpha$ , in eq. (5.7): we reviewed the studies included in the meta-analysis by Atakisi, examining the fits they obtained and their conclusions regarding shell half dose, which they refer to as *local half dose*. Based on their meta-analysis, they suggested fits for eq. (2.19) with  $\alpha \in [1.4, 1.86]$ , each value optimised for a different data set.  $K$  was not specified, but is assumed to be constrained in some way, in order to obtain an unambiguous value for  $\alpha$ . Generalising their results, they proposed shell half doses of 2 MGy to 3 MGy at 1 Å, 8 MGy at 2 Å, 16 MGy at 3 Å, and 30 MGy at 4 Å.

Based on this, we fit eq (5.7) to the data of Liebschner *et al.*, with  $\alpha = 1.7$ , and compared the resulting values of  $\kappa$  against the general shell half doses proposed by Atakisi *et al.*. We found excellent agreement when  $\kappa = 2d^{0.25}$ , i.e. when  $D_{1/2}(d) = 2d^{1.95}$ . We repeated this approach for the high-resolution results ( $d < 5 \text{ Å}$ ) summarised in fig. 7 of [86], setting  $\alpha = 1.86$  as per Atakisi *et al.*'s best fit, and again obtained values for  $\kappa$  that were very closely approximated by taking  $D_{1/2} = 2d^2$ . We re-plotted these results in fig. 5.2, using linear axes for both resolution,  $d$ , and shell half dose,  $D_{1/2}$ . In their plot Atakisi *et al.* used log axes in both cases; this was necessary as they were plotting shells with resolution between 1 Å and 100 Å, and, correspondingly, shell half

doses between a few MGy and several 100s of MGy. With a much smaller range of resolutions, linear axes are possible, and fig. 5.2 highlights two things:

1. The variance in half dose as a function of resolution in the experimental results;
2. How close to each other fit-curves with different values of  $\alpha$  can be. We included the best-fit curve we obtained ( $D_{1/2}(d) = 3.5d^{1.6}$ ) and the curve generated from the final parameters we chose to use ( $D_{1/2}(d) = 2d^2$ ); they are very similar to each other. We also obtained a fitted-curve of  $D_{1/2}(d) = 2.22d^{1.86}$  using the  $\alpha$  value (1.86) reported by Atakisi *et al.*, it was found to be very close to both the best fit and the curve for Model A, we therefore excluded it from the graph for clarity.

The best fit we obtained has an exponent of  $\alpha = 1.6$ , while Atakisi *et al.* reported a best fit exponent of  $\alpha = 1.86$ . From their graph (fig. 7 in [86]), this fit seems to include results with resolutions down to 10 Å, while our fit is based on results down to only 5 Å. This difference in resolution cut-off is likely sufficient to explain the difference in best fit; as noted in point 2, above, the variation in fit curves with these different exponents is actually very small in the resolution range  $d \in [1 \text{ Å}, 5 \text{ Å}]$ .

Through the course of our analysis, we determined that there is no utility, as far as predictive modelling goes, in specifying the exponent,  $\alpha$ , to better-than-integer precision. The same goes for the multiplicative coefficient,  $\kappa$ . The difference between curves with  $\alpha = 2$  and  $\alpha = 1.6$  is well within the spread of experimental values used to derive these curves, as seen in fig. 5.2, so any added precision is below the error threshold. Selecting  $\alpha = 2$  tends to predict a slightly lower spot half dose than shown in most of the experimental data, so it provides a conservative prediction and a given sample is more likely to have a longer lifetime in the beam rather than shorter, which, in the context of experiment planning is preferable. Therefore the values of  $\alpha = 2$ ,  $\kappa = 2$  were acceptable to analyse our data and should be acceptable for general predictive use for a range of samples at resolutions between 1 Å and 5 Å, pending the accumulation of experimental evidence to the contrary.

### 5.5.2 Comparison of Model A & Model H in Reference to Our Experimental Data Sets

Having assigned fixed coefficient to Model A, we now had two models in the same form that could be directly compared to each other, with reference to experimental data. In the previous chapter, we described experiments that measured spot fading of lysozyme crystals at 12.66 keV and 26.00 keV. We chose to use the spot fading results from the latter of our two experiments (at NSLS-II) as an initial test data set for comparing the two models, Model A and Model H.

The main objective was to see whether one of these models offered a more accurate prediction of experimental spot fading.

The loss of intensity (by resolution shell) in our experimental data is presented in figs. 5.3 (for measurements at 12.66 keV) and 5.5 (for measurements at 26.00 keV), with trend lines showing the predictions of Model A and Model H, and a best-fit exponential curve. A notable feature of both sets of our experimental data is the high degree of variance in the results. This variance is greater in the 26.00 keV data set than the 12.66 keV data set and greater, within both data sets, at higher resolutions. The crystals used in our experiment were much smaller than those used in previous spot fading experiments and hence the associated Bragg peaks were not particularly strong. Bragg peaks were also relatively sparse and we had few data sets available, which works against the capacity of averaging to smooth out variation from peak to peak. As a result the data typically had a lower SNR compared to the larger crystals used for conventional MX. However, the mean is well described by the best-fit trend line.

Among the 12.66 keV data, the best-fit line is consistently positioned between the two models until the last set, corresponding to the lowest resolution shell (5 Å to 10 Å). In all cases, the half dose of the best fit is within a factor of two of both models, indicating that both are able to predict intensity loss with reasonable accuracy. Model A consistently predicts faster spot fading than the mean (and Model H, with the exception of the lowest resolution shells). For the higher resolution shells (with upper edges at 2 Å and 2.5 Å) Model A shows steeper spot fading than almost all the data sets collected, with the exception of some outliers. This provides a good estimate of the minimum life-time of the crystals and, as such, it is a 'better' model to employ in experiment planning. If a conservative estimate is made and more sample is available than is required to obtain a full data set, the experiment is still a success; if an optimistic estimate is made and less sample than required is available, the structure cannot be solved. Model H consistently predicts slower spot fading than observed in at least half of the experimental data.

In addition to showing a greater spread of intensity data than the 12.66 keV data sets, the 26.00 keV data sets consistently show a steeper best-fit line, indicating faster spot fading, compared to the predictions of both models. In these cases, Model A universally shows better agreement with the mean indicated by the best-fit, though again both models predict shell half doses within a factor of 2 of the best-fit curve.

While our data would certainly recommend Model A over Model H as being more likely to provide a conservative estimate of spot fading, to conclusively determine which is the most accurate model, A or H, additional data are required. Shot noise and other sources of interference are always present and models should allow for this. Such noise speaks to the notion that being

over-precise in the selection of coefficients (e.g.  $\alpha$ ) is not likely to be useful in the general case. Due to the limited availability of data and the lower SNR which results from the smaller crystal size, we also analysed previously published experiments, which allowed a more definitive comparison of the two models. The results of this meta analysis are discussed in the next section.

### 5.5.3 Comparison of Model A & Model H in Reference to Prior Spot Fading & Shell Half Dose Studies

An analysis of the effectiveness of predictive models to cope with variations in experimental parameters and data quality was carried out by analysing previously published damage studies covering several different samples and experimental set ups. This meta-analysis included both a detailed reanalysis of individual spot fading experiments and wider-examination based on the shell half dose metric.

We reanalysed the results of spot fading experiments by Bourenkov and Popov [141], Warkentin and Thorne [142], and Liebschner *et al.* [75]. These results are described in section 5.4.4. A summary of the results can be seen in fig. 5.7, which plots a representative sample of the results at different resolutions against Model A and Model H. A comparison of both models to all the results from the three studies is shown in figs. B.1 – B.5 in appendix B. With the exception of the lower-resolution (7.5 Å to 8 Å) data of Bourenkov and Popov, all the experimental data plotted in fig 5.7 exhibit clear exponential decay. Compared to the data collected from micro-crystals (figs. 5.3 and 5.5), the data in from conventional MX generally had better SNR. However there was significant variation between the results of the different studies, within a given resolution range. For the most part, the exponential decay is better described by Model A than by Model H. In one case (Warkentin and Thorne [142],  $d = 1.8$  Å) the observed spot fading was even steeper than predicted by Model A. Model H only provided a good description of some of the FAE data of Bourenkov and Popov, with most of their results falling somewhere between Model A and Model H.

This meta-analysis includes four different samples and it is unsurprising that there is some variation in spot fading results between them. While it has been observed [71, 119, 145] that change in B-factors (another global damage metric) are similar across several different proteins, site-specific damage is much more variable [72, 146, 147]. Site specific damage alters the ratio between individual Bragg peaks in a way which is not accounted for in global damage models. There is even variation in the rate of spot fading between the two studies that used thaumatin [75, 142]. There were several minor variations in experimental protocols that may account for this difference, including different sample-preparation, and beam configurations. However, for

all resolution shells, both studies using thaumatin had results within a factor of two of the predictions of Model A (The half doses of Liebschner *et al.* were 10–20% greater than Model A, those of Warkentin and Thorne were 30–40% lower). It has been observed [71] that several experimental parameters, including crystal volume and beam flux, often have error-bars that correspond to a factor of two difference in observed damage metrics, so anything within this range may be taken as accurately described by the model. The majority of the data of Bourenkov and Popov also fits within this factor of two. The main requirement of a predictive model such as Model A and Model H is that it is broadly applicable. Model A meets that requirement for the studies analysed here: several results deviate from its predictions, but all are well within a factor of 2 difference.

In reference to the data from our NSLS-II experiments, we found that Model A was preferred based on its provision of a more conservative estimate of crystal life time. However, neither model could really be said to be a ‘better’ description of the experimental spot fading than the other. However, in reference to the three studies in our meta-analysis, Model A is preferred as a more accurate predictor of spot fading. Ignoring the low resolution (7.5 Å and 8 Å) results, which do not fade sufficiently to observe a reliable trend, half the data sets seen in fig. 5.7 show spot fading that very closely matches the curves of Model A. Only one of the data sets follows the curve of Model H. The remaining 3 data sets differ from both curves, but have half doses that are closer to the predictions of Model A (differing by 40–60%) than model H (differing by 70–300%). Model A again provides a conservative estimate of crystal life time; only the data of Warkentin and Thorn [142] fades more quickly than the predictions of Model A. All of the data we have analysed then points to Model A providing a more useful prediction of spot fading than model H for the high resolution ( $d < 5$  Å) Bragg spots likely to be of interest in MX.

### **The Impact of Non-Uniform Illumination on the Accuracy of Spot Fading Models**

Modern MX experiments are often performed with small ( $\sim 1$  μm) crystals and Gaussian-profile beams, and invariably involve rotation of the crystal. These characteristics lead to non-uniform illumination of the sample, which results in non-uniform damage. It has been observed [66, 76, 86] that such conditions can result in an incorrect estimation of the effective dose state of the crystal, that, if not accounted for, can result in a misunderstanding of the relationship between ‘true dose’ and diffracted intensity. The basic idea is that if part of a crystal is not contributing significantly to diffraction, the amount of dose it has absorbed is not particularly relevant and shouldn’t be considered when expressing diffracted intensity as a function of dose. This gives rise to the notion of an *effective dose*,  $D_{Eff} \leq D$ , where dose is weighted relative to the strength of the diffracted signal from the same region of a crystal. The idea is illustrated by Warkentin *et*

al. [76] with the following example: consider a perfect top-hot beam with a square profile and a crystal that is larger than the beam such that only part of it is illuminated. The illuminated crystal region is exposed until diffraction is no longer observed. The beam is then translated by half its width to illuminate a new region, half of which has already been sufficiently exposed that it does not diffract at all, and half of which is fresh sample, having absorbed no dose. The average dose to the whole illuminated region is now half the initial dose delivered to the now non-diffracting region, however the *effective dose*—that is, the dose relative to the diffracted signal—is zero, because *all* the diffraction is now coming from a region of the crystal that has received no dose up until this moment. In real-world experiments, the effect is continuous rather than discrete. The contribution to diffraction of a given region of the crystal is affected by two factors: the intensity of illumination and the amount of dose it has absorbed. Given an undamaged crystal in a Gaussian beam, the central region will initially contribute more strongly to diffraction than the beam tails. Given two regions of a crystal with different dose-states, exposed to a uniform flux, the less damaged region will diffract more strongly than the more damaged region. The complexity of the situation is that the two factors affecting diffracted intensity are not independent. As the crystal is exposed, the majority of diffraction initially comes from the regions exposed to the centre of the beam, but those regions also accrue damage most quickly, and later in a high-dose (i.e. comparable to the half dose) exposure, the majority of the diffraction comes from regions in the fringes of the beam, that are much less damaged than the average dose state of the crystal would tend to indicate (see, for example, our discussion of the evolution of diffraction volume in Chapter 3, which is in agreement with the observations of Warkentin *et al.* [76]). The situation is more complex when rotation is taken into account, as the illumination of a given region of the crystal is different at different moments in time as well as different positions in space, but, in general, there will still be, at any point in the exposure, a distribution of local dose states across the crystal, and therefore a distribution of diffraction contributions. If this difference is not accounted for, it may seem that diffracted intensity fades more slowly, as a function of dose, than is really the case. Therefore, accounting for non-uniform illumination could be important in the context of accurately and precisely describing the loss of intensity in an experiment.

RADDOSE-3D [66, 67] does partially allow for the situation described above in its dose estimates. It produces several dose metrics, one of which is the so-called *diffraction-weighted dose* (DWD) [66], intended to address the impact of non-uniform illumination on the dose–diffraction relationship. RADDOSE-3D’s DWD is calculated as

$$DWD(t) = \frac{\int_0^t \int_V D(\mathbf{r}, t) F(\mathbf{r}, t) d^3r dt}{\int_0^t \int_V F(\mathbf{r}, t) d^3r dt}, \quad (5.8)$$

where  $D(\mathbf{r}, t)$  is the absorbed dose,  $F(\mathbf{r}, t)$  is the incident flux-density, and the time and space integrals are taken over the period and volume of interest, typically a whole exposure and the entire crystal region. It has been pointed out [76] that this metric is actually weighted by *incident flux-density* not diffraction. It therefore goes only half-way to addressing the impact of non-uniform illumination: diffraction *is* proportional to incident flux-density, but it is also affected by the local damage state of the crystal, which is not accounted for in RADDOSE-3D's DWD. As is, the DWD estimate produced by RADDOSE-3D should provide a good indication of the dose-diffraction relationship for low dose-states, but as the exposure progresses and more diffraction comes from the region of *lower incident flux-density* and therefore lower dose, the problem described above will remain. Warkentin *et al.* [76] proposed a revised version of the metric by adding a term,  $S(D(\mathbf{r}, t))$ , which is the diffracted flux per unit illuminated volume per incident flux-density. Equation (5.8) then becomes

$$DWD(t) = \frac{\int_0^t \int_V D(\mathbf{r}, t) S(D(\mathbf{r}, t)) F(\mathbf{r}, t) d^3r dt}{\int_0^t \int_V S(D(\mathbf{r}, t)) F(\mathbf{r}, t) d^3r dt}. \quad (5.9)$$

This dependence on the diffracting power of the crystal is not currently incorporated into RADDOSE-3D. Inclusion of this parameter would likely result in dose predictions that allowed a more accurate description of spot fading, especially when the total dose is very high, relative to the half dose. It does not *necessarily* follow, however, that the impact of non-uniform illumination must be accounted for in a predictive model. In our meta-analysis, we determined that the effect of its inclusion or exclusion did not make a significant difference in prediction of half dose. Therefore, whether a particular data set was better described by Model A or Model H was independent of whether its illumination conditions had been accounted for in determining dose.

We compared shell half dose data from four studies [75, 78, 141, 142], by using WebPlotDigitizer [144] to extract data from two sources (per study):

1. The spot fading graphs from the original publications;
2. The graph of shell half doses as a function of resolution in the meta-analysis by Atakisi *et al.*.

In the latter case, Atakisi *et al.* modelled, in detail, the experimental conditions of the original data collection, and recalculated the doses to account for the specific illumination conditions of each study, so their half dose values account for non-uniform illumination. In the former case, we used the doses as they appeared on the graphs of the original sources, and did not adjust them to account for non-uniform illumination. We obtained exponential best-fit curves to the experimental data and used those fit curves to determine the shell half doses. For all data sets

across the four studies, the shell half dose values we obtained via the two methods were within 12%, and the majority were within 2%. So if the goal is to get a coarse predictive measure of the life-time of a crystal, in the form of half dose, a precise model of the non-uniformity of the illumination is probably not required. In our current investigation, the goal is to determine whether this coarse prediction is better provided by Model A or Model H and in this case also, factoring in non-uniform illumination does not seem to be important. If the goal is to describe, as precisely as possible, the loss of intensity in Bragg spots, so that an attempt can be made to reconstruct their undamaged relative intensities, then the illumination conditions should be taken into account, as the intensities in the high-dose tail-end of the data sets are likely to be strongly affected by this non-uniformity and the resulting change in the dose–diffraction relationship. The impact of illumination conditions on the dose-state of a crystal and the resulting damage is also no doubt of importance in the development of a detailed understanding of radiation damage in modern MX experiments, which is another goal in studying spot fading. Experiment design should take into account the potential impact of the beam-profile, however its inclusion in a general predictive model of spot fading does not seem either necessary or desirable.

### Model H in Prior Studies of Low Resolution Bragg Spots

The empirical spot fading model referred to in this chapter as either the *Howells-Holton model* or *Model H*, was originally proposed by Holton [71], based on the findings of Howells *et al.* [87]. In their investigation of the resolution-dependence of spot fading as a function of dose, Howells *et al.* were interested in the application of x-ray diffraction microscopy (XDM) to biological samples, where radiation tolerance was expected to be the limiting factor to attainable resolution. Their resolution range of interest was 30 Å to 100 Å, one to two orders of magnitude lower than the resolutions of interest in macromolecular crystallography (MX) (1 Å to 2 Å), where anything lower than around 30 Å is typically obscured by the beamstop and/or the transmitted radiation. In other words there is little to no overlap between the resolution ranges Howells *et al.* were most interested in and those typically of interest in MX. Nevertheless, in their meta-analysis of spot fading information, they included results from electron and x-ray diffraction methods, in addition to XDM, covering a resolution range of 2 Å to 600 Å. Only one third of the results analysed involved diffraction spots with resolutions higher than 5 Å. Among these higher resolution results, the maximum tolerable dose—as they referred to the half dose—varied between 5 MGy and almost 80 MGy. With this degree of variation, many of the results would fall some distance from *any* fit. The lower resolution results showed a more consistent linear relationship between resolution and tolerable dose, with a gradient between 6 MGy Å<sup>-1</sup> and 10 MGy Å<sup>-1</sup>. The Howells criterion was selected based on these low resolution results. Extrapolated to higher resolutions,

the results fell, in approximately equal measure, on either side of this slope. So the model chosen provided a good description of low resolution results and of the mean of a very noisy selection of high resolution results. This makes it well suited to its original target, XDM, and for resolutions around 2 Å, it produces predictions that were similar to other dose-limit recommendations by e.g. Henderson (20 MGy) [16] and Owen (30 MGy) [119].

Coughlan *et al.* [28] measured spot fading for several low resolution Bragg peaks (13 Å to 17 Å) of lysozyme crystals, up to very high doses (100s of MGy) and, like Howells *et al.*, observed spot half doses consistent with Model H. Also like previous results, there was a reasonable degree of scatter about the linear fit, suggesting that an increase in the precision of the model could have no utility in addressing experimental data collected from micro-crystals, where a degree of noise is to be expected. Nave *et al.* [143] observed a spot half dose of 120 MGy for a 12.6 Å Bragg spot and 300 MGy for a 35 Å spot, both results falling slightly shy of the 10 MGy Å<sup>-1</sup> of the Howells criterion. At this stage, there is a reasonable body of evidence regarding the loss of intensity of individual Bragg spots and/or resolution shells as a function of dose. The general consensus, across disciplines where radiation-sensitivity is a limiting factor in imaging, is that the Howells-Holton model provides a good estimate of the tolerable dose for biological samples at lower resolutions, but is not as accurate at higher resolutions ( $d < 5$  Å). In comparisons to model A, plotted curves show that the two trend lines cross at 5 Å, so it is perhaps not surprising that this seems to be the cutoff for one model providing a better prediction than the other. Beyond 5 Å, Model A describes shallower spot fading than has been observed in experiments [26, 27, 87, 119, 143]. In general, the relationship between resolution and dose seems to be itself resolution-dependent, being roughly proportional to the square of resolution at higher resolutions and linear at lower resolutions.

### Non-Exponential Trends in Spot Fading

Both models, A and H, predict purely exponential loss of intensity in Bragg spots / resolution shells (and integrated intensity, for that matter). Purely exponential behaviour is in good agreement with experimental evidence [71, 76, 86]. However, there is also evidence that the intensity loss of individual Bragg peaks and, by extension, resolution shells and integrated intensity, can depart significantly from this exponential trend [76, 83, 148, 149]. Here we discuss:

- The non-exponential trends we saw in our data, with reference to previous observations of the same trends, and possible explanations of them;
- The implications of non-exponential behaviour for using purely exponential models to predict spot fading.

We previously discussed, in section 5.5.3, how non-uniform illumination can lead to apparent non-exponential fading, but this is an artefact of the way the dose is estimated. If the dose estimate is weighted by diffraction contribution, the behaviour is purely exponential. We will not revisit this variety of non-exponential behaviour here.

In our data sets, non-exponential behaviour was observed in a significant number of peaks, which translated to non-exponential behaviour in resolution shells. The primary trend observed was that peak intensities initially increased in diffracted intensity prior to decaying in an exponential fashion. When this behaviour was observed in one or more bright peaks, it led to bumps or plateaus in the summed (or averaged) intensities of that resolution shell. This behaviour was particularly prevalent in the highest ( $\leq 2 \text{ \AA}$ ) and lowest ( $\geq 10 \text{ \AA}$ ) resolution ranges. Where possible, we excluded peaks from our analysis that didn't allow for a reasonable fit to an exponential curve. However, these trends were very common in our data and we could not entirely exclude them, so we included some peaks that showed an initial increase when a reasonable fit could still be obtained for the summed/average resolution shell data.

Owen *et al.* [148, 149] observed a similar non-exponential effect in crystals of thaumatin, bovine enterovirus serotype 2, and foot and mouth disease virus imaged at room temperature. The integrated intensities initially exhibited a plateau of approximately 0.5 MGy, before decaying exponentially. They described this plateau as a 'lag phase', indicating a brief lull prior to the onset of damage. They hypothesized that the lag phase was due to delayed diffusion of radical species at high dose rates in room temperature MX. However, radical diffusion is impeded by cryogenic cooling, and this lag phase is not seen universally in cryo MX. Teng and Moffat [78] observed similar results at 100 K. Their results seem to exhibit behaviour which was counter to the typical exponential behaviour: fading is initially very slow (up to about 8 MGy) and then becomes steeper up to a total dose of 16 MGy. However, this total dose is very low and was only sufficient to reduce a few of the highest resolution shells to half intensity. It is possible that an exponential trend may have emerged over a higher dose range, which would match the behaviour observed by Owen *et al.*. Teng and Moffat also plotted other global damage metrics: unit cell expansion and Wilson B-factors. Both showed immediate evidence of damage, precluding the interpretation of slow spot fading as a lag in the onset of damage. A possible interpretation of all these data is that expansion of the unit cell may cause sufficient change in the Bragg condition to partially offset the inherent loss of spot intensity due to radiation damage. A characteristic of imperfect, finite crystals is that Bragg peaks have finite dimensions in reciprocal space, they are not perfect geometric points. This property will be significant in micro-crystallography where the diffraction region is often on the order of a few square-microns. In this case, the measured

intensity of a Bragg peak will be due to its actual total intensity, and also the extent of its intersection with the Ewald sphere (which also has finite thickness, owing to the bandwidth of the beam). If a peak does not initially perfectly intersect the Ewald sphere, a slight change in the Bragg condition may mean that though its total intensity is reduced with dose, its measured intensity initially increases as its alignment with the Ewald sphere improves. In this case, an increase in measured intensity is indicative of radiation damage.

This change in the Bragg condition, driven by unit cell expansion, or, more generally, a 'redistribution of electron density in reciprocal space', has also been proposed by Warkentin *et al.* [76] as a factor in the non-exponential behaviour of some Bragg peaks. They observed that the likelihood of reciprocal lattice points being positioned such as to cause Bragg peaks to initially increase in intensity was influenced by crystal orientation in the beam, which will be more noticeable when crystals are not rotated (or are rotated very little), and is likely to be more problematic in serial synchrotron crystallography. These conditions mean that the rate of mosaicity and unit cell change will be larger relative to the native drop in Bragg peak intensity due to radiation damage. In our experiments, we used flash-cooled polymer chips to mount crystals, and it is possible that a preferential alignment of the crystals on the chips may be responsible for the high prevalence in our data sets of peaks showing early increases in intensity. During the experiment, the volume of reciprocal space sampled was too small to determine if preferential alignment was occurring. However, preferential alignment was observed when the crystals were mounted on TEM grids (see section 4.2.1); it is possible they were similarly aligned on the polymer chips used at NSLS-II. Warkentin *et al.* also suggested that this non-exponential behaviour might be more obvious when the beam was smaller than the crystal as undamaged regions of the sample might constrain the expansion of the damaged region. This was not the situation for our data, as the crystals and the beam were approximately the same size, though it is possible that a similar effect was present due to the more rapid damaging of the centre of the sample, exposed to the most intense part of the Gaussian beam.

It is also possible that the same effect might manifest the converse behaviour: a change in the Bragg condition may cause a peak's intersection with the Ewald sphere to worsen, causing a sharper than expected drop in measured intensity. This steep drop off was also observed in several peaks in our data sets, though it was not as prevalent as an early rise in intensity. With well-diffracting crystals providing a large number of Bragg peaks in each frame, the contrary fading of peaks might be expected to average out when looking at integrated intensities or resolution shells, and the effected peaks can be excluded from individual analysis. Though in the case where the number of peaks is limited, as it was in our data sets, this rapidly means the discarding of the entire data set, especially if a preferential orientation has resulted in a high

prevalence of non-exponential behaviour.

In the context of the current work, the question is how to deal with these non-exponential trends when using an exponential model to predict spot fading. A reasonable answer is to ignore it: the models are intended to make broadly applicable predictions of the rate of spot fading and it is expected that experimental results will not match them precisely. In our data, we observed reasonably good exponential fits even for resolution shells that showed low-dose non-exponential behaviour and Warkentin *et al.* [76] observed that when the initial rises/plateaus are excluded from fitting, the half doses obtained are similar to those of peaks/shells that show purely exponential fading. Non-exponential decay (of individual Bragg peaks, resolution shells, and integrated intensities) offers an opportunity for interesting insights into the damage mechanisms underlying spot fading and other damage metrics. Future models intended to precisely describe, rather than predict, damage metrics could benefit from a detailed study of the conditions under which intensity plateaus are observed, and how these plateaus relate to other damage metrics like unit cell expansion. However, in the context of the present Chapter, non-exponential behaviour does not impact the question of whether one exponential model or another (e.g. Model A vs Model H) is a better predictor of *typical* spot fading behaviour.

#### 5.5.4 Model H & Integrated Half Dose Values

We observed (see figs. 5.4 and 5.6) that Model H provided a reasonable fit to our integrated intensity data, when  $d$  was taken to be the maximum resolution to which the undamaged crystal diffracted,  $d_{MAX}(0)$ . While this is not the interpretation of the model that was originally intended [71, 87] (it purported to model the intensity loss of a Bragg spot (or spots) of a given resolution), it is an interpretation that fits a reasonable body of results reported in the literature. Teng and Moffat [79] measured loss of integrated intensity for lysozyme crystals initially diffracting to 1.6 Å that is well described by Model H, if  $d$  is interpreted in this way. The same is true for integrated intensities measured by Liebschner *et al.* [75] for thaumatin crystals initially diffracting to 1.22 Å. We have replotted the data of these two studies (we have normalised the intensities) in fig. 5.8. Alongside the experimental data, we have plotted Model H, with  $d = d_{MAX}(0)$ . In both cases Model H shows excellent agreement to the measured integrated intensities.

Model H can be compared to a larger body of past results with reference to the half dose,  $D_{1/2}$ , metric, which is often reported, even when detailed intensity vs. dose data is absent. Table 5.3 summarises the relationship between  $d_{MAX}(0)$  and  $D_{1/2}$ , for integrated intensity measurements reported in the literature over the last two decades (including the two studies highlighted in fig. 5.8), comparing results to the predictions of Model H.

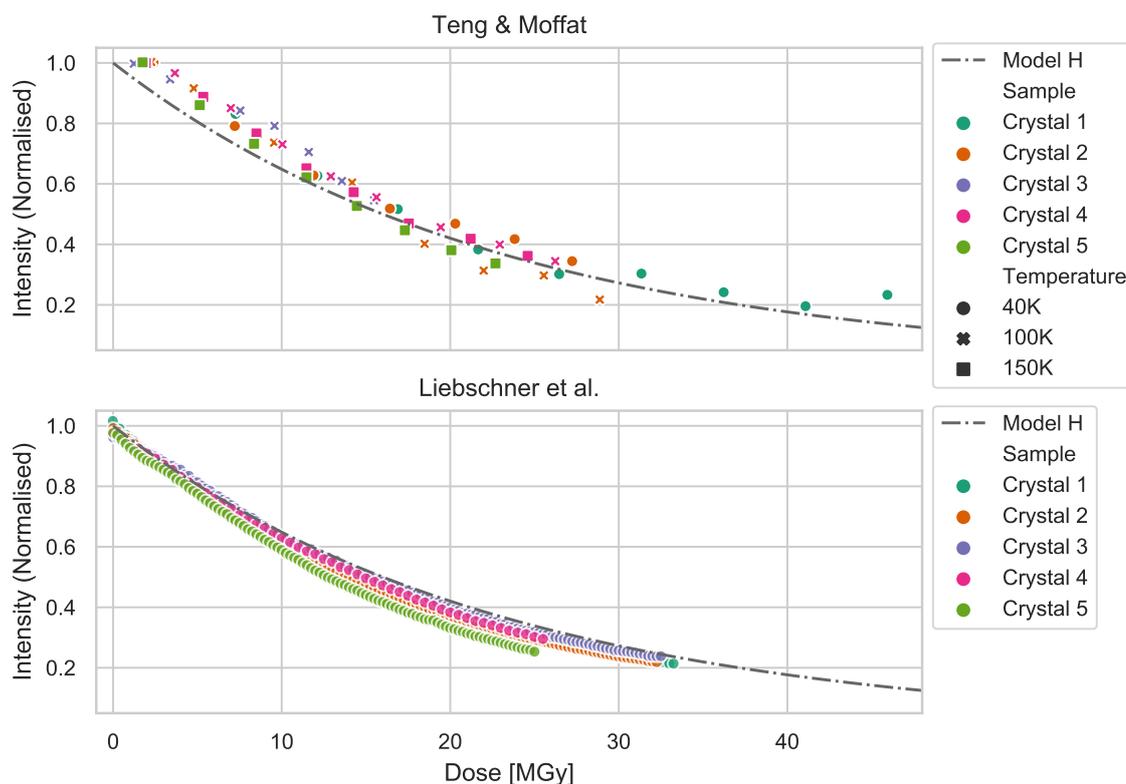


Figure 5.8: **Loss of integrated intensity & the Howells-Holton model at high resolutions.** The experimental data ( $\bullet$ ,  $\blacksquare$ ,  $\times$ ) are for lysozyme crystals diffracting to  $1.6 \text{ \AA}$  from Teng & Moffat [53], and thaumatin crystals diffracting to  $1.22 \text{ \AA}$  from Liebschner *et al.* [75]. Both data sets are for the integrated intensity of the entire diffraction pattern. Both show excellent agreement with Model H.

All the results shown in table 5.3 were for cryogenic sample conditions with standard x-ray beam energies ( $E \approx 13 \text{ keV}$ ). A range of protein crystals were used, including lysozyme, apoferritin, holoferritin, and thaumatin. Across the range of maximum resolutions ( $1.22 \text{ \AA}$  to  $4 \text{ \AA}$ ), the half doses for integrated intensity show very good agreement with the prediction of Model H:  $D_{1/2} = 10d$ .

The indication of a sizeable collection of data then, is that the Howells-Holton model is reasonably accurate for predicting spot fading of low-resolution spots, while a predictive version of the Atakisi model shows much better agreement with individual spot fading at higher resolution. In the high resolution regime, however, it does seem that the Howells-Holton model retains utility in predicting the rate of integrated intensity loss. This is in line with our other findings, as discussed in sections 5.5.2 and 5.5.3, above.

Table 5.3: **Integrated half dose values & the Howells-Holton model of spot fading.** Half doses are shown for integrated intensities of several previous studies, alongside the prediction of the Howells-Holton model (Model H:  $D_{1/2} = 10d$ ), if  $d$  is taken to be  $d_{MAX}(0)$ , the maximum resolution to which the undamaged crystal diffracts. All results shown are for standard beam energies ( $E \approx 13$  keV) and cryo-cooled samples. Across the studies,  $d_{MAX}(0)$  is between 1.22 Å and 4 Å. The data show very good agreement with the Model H.

Source	$d_{MAX}(0)$ [Å]	$D_{1/2}$ [MGy]	
		Exp. Data	Model H
Liebschner <i>et al.</i> [75]	1.22	16.6	12.2
Warkentin <i>et al.</i> [76]	1.4	10	14
Warkentin <i>et al.</i> [76]	1.6	13	16
Teng & Moffat [78, 79]	1.6	17	16
De la Mora <i>et al.</i> [96]	1.9	15.3	19
Burmeister [72]	2.0	21	20
Liebschner <i>et al.</i> [75]	2.45	20.4	24.5
Liebschner <i>et al.</i> [75]	2.08	18.7	20.8
Owen <i>et al.</i> [119]	4.0	43	40

## 5.6 Summary

Models of intensity loss for individual Bragg peaks, resolution shells, and integrated intensities are an important component of the study of radiation damage. Models may be used to predict the rate at which Bragg peaks will fade, in order to estimate prior to an experiment how much sample should be available. They can also be used to help correct peak intensities to zero-dose equivalents, in order to extend the lifetime of samples, and as an aid in understanding the physical processes underlying radiation damage. Spot fading models have been employed with these purposes in mind since the early days of MX.

An important characteristic of spot fading is that it is resolution-dependent, with higher resolution peaks fading more quickly than lower resolution peaks. Here we have compared the predictions of two resolution-dependent spot fading models against intensity data collected at two x-ray energies (12.66 keV and 26.00 keV) and previously published results. Model H, due to Howells [87] and Holton [71], predicts spot fading proportional to  $\exp(-d^{-1})$ ; Model A, adapted from a non-predictive model by Atakisi [86], predicts spot fading proportional to  $\exp(-d^{-2})$ . Both models provided an adequate fit to our own data (at both energies) for resolution shells from 2 Å to 5 Å, with Model A preferred as it offers a more conservative estimate, a useful characteristic in a planning tool. In reference to our meta-analysis of previously published data,

Model A provides a superior fit for high resolution shells (up to 5 Å), while Model H provides a better fit for lower resolution data (down to 100 Å). Overall, it is expected that a good prediction of intensity loss for a shell with resolution  $d$ , will be obtained from Model A (eq. (5.7)). Interestingly, we found that Model H (eq. (2.18)) is a good indicator of loss of *integrated intensity* for diffraction data when  $d$  is taken as the maximum resolution of an undamaged crystal. For applications requiring a lower resolution than MX (for example, XDM, with a desired resolution of 30 Å to 100 Å), Model H is a good predictive model. A more general model would require that the resolution-dependence was itself resolution-dependent, but there is not necessarily any utility in such a model, given the accuracy of existing models within the kinds of resolution ranges typically desired in different disciplines.

Using modelling to correct measured peak intensities to their zero-dose values is a more complex task than predicting spot fading rates. Here, the priority is accurate and precise fitting to a specific set of data rather than the lower-precision, but widely-applicable approach of a predictive model. In this case, neither of the models examined here offers much beyond a starting point, though Model A is likely to provide a more fitting starting point for data collected at higher resolutions.

Significant deviation in spot fading from the norm of exponential decay has been observed by experimenters [76, 78, 83, 148, 149]. Non-exponential spot fading is not modelled by either Model A or Model H. This variation from typical average behaviour can itself provide, or at least provoke, insights into the mechanisms underlying the loss of measured intensity for individual spots, resolution shells, and whole diffraction patterns. Redistribution of electron densities in reciprocal space can result in an appearance of plateaus in spot fading that are nevertheless an indicator of damage, not the temporary avoidance of it. Non-uniform illumination can result in the diffraction contributions of different parts of a crystal changing over the course of a series of exposures and this requires a reassessment of the dose-state of the crystal, weighted by the diffraction contribution. Given the trend in MX toward micro-crystals and beams, this effect cannot be ignored in the analysis of diffraction data, though it may not be of critical importance in predicting spot fading unless particularly precise estimates are required ahead of time.

The meta-analysis and original results presented here cover a small range of samples, focused on commonly-used model proteins such as lysozyme and thaumatin. The range of crystal sizes is more significant, from the 2 µm micro-crystals of our own experiments to the 100 µm to 600 µm crystals used by Liebschner *et al.* [75] and Warkentin *et al.* [76]. A large range in radiation tolerance across protein targets exists and there is potential benefit to be had by extending the

range of samples that are used to validate spot fading models. Further experiments with micro-crystals, again using multiple protein samples, would help to quantify the peculiar characteristics of these crystals, relative to the larger crystals used in conventional MX. The fitting parameters we selected for Model A, that give the relationship  $D_{1/2} = 2d^2$ , make for a good description of the data analysed in this chapter. If a larger survey of spot fading for different protein targets were conducted, results could be grouped by radiation tolerance or other protein-protein similarities. This may reveal trends in spot fading at a finer detail than possible with only a handful of samples. It may be that different families of proteins require different coefficients to be used in modelling spot fading, i.e. metallo-proteins, which are highly radiation sensitive due to the increase cross-section of heavy atoms, are likely to show steeper spot fading than model proteins like lysozyme. These proteins will also be highly susceptible to specific damage, which will result in more pronounced intensity loss to high resolution Bragg spots.

We have focused here, as have previous researchers before us, on cryogenically cooled samples. The dose-dependence of spot fading has been shown to be independent of dose-rate under these conditions. This dose-rate independence is not seen for room temperature crystallography [82, 148, 150], which is experiencing a resurgence in popularity as a complement, or even alternative, to cryo-crystallography [81]. Room temperature results do, however, exhibit the same general exponential trend, it is simply that the half dose (considered globally or for a particular resolution) is both much smaller and dose-rate dependent [82, 150, 151]. A larger body of results, systematically addressing a variation in dose-rate, could be used to adapt Model A for use in predicting spot fading in room temperature crystallography. Given the much reduced sample lifetimes at room temperature as compared to cryogenic temperatures, a predictive model may be seen as even more important to experimental planning under such conditions.

## Chapter 6

# Thesis Summary & Further Work

The overarching motivation for understanding and investigating radiation damage is the desire to understand the function of proteins via atomic-resolution imaging of their structures. X-ray crystallography is a mature discipline that has been responsible for the majority of protein structures solved to date [5]. Despite its success, there still remains a significant amount of work to be done to further improve its utility. There are several fundamental bottle-necks in the process of structure solution via MX, including the difficulty growing large, well-ordered crystals, and the radiation damage that accrues during experiments, limiting the resolution at which proteins can be imaged. Technological progress at synchrotrons and XFELs has provided the means to extend the usable lifetime of protein crystals and thus enable MX experiments to be conducted with smaller and less well-ordered crystals. With these advancements come new challenges to designing and executing successful experiments and radiation damage remains a relevant and important issue.

In this thesis, we have investigated the progression of radiation damage in MX, with a particular focus on micron-scale crystals and the potential for PE escape to reduce the rate at which they are damaged. Our computational simulation of damage in micro-crystals produced dose predictions accounting for PE escape, which was not, at the time, included in RADDPOSE [66, 67], the dose prediction software most used in MX. The experiments we performed were among the first in the world to systematically test the prediction that performing MX experiments at high x-ray energies could result in a reduction in damage rates to micro-crystals. This thesis also contained an in depth review and comparison of spot fading models, including a meta-analysis of results available in the literature.

In this concluding chapter, a summary is provided of the work that was conducted and the results obtained. The summary is followed by recommendations for further work that could

help overcome some of the challenges, in the context of radiation damage, currently facing the field of structural biology.

## 6.1 Summary of Work & Results of This Thesis

### 6.1.1 Simulations Predict Significant Damage Reduction to Micro-crystals Due to Photoelectron Escape

In Chapter 3, we presented a discrete spatially- and temporally-resolved simulation of radiation damage that incorporated the effect of PE motion within a crystal and its surrounding solvent. The simulation was used to investigate the impact PE motion might have on the rate at which radiation damage accrues in a crystal due to variations in sample and beam size, and beam energy. Our simulation pipeline incorporated Monte Carlo modelling of electron trajectories via CASINO [106, 107], an application for modelling SEM experiments. Monte Carlo modelling has been previously used in simulations to explore PE motion in protein crystals. Some previous models have dealt with estimates of penetration depth [11, 12] without accounting for the characteristics of the beam, or a full spatially-resolved image of dose distribution, while others have made a partial model of the interaction region but conflated the x-ray beam with the electron beam modelled in CASINO [104, 111]. Our own simulation extends on the work of previous researchers by explicitly modelling the entire interaction region, and carefully separating the two-stage process of an x-ray photon interacting with a *K*-shell electron in the sample in order to produce a PE, and the subsequent motion, through the material, of that PE. Our simulation further builds on the spatially-resolved dose distribution to provide a temporally-resolved prediction of Bragg spot intensity loss during an x-ray exposure. To our knowledge, our simulation was the first example of a tool that incorporated PE motion into a detailed model of experimental dose distribution.

We executed our simulation with variations in beam size (1  $\mu\text{m}$  to 5  $\mu\text{m}$ ), beam energy (9 keV to 30 keV), and crystal size (1  $\mu\text{m}$  to 5  $\mu\text{m}$ ), in order to examine the impact of each of these parameters on the rate at which radiation damage accrues, with particular attention given to the difference in damage due to PE motion. For all combinations of parameters modelled, our results that did not include PE escape aligned well with those of RADDOSE-3D, with small variation due to our inclusion of Compton scattering. At standard MX energy (12 keV), the model of a 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  (FWHM) beam that included PE escape showed a 4-fold increase in lifetime over the model that did not account for PE escape. The difference between models that did and didn't incorporate PE escape diminished with increasing beam size. The difference was observed to be negligible at length scales larger than 5  $\mu\text{m}$ . This reduction was seen whether or not

the crystal size also increased with the beam size, as a large beam and smaller crystal resulted in scattering of PEs *into* the crystal from the surrounding solvent, offsetting the benefit of those that escaped the crystal. For a 1  $\mu\text{m}$  crystal in a 1  $\mu\text{m}$  beam, increasing the beam energy resulted in an increasing difference between results that incorporated PE escape and those that did not. The improvement seen due to PE escape was greatest between 9 keV and 20 keV, and continued to increase up to 30 keV, but at a reduced rate. By 20 keV much of the energy from PEs produced in the crystal has already escaped its bounds and any further increase in initial energy mainly serves to scatter the electrons even further into the solvent.

An overview of the results of our simulation that summarises the impact of all three variables—beam size, beam energy, and crystal size—can be seen in fig. 4 of the paper included in section 3.4, which shows the diffraction efficiency (intensity of coherently-scattered x-rays per unit absorbed dose) of several configurations, normalised to the diffraction efficiency of a crystal in a 12 keV beam, when PE motion is ignored. All the configurations modelled see an increase in diffraction efficiency with beam energy, with the maximum increase being a factor of around 30. A 1  $\mu\text{m}$  crystal in a larger beam (1.5  $\mu\text{m}$ , 2  $\mu\text{m}$  and 5  $\mu\text{m}$ ) saw a slightly better improvement in diffraction efficiency than a larger crystal matched to the beam size, suggesting that higher energy x-rays might be preferable even if solvent illumination cannot be entirely excluded.

The work on our simulation informed the design of subsequent chapters of this thesis. Based on the results and questions raised in the design and interpretation of the output, we conducted experiments to test the impact of x-ray beam energy on damage to micro-crystals (Chapter 4), and examined the assumptions of the models used to describe Bragg spot fading (Chapter 5).

### 6.1.2 Early Experiments Highlight Technical Challenges in Micro-crystallography at Higher Beam Energies

The two experiments presented in Chapter 4 were among the first to explicitly and systematically examine the relationship between beam energy and damage rates in micro-crystallography. The NSLS-II experiment was performed with the aid of a pre-market cadmium-telluride detector—the Eiger 1M CdTe. This was one of the first uses of this detector. The results of these experiments suggest the detector could be useful for analysing high-energy diffraction data collected from micro-crystals. The challenges faced also provide useful insight into the design and execution of future experiments.

Our first experiment was conducted at the P11 beamline of PETRA-III. We recorded diffraction data at six energies (12 keV, 18 keV, 20 keV, 22 keV, 24 keV and 26 keV) from 5  $\mu\text{m}$  crystals of

tetragonal lysozyme, mounted on silicon chips [54]. The experimental conditions at PETRA-III were not optimal, and useful diffraction data was not obtained. However, the issues encountered provided the opportunity to learn from this experiment and implement an improved experimental design at NSLS-II. In particular, the PETRA-III experiment highlighted the importance of a small beam focus size, and the technical challenge of achieving it; and the need to measure diffraction with a detector that maintains high quantum efficiency at high beam energy, such as the Eiger 1M CdTe. Such a detector was not available to us at the time of the PETRA-III experiment, but as they are now being brought to market, they should be more accessible in the future.

Our second experiment was conducted at the 17-ID-2 FMX beamline of NSLS-II. We recorded diffraction data at two energies (12.66 keV and 26.00 keV) from 2  $\mu\text{m}$  crystals of tetragonal lysozyme, mounted on optically transparent polymer chips designed and produced by the Perry lab. The detector was an Eiger 1M CdTe cadmium-telluride detector. The beam size was optimised to produce a 3  $\mu\text{m}$   $\times$  2  $\mu\text{m}$  focus at both energies used and a low-intensity raster scanning was used to identify well diffracting crystals on the chip. These factors addressed many of the issues faced at the PETRA-III experiment and resulted in an improvement in the quality of the diffraction data collected.

The NSLS-II experiment pushed the capacities of the beamline and required several interconnected systems to operate at their practical limits, in concert. The preliminary data obtained had a low SNR, but we did observe a reduced rate of spot fading at 26.00 keV compared to 12.66 keV, in keeping with the expectation of damage reduction via PE escape. Based on the spot fading rate of the experimental data at both energies and a comparison to spot fading results reported in the literature, we determined that the rate of intensity loss assumed in our initial modelling was too slow to accurately describe real-world data. Several challenges during beamtime further contributed to our understanding of the work required to make high energy MX experiments routinely viable. Some difficulties that arose during the experiment were:

- The chip design had to be adapted on site to withstand the force of the cryostream without movement.
- Beam configuration and focusing was still a time-intensive operation and needs to be optimised.
- Static measurements were not possible, and a minimal rotation (0.002°) had to be used between exposures, due to limitations in the data acquisition software.
- The photoionisation cross section of our solvent was higher than that of the crystal.

The results of our NSLS-II experiment are summarised in fig. 4.13. Due to the low diffraction spot counts and significant noise, we restricted our final analysis to the 10 brightest peaks at each energy that individually demonstrated more-or-less exponential fading with dose. Among these peaks, we observed a reduced rate of spot fading at 26.00 keV compared to 12.66 keV.

### 6.1.3 Meta-Analysis Demonstrates Capabilities & Limitations of Predictive Models of Radiation Damage

In the work presented in Chapter 5, we proposed a predictive model (Model A) based on Atakisi's assertion of a roughly  $d^2$  dependence for  $D_{1/2}$ , refining empirical constants based on a meta-analysis of recent spot fading results. We compared the predictive utility of the *Howells-Holton model* (Model H,  $D_{1/2} \propto d$ ) and our version of the *Atakisi model* against data captured in our experiments from Chapter 4 and several other data sets from the literature.

The micro-crystallography data we had available from our NSLS-II experiment had a low SNR and was restricted to a maximum resolution of 2 Å at 12.66 keV and 3 Å at 26.00 keV, with a minimum resolution of 10 Å at both energies. Both Model A and Model H provided a good description of the mean behaviour, differing from a line of best-fit by no more than a factor of 2 across all data sets at both energies. Compared to several previously-published low-noise data sets [75, 141, 142], Model A provided a better match to most of the data, however there were also several data sets that matched Model H better or sat somewhere between the two models, or showed steeper fading than either of the models predicted.

Interestingly, though Model A tended to provide a better fit to the fading of individual Bragg peaks (or groups of peaks in a single resolution shell) than Model H, the latter was found to provide a good fit to the fading of an entire diffraction pattern, when  $d$  was taken to be the maximum resolution of the undamaged diffraction pattern.

Overall, Model A was preferred over Model H as a better tool for predicting the fading of individual Bragg spots or resolution shells, particularly at higher resolutions. Model A predicts faster spot fading than Model H for  $d < 5$  Å so using Model A as a predictive tool will provide a more conservative estimate of the sample required for an experiment.

## 6.2 Suggestions for Further Work

During the course of this research a number of interesting follow up questions and directions for future investigations beyond the scope of the current project have emerged. Here we propose potential extensions of the research reported in this thesis.

### 6.2.1 Extended Experimental Study of Radiation Damage at High Energy

The results obtained at NSLS-II were promising and, combined with the insights regarding the technical challenges faced at beamtime, can provide a starting point for an in depth experimental study of the impact of beam energy on radiation damage. The initial goal of future experiments should be to probe a wider range of experimental targets and conditions. Further experiments will be made feasible by the reduced emittance of 4<sup>th</sup> generation synchrotrons, and the increasing availability of CdTe detectors, such as the Eiger 1M CdTe used in the NSLS experiment. Experiments like the ones reported in Chapter 4 should be repeated with different protein samples and different crystal sizes.

Thaumatococcus would be a logical choice for an initial sample to investigate further. Like lysozyme, thaumatococcus is a commonly used model protein with well-established crystallisation protocols [152], and micro-crystals (approximately  $1.5\ \mu\text{m} \times 3\ \mu\text{m}$ ) have previously been used for diffraction experiments [153]. High quality spot fading data are available for large crystals [75, 76, 142], as seen in the meta-analysis in Chapter 5; these could provide a point of comparison for data collected from thaumatococcus micro-crystals, further elucidating the specific radiation damage characteristics of micro-crystals. Also of interest is the spot fading of proteins that are particularly prone to site-specific damage, such as metalloproteins. Due to the impact of specific damage, certain Bragg spots will fade at a rate that is dependent not only on resolution, but azimuthal angle, an effect that is not currently accounted for in models of spot fading.

Crystals in the range  $1\ \mu\text{m}$  to  $5\ \mu\text{m}$  are of greatest interest initially. The simulations presented in Chapter 3 predict the greatest benefit from PE escape below  $5\ \mu\text{m}$ , as does the simulation of Dickerson and Garman [19]. It would be worthwhile to investigate larger crystals, up to around  $20\ \mu\text{m}$ , as they may also benefit from PE escape at higher x-ray energies [19], however the predicted benefit is greater for smaller crystals and these should be the primary focus.

In addition to measuring spot fading with different protein samples and crystal sizes, additional x-ray energies should be investigated. The simulations of Dickerson and Garman [19]—which accounted for QE of the detector used—recommended an ideal x-ray energy of approximately 26 keV. Measuring spot fading for a given crystal sample at this energy and above (up to approximately 30 keV) would provide experimental validation of the simulations, and help to establish recommended protocols for future high energy MX experiments. Once a robust body of spot fading data was available, future investigations could also focus on the feasibility of conducting, for example, phasing measurements at higher energies, which would have implications for sample preparation.

Ideally, future experiments in high energy MX will be conducted at a range of beamlines world-wide. As our PETRA-III experiment demonstrated, obtaining an appropriate small (1  $\mu\text{m}$  to 3  $\mu\text{m}$ ) focal spot during beamline testing does not necessarily translate to routine availability of such a beam focus. Obtaining a small beam spot at high energy is an additional challenge. Our experiment at NSLS-II pushed the limits of the beamline's capacity. Further experiments, at other facilities, should push the limits of other beamlines, establishing protocols to routinely achieve the beamline configurations necessary to take advantage of PE escape via high energy MX.

### 6.2.2 Reducing the Photoionisation Cross Section of the Solvent

The need to minimise the illumination of solvent has been noted since the very earliest simulations of PE motion [11], as PEs produced in the solvent can be scattered into the crystal, increasing, rather than reducing, the rate of damage. As we noted in section 4.4.7, even if the volume of illuminated solvent *around* the crystal is minimised by either draining the solvent or tightly focusing the beam on the crystal, it is impossible to minimise the illumination of the solvent *within* the crystal solvent channels, which will typically account for 30-60% of the crystal volume. Given that this volume will always be illuminated along with the crystal, the question is whether it is possible to minimise the amount of PEs produced here by reducing the cross section for photoionisation.

The primary driver of the relatively high photoionisation cross section of a protein solvent was, in our case, the salt content. Our solvent had a 1 M concentration of NaCl. We conducted a brief trial to reduce the salt content of the storage solvent after the crystals had been produced but this invariably resulted in the crystals dissolving within a few minutes. Given the collection of variables involved in getting crystals to form in the first place and retain their form long-term, it is likely that any reduction in salt content would have to be balanced by some addition to retain the concentration and distribution of ions in the solvent, at the very least. If this can be accomplished while reducing the cross section of the solvent, it could improve the chances of reducing damage in micro crystallography experiments via PE escape. Finding suitable solvent conditions would likely involve large-scale laboratory trials and could form the focus of future research.

### 6.2.3 Adapting Current Models for Room Temperature Crystallography

Over the last two decades, the vast majority of MX experiments have been performed at cryogenic temperatures, where damage rates are dependent on dose but independent of *dose rate*. Recently, there has been a renewed interest in conducting MX experiments at room temperature

(RT), where the dose rate is important [76, 92, 148, 150]. Regardless of the dose rate, however, total tolerable doses at RT are, at best, around two orders of magnitude lower than at cryogenic temperature. The introduction of serial crystallography techniques, first at XFELs and now at synchrotrons, mean that the total dose required for structure solution can be spread across many crystals, making RT crystallography viable with micro-crystals. There are several reasons why RT crystallography may be preferred, chiefly that results may be more biologically meaningful due to the fact that conformational heterogeneity and functional flexibility can be preserved at RT, while they are frozen out at cryogenic temperatures [81]. Preservation of functional flexibility allows the study of time-resolved protein dynamics. Fluid delivery systems such as the Lipidico [58, 59] injector used at the Australian Synchrotron only operate at room temperature.

As RT crystallography appears to be in resurgence and given the increased radiation-sensitivity of protein crystals at RT, it will be useful to model dose and dose-rate dependence under these conditions. Current predictive models of radiation damage are based on the dose-rate independence of cryogenic crystallography and cannot be used 'as is' to predict damage rates for room temperature experiments. Recently, Hadian-Jazi *et al.* [151] observed multiple hits on micro-crystals during an RT crystallography experiment using a high-viscosity injector. A summary of their results shows exponential fading of Bragg spots with dose. A handful of spot fading experiments have been conducted comparing the impact of temperature on half dose [76, 82, 150]. A coarse meta-analysis is possible from this existing data, and an average half dose of 0.57 MGy has been suggested by De la Mora *et al.* [82], but a more thorough investigation, comparing several different protein targets and dose-rates could be used to construct an empirical model that could enable a prediction of damage rates at RT, as a function of both dose and dose rate. This analysis could be based on the modelling approaches of Holton and Atakisi that were used and adapted in Chapter 5.

## Appendix A

# Example RADDPOSE-3D Input File

We used RADDPOSE-3D as part of our simulation, presented in chapter 3, and to estimate doses in the experiments described in chapter 4. The following document is an example of a RADDPOSE-3D input text file, specifying the parameters used in the dose calculations. The values in the example are from one of the iterations of our simulation execution.

```
#####  
#                               Crystal Block                               #  
#####
```

```
# Crystal: 1um x 1um x 1um  
# Cryoprotectant: Glycerol/PEG400, 9% NaCl
```

Crystal

```
Type Cuboid # Cuboid  
Dimensions 1.0 1.0 1.0 # Dimensions of the crystal in X,Y,Z in µm.  
PixelsPerMicron 50 # The computational resolution  
AbsCoefCalc RD3D # Tells RADDPOSE-3D how to calculate the  
# Absorption coefficients
```

```
UnitCell 79.24 79.24 38.01 # unit cell size: a, b, c  
# alpha, beta and gamma angles default to 90°  
NumMonomers 8 # number of monomers in unit cell  
NumResidues 129 # number of residues per monomer  
ProteinHeavyAtoms S 10 # heavy atoms added to protein part of the  
# monomer, i.e. S, coordinated metals,  
# Se in Se-Met  
SolventHeavyConc Na 772 Cl 772 # concentration of elements in the solvent  
# in mmol/l. Oxygen and lighter elements  
# should not be specified  
SolventFraction 0.36 # fraction of the unit cell occupied by solvent
```

```
#####  
#                               Beam Block                               #  
#####
```

Beam

```
Type Gaussian # Gaussian profile beam  
Flux 1e10 # in photons per second (2e12 = 2 * 10^12)  
FWHM 1.0 1.0 # in µm, vertical by horizontal for a  
Gaussian beam  
Energy 9.0 # in keV
```

```
Collimation Rectangular 4.4 4.4 # Vertical/Horizontal collimation of the beam  
# For 'uncollimated' Gaussians, 3xFWHM  
# recommended
```

```
#####  
#                               Wedge Block                               #  
#####
```

```
Wedge 0.0 0.0 # Start and End rotational angle of the crystal  
# Start < End  
ExposureTime 1 # Total time for entire angular range in seconds  
# AngularResolution 0.01 # Only change from the defaults when using very  
# small wedges, e.g 5°.
```

## Appendix B

# Spot Fading Meta-Analysis: Additional Plots

As part of the research presented in chapter 5, we conducted a meta-analysis of resolution-dependent spot fading results from the recent literature. Figure 5.7 contains an indicative sample of results from Bourenkov and Popov [141], Warkentin and Thorne [142], and Liebschner *et al.* [75], where there was some overlap in the resolution bands observed by the three studies. These results are compared to the two spot fading models, Model A and Model H, that we investigated in that chapter. We include here the full results from Bourenkov and Popov (figs. B.1, B.2, B.3), Warkentin and Thorne (fig. B.4), and Liebschner *et al.* (fig. B.5).

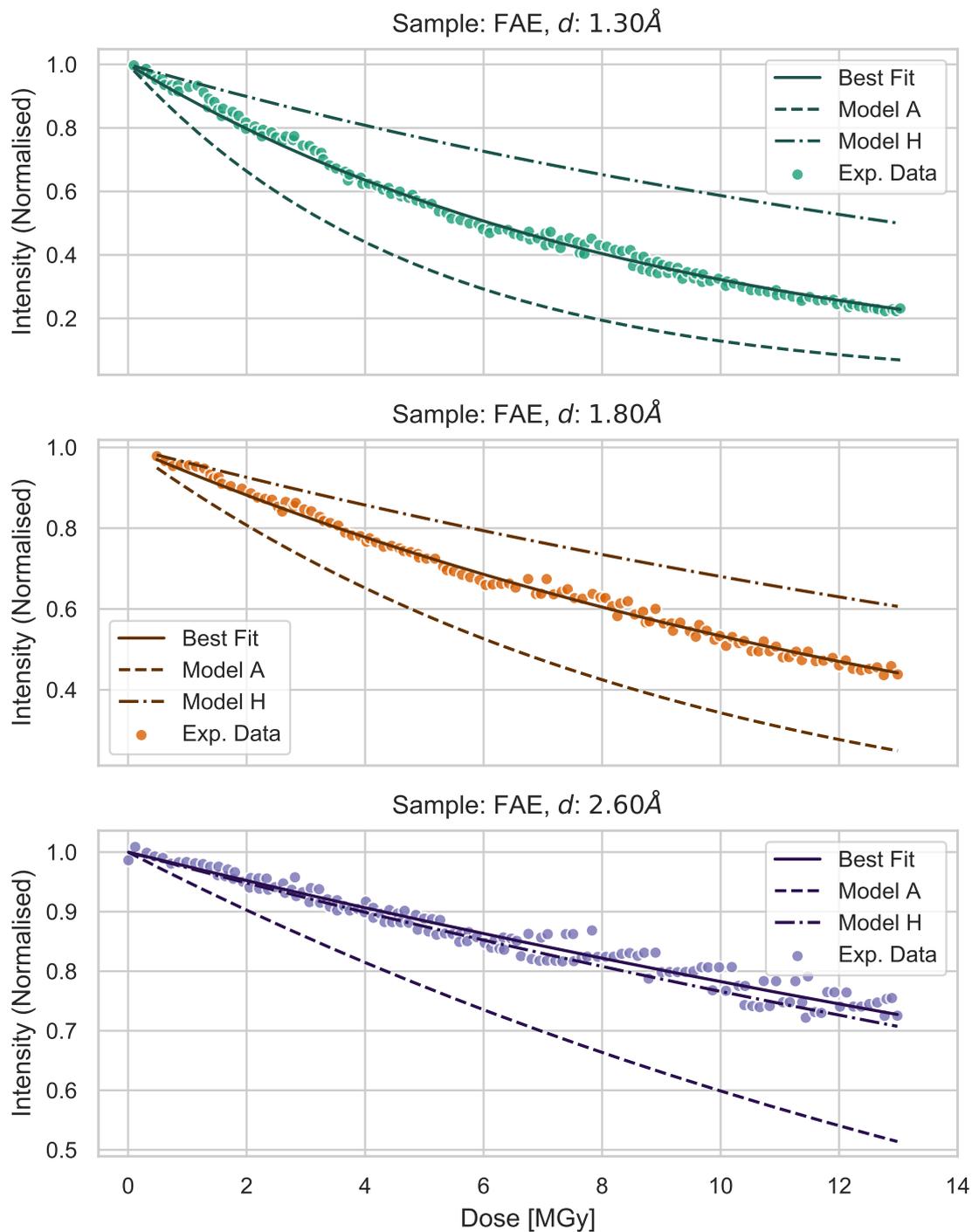


Figure B.1: Shell intensity data for FAE crystals of Bourenkov and Popov, compared to Model A and Model H. At higher resolutions, the best fit for the data is between Model A and Model H; for  $d = 2.6\text{\AA}$  the best fit is close to Model H. The total dose is low (approx. 13 MGy), sufficient to reduce high resolution peaks to half intensity but not far beyond.

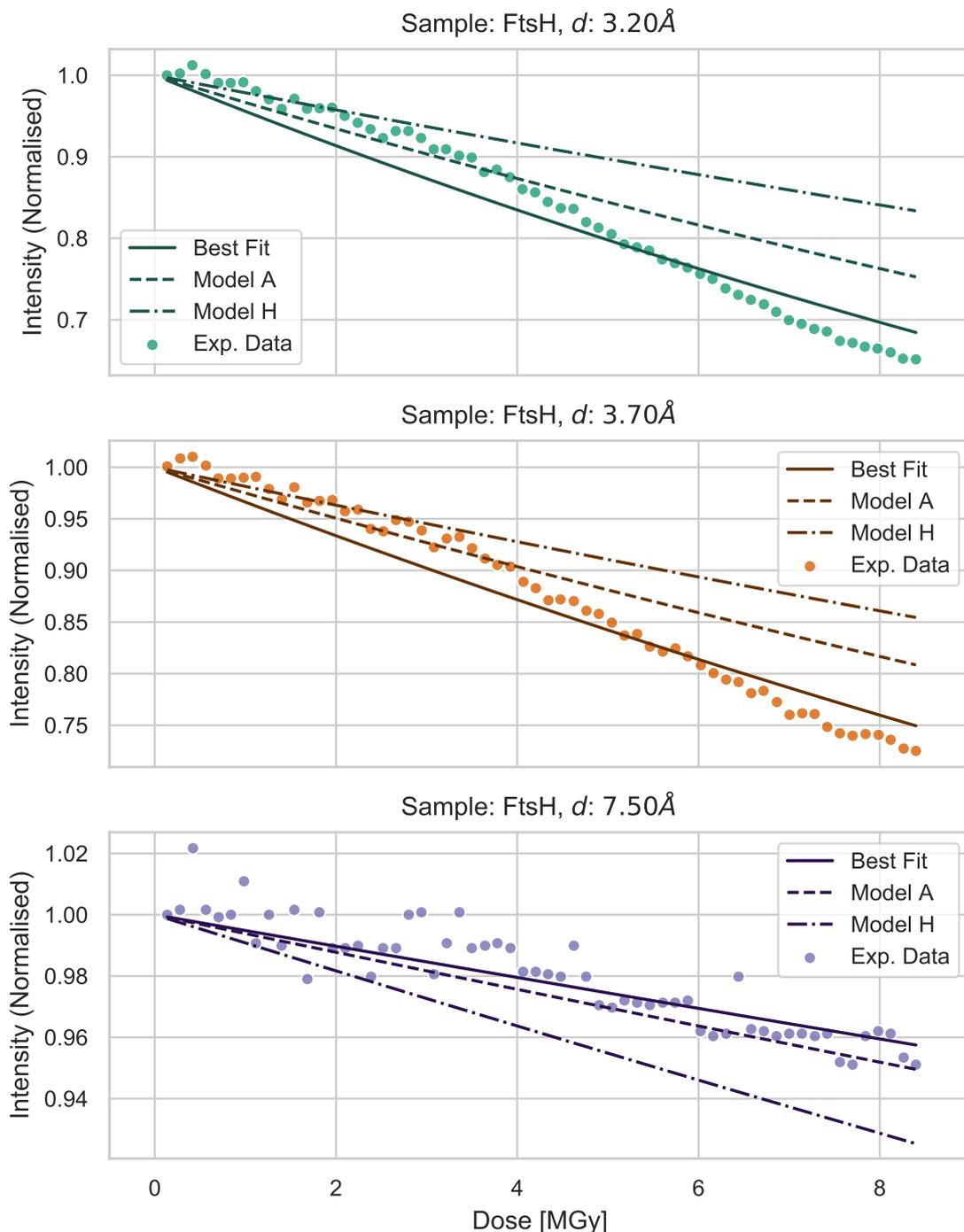


Figure B.2: **Shell intensity data for FtsH crystals of Bourenkov and Popov, compared to Model A and Model H.** The total dose (approx. 9 MGy) is very low and none of the resolution shells measured reach half initial intensity. There is a distinct pattern in the data at 3.2 Å and 3.7 Å that may match the initial plateaus observed by Owen *et al.* [149] and Warkentin *et al.* [76], however without a longer exposure this cannot be known; the data may also exhibit an oscillating component due to some systematic error in collection. In so far as the data fit an exponential curve, that curve does not match either Model A or Model H; the higher resolution shells fade more quickly than both models, the lower resolution shells more slowly.

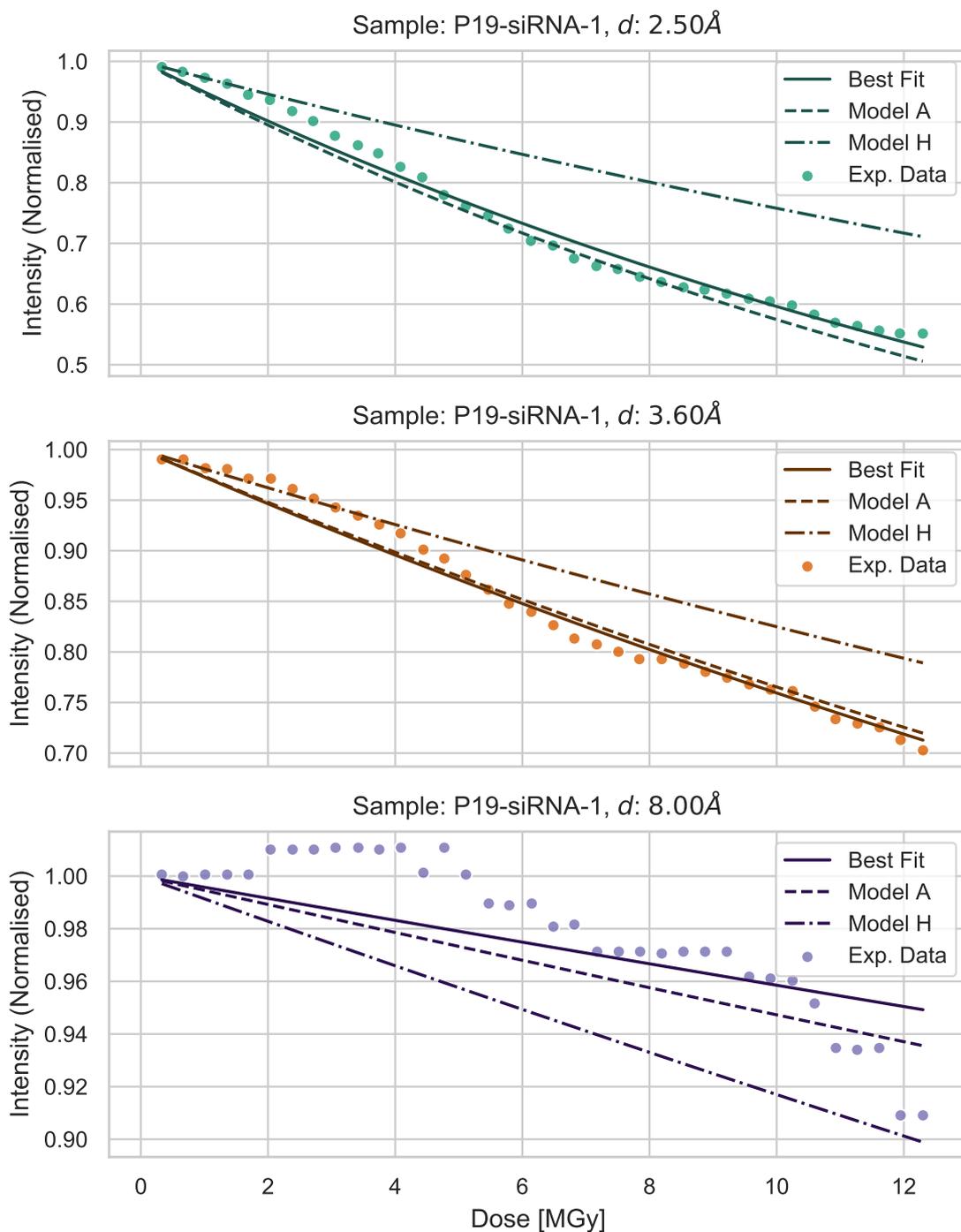


Figure B.3: Shell intensity data for P19-siRNA-1 crystals of Bourenkov and Popov, compared to Model A and Model H. The higher resolution shells exhibit fading at a rate that is consistent with Model A, though with an oscillating component that may be due to systematic error in data collection. The lower resolution shell is noisy and does not match either model well; as the total dose is well below the expected shell half-dose, it is not possible to make a meaningful comparison to the models.

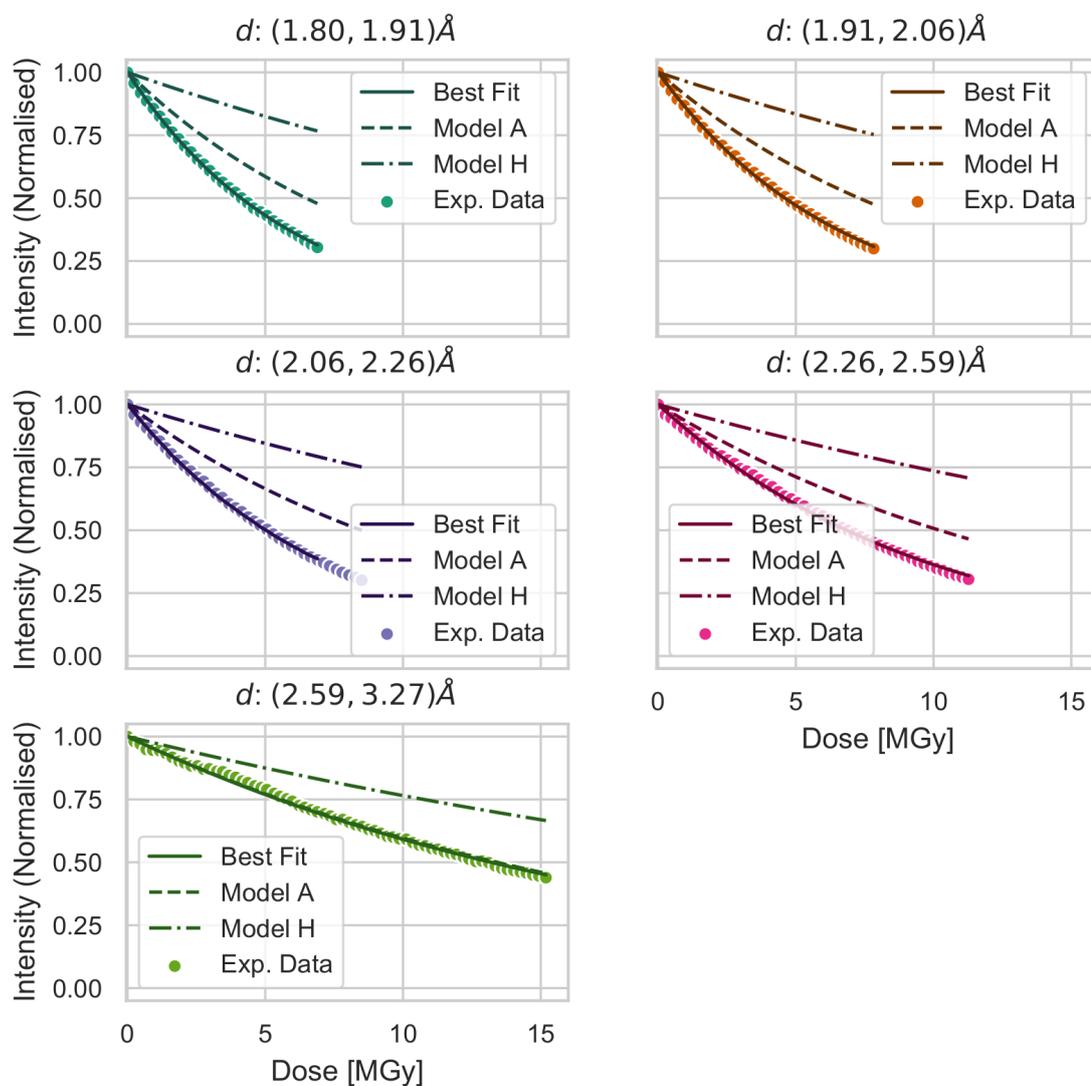


Figure B.4: **Shell intensity data for thaumatin crystals of Warkentin and Thorne, compared to Model A and Model H.** For all but the lowest resolution shell (2.59 Å to 3.27 Å), the measured intensities fade faster than the predictions of both Model A and Model H; for the lowest resolution shell the best fit matches Model A.

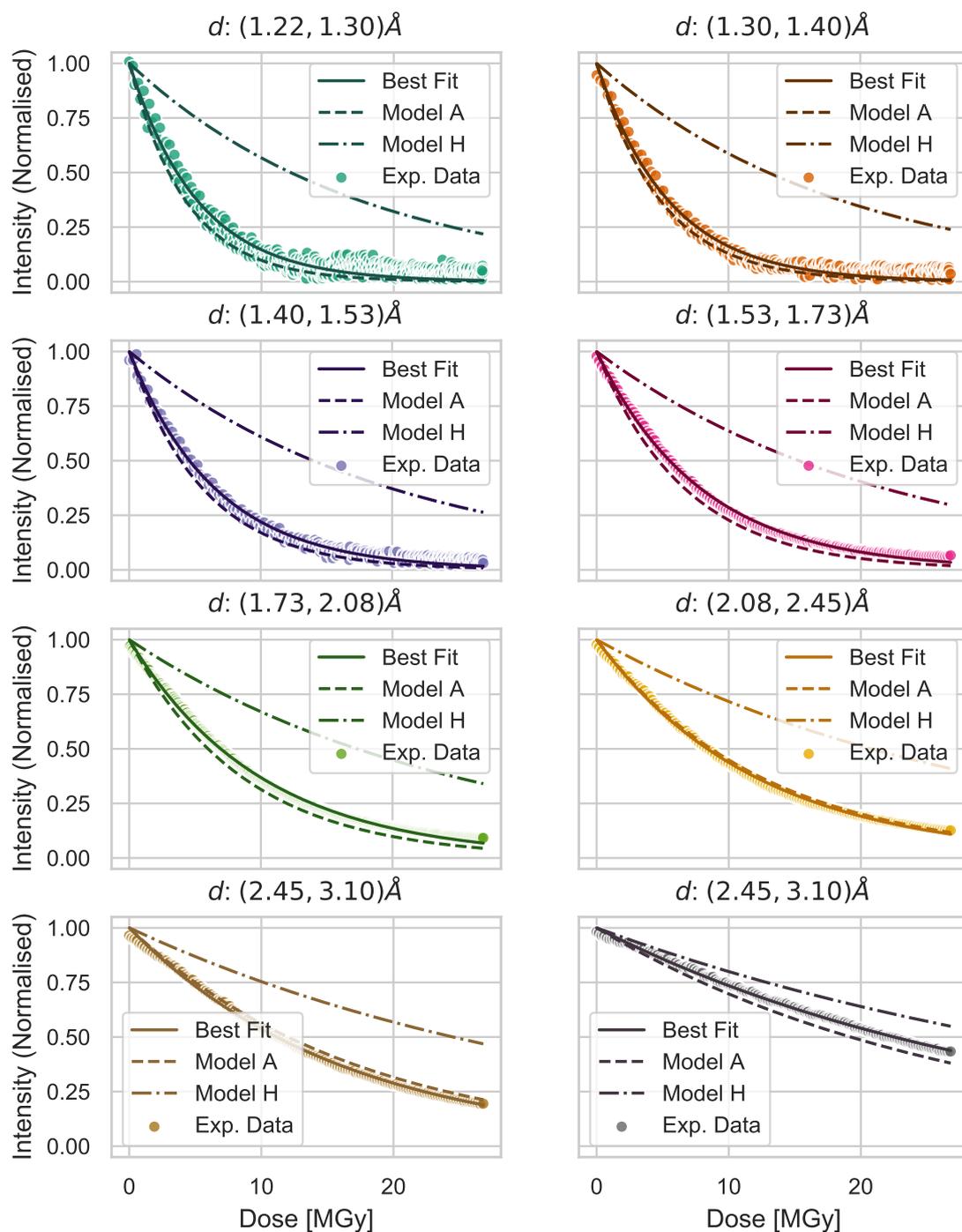


Figure B.5: Shell intensity data for thaumatin crystals of Liebschner *et al.*, compared to Model A and Model H. At all resolutions, the best fit matches Model A very closely. Model H provides a reasonable description for the data at the lowest resolution only, where the difference between the two models is reduced.

# Appendix C

## Image Permissions

Some images used in this thesis are the work of others, or adapted from the work of others. These have first appeared in the scientific literature or online. Where such images have been used, it has been by permission of the creator or publisher, or under the auspices of a creative commons license, and their use has been acknowledged in the relevant figure captions. A complete list of images used is collected here.

### C.1 Images Used Under the Auspice of a Creative Commons License

Figure 2.3c contains an image created by Rob Lavinsky. It is used, without modification, under the Creative Commons CC-BY-SA 3.0 license (<https://creativecommons.org/licenses/by-sa/3.0/>). The image is hosted by Wikipedia (<https://en.wikipedia.org/wiki/Halite#/media/File:Halite-249324.jpg>).

Table 2.1 contains images created DrBob, Daniel Mayer, Rocha, and Stannered. They are used under the Creative Commons CC-BY-SA 3.0 license (<https://creativecommons.org/licenses/by-sa/3.0/>). The images have been modified for inclusion in this thesis by digital tracing and re-rendering, which has removed some content from the original images. The original images can be found on Wikipedia ([https://en.wikipedia.org/wiki/Bravais\\_lattice](https://en.wikipedia.org/wiki/Bravais_lattice)).

Figure 2.9a contains an image created by Thomas Shafee. It is used under the Creative Commons CC-BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/>). The image has been modified by cropping the original. The original image is hosted by Wikipedia ([https://en.wikipedia.org/wiki/Enzyme#/media/File:Enzyme\\_structure.svg](https://en.wikipedia.org/wiki/Enzyme#/media/File:Enzyme_structure.svg)).

Figure 2.12 contains an image published in the Archives of Biochemistry and Biophysics [45]. Reuse of the contents of this article is permitted under the Creative Commons CC-BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/>).

## C.2 Images Used by Permission of the Owner

Figure 2.5 is composed of 10 individual images created by Kevin Cowtan from the University of York. The images are hosted on the web in 'Kevin Cowtan's Picture Book of Fourier Transforms': <http://www.yesbl.york.ac.uk/~cowtan/fourier/fourier.html>. They are used by permission of the creator, which was granted by email on 3<sup>rd</sup> June, 2021. The contrast of the images has been increased for improved visibility.

Figure 2.11 contains two diffraction images published in the Journal of Synchrotron Radiation [44]. This journal is published by the International Union of Crystallography (IUCr), who have granted blanket permission to reuse figures in scientific publications, including theses. The permission terms can be found online: <https://journals.iucr.org/services/permissions.html>.

Figure 2.13 contains an image published in Nature [9]. The image is reused by permission from the publisher, Springer Nature, granted via the RightsLink (<https://www.copyright.com/publishers/rightslink/>) service on 9<sup>th</sup> December, 2021.

Figure 2.14 contains two images published in Synchrotron Radiation News [74]. The images are reused by permission from the publisher, Taylor & Francis, granted via the RightsLink (<https://www.copyright.com/publishers/rightslink/>) service on 10<sup>th</sup> December, 2021.

# Bibliography

- [1] C. Branden and J. Tooze. *Introduction to Protein Structure*. 2nd. Garland, 1999.
- [2] B. Rupp. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. Garland Science, 2010.
- [3] A. Warshel et al. "Modeling Electrostatic Effects in Proteins (Review)". In: *Biochimica et Biophysica Acta* 1764 (2006), pp. 1647–1676. DOI: <https://doi.org/10.1016/j.bbapap.2006.08.007>.
- [4] H. M. Berman et al. "The Protein Data Bank". In: *Nucleic Acids Research* 28 (), pp. 235–242. DOI: [10.1093/nar/28.1.235](https://doi.org/10.1093/nar/28.1.235).
- [5] RCSB. *PDB Statistics*. 2000. URL: <https://www.rcsb.org/stats/summary>.
- [6] The UniProt Consortium. "UniProt: the Universal Protein Knowledgebase in 2021". In: *Nucleic Acids Research* 49.D1 (2020), pp. D480–D489. DOI: [10.1093/nar/gkaa1100](https://doi.org/10.1093/nar/gkaa1100).
- [7] S. D. Durbin and G. Feher. "Protein Crystallization". In: *Annual Review of Physical Chemistry* 47.1 (1996), pp. 171–204. DOI: [10.1146/annurev.physchem.47.1.171](https://doi.org/10.1146/annurev.physchem.47.1.171).
- [8] N. E. Chayen. "Turning Protein Crystallisation From an Art into a Science". In: *Current Opinion in Structural Biology* 14.5 (2004), pp. 577–583. DOI: <https://doi.org/10.1016/j.sbi.2004.08.002>.
- [9] H. N. Chapman et al. "Femtosecond X-ray Protein Nanocrystallography". In: *Nature* 470 (2011), pp. 73–77. DOI: [10.1038/nature09750](https://doi.org/10.1038/nature09750).
- [10] F. Stellato et al. "Room-Temperature Macromolecular Serial Crystallography Using Synchrotron Radiation". In: *IUCrJ* 1 (2014), pp. 204–212. DOI: [10.1107/S2052252514010070](https://doi.org/10.1107/S2052252514010070).
- [11] C. Nave and M. A. Hill. "Will Reduced Radiation Damage Occur With Very Small Crystals?" In: *Journal of Synchrotron Radiation* 12.3 (2005), pp. 299–303. DOI: [10.1107/S0909049505003274](https://doi.org/10.1107/S0909049505003274).
- [12] J. A. Cowan and C. Nave. "The Optimum Conditions to Collect X-ray Data From Very Small Samples". In: *Journal of Synchrotron Radiation* 15.5 (2008), pp. 458–462. DOI: [10.1107/S0909049508014623](https://doi.org/10.1107/S0909049508014623).

- [13] E. F. Garman. "Radiation Damage in Macromolecular Crystallography: What is it and Why Should We Care?" In: *Acta Crystallographica Section D* 66.4 (2010), pp. 339–351. DOI: 10.1107/S0907444910008656.
- [14] E. F. Garman. In: *Protein Crystallization 2nd Ed.* Ed. by T. M. Bergfors. La Jolla: International University Line, 2009, pp. 181–196.
- [15] E. M. H. Duke and L. N. Johnson. "Macromolecular Crystallography at Synchrotron Radiation Sources: Current Status and Future Developments". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466.2124 (2010), pp. 3421–3452. DOI: 10.1098/rspa.2010.0448.
- [16] R. Henderson. "Cryo-Protection of Protein Crystals against Radiation Damage in Electron and X-Ray Diffraction". In: *Proceedings of the Royal Society of London B: Biological Sciences* 241.1300 (1990), pp. 6–8. DOI: 10.1098/rspb.1990.0057.
- [17] T. Ursby et al. "BioMAX – The First Macromolecular Crystallography Beamline at MAX IV Laboratory". In: *Journal of Synchrotron Radiation* 27.5 (2020), pp. 1415–1429. DOI: 10.1107/S1600577520008723.
- [18] J. Christensen et al. "Radiation Damage in Small-Molecule Crystallography: Fact Not Fiction". In: *IUCrJ* 6.4 (2019), pp. 703–713. DOI: 10.1107/S2052252519006948.
- [19] J. L. Dickerson and E. F. Garman. "The Potential Benefits of Using Higher X-ray Energies for Macromolecular Crystallography". In: *Journal of Synchrotron Radiation* 26 (4 2019), pp. 922–930. DOI: 10.1107/S160057751900612X.
- [20] Q. Liu and W. A. Hendrickson. "Contemporary Use of Anomalous Diffraction in Biomolecular Structure Analysis". In: *Methods in molecular biology* 1607 (2017), pp. 377–399. DOI: 10.1007/978-1-4939-7000-1\_16.
- [21] H. Marman, C. Darmanin, and B. Abbey. "The Influence of Photoelectron Escape in Radiation Damage Simulations of Protein Micro-Crystallography". In: *Crystals* 8.7 (2018). DOI: 10.3390/cryst8070267.
- [22] J.J. Hubbell and S.M. Seltzer. *X-Ray Mass Attenuation Coefficients*. National Institute of Standards and Technology. 2004. DOI: <https://dx.doi.org/10.18434/T4D01F>. URL: <https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients>.
- [23] A. H. Compton. "A Quantum Theory of the Scattering of X-rays by Light Elements". In: *Phys. Rev.* 21 (5 1923), pp. 483–502. DOI: 10.1103/PhysRev.21.483.
- [24] A. M. Glazer. *Crystallography: A Very Short Introduction*. Oxford University Press, 2016.
- [25] A. Bravais. "Mémoire sur les systèmes formés par les points distribués régulièrement sur un plan ou dans l'espace". In: *J. École Polytech* 19 (1850), pp. 1–128.

- [26] H. D. Coughlan et al. "Three-Dimensional Reconstruction of the Size and Shape of Protein Microcrystals Using Bragg Coherent Diffractive Imaging". In: *Journal of Optics* 18 (2016), p. 054003. DOI: 10.1088/2040-8978/18/5/054003.
- [27] H. D. Coughlan et al. "Radiation Damage in a Micron-Sized Protein Crystal Studied via Reciprocal Space Mapping and Coherent Diffractive Imaging". In: *Structural Dynamics* 2 (2015), p. 041704. DOI: 10.1063/1.4919641.
- [28] H. D. Coughlan et al. "Bragg Coherent Diffraction Imaging and Metrics for Radiation Damage in Protein Micro-Crystallography". In: *Journal of Synchrotron Radiation* 24 (2017), pp. 83–94. DOI: 10.1107/S1600577516017525.
- [29] A. McPherson. "Introduction to Protein Crystallization." In: *Methods* 34 (3 2004), pp. 254–265. DOI: 10.1016/j.ymeth.2004.03.019.
- [30] P. Laschtschenko. "Über die keimtötende und entwicklungshemmende Wirkung von Hühnereiweiß". In: *Zeitschrift für Hygiene und Infektionskrankheiten* 64 (1 1909), pp. 419–427. DOI: 10.1007/BF02216170.
- [31] A. Fleming. "On a remarkable Bacteriolytic Element Found in Tissues and Secretions". In: *Proc. R. Soc. Lond. B.* 93 (653 1922), pp. 306–317. DOI: 10.1098/rspb.1922.0023.
- [32] C. C. F. Blake et al. "Structure of Hen Egg-White Lysozyme: A Three-dimensional Fourier Synthesis at 2 Å Resolution". In: *Nature* 206 (4986 1965), pp. 757–761. DOI: 10.1038/206757a0.
- [33] L. N. Johnson and D. C. Phillips. "Structure of Some Crystalline Lysozyme-Inhibitor Complexes Determined by X-Ray Analysis At 6 Å Resolution". In: *Nature* 206 (4986 1965), pp. 761–763. DOI: 10.1038/206761a0.
- [34] C. F. Macrae et al. "Mercury CSD 2.0– New Features for the Visualization and Investigation of Crystal Structures". In: *Journal of Applied Crystallography* 41.2 (2008), pp. 466–470. DOI: 10.1107/S0021889807067908.
- [35] M. Eriksson et al. "The MAX IV Synchrotron Light Source". In: *Proceedings of IPAC2011*. 2011, pp. 3026–3028.
- [36] MAX-IV Laboratory. *First light at MAXIV!* 2015. URL: [https://twitter.com/maxiv\\_laboratory/status/661454023585865728?lang=en](https://twitter.com/maxiv_laboratory/status/661454023585865728?lang=en) (visited on 03/03/2021).
- [37] K.-J. Kim. "Characteristics of Synchrotron Radiation". In: *AIP Conference Proceedings* 184.1 (1989), pp. 565–632. DOI: 10.1063/1.38046.
- [38] H. Onuki and P. Elleaume. *Undulators, Wigglers and Their Applications*. Taylor Francis, 2002. DOI: 10.1201/9780203218235.
- [39] P. F. Tavares et al. "The MAXIV Storage Ring Project". In: *Journal of Synchrotron Radiation* 21.5 (2014), pp. 862–877. DOI: 10.1107/S1600577514011503.

- [40] L. Liu et al. "The Sirius Project". In: *Journal of Synchrotron Radiation* 21.5 (2014), pp. 904–911. DOI: 10.1107/S1600577514011928.
- [41] G. Rosenbaum, K. C. Holmes, and J. Witz. "Synchrotron Radiation as a Source for X-ray Diffraction". In: *Nature* 230 (5294 1971), pp. 434–437. DOI: 10.1038/230434a0.
- [42] A. Harmsen, R. Leberman, and G.E. Schulz. "Comparison of Protein Crystal Diffraction Patterns and Absolute Intensities From Synchrotron and Conventional X-ray Sources". In: *Journal of Molecular Biology* 104.1 (1976), pp. 311–314. DOI: 10.1016/0022-2836(76)90017-6.
- [43] J. C. Phillips et al. "Applications of Synchrotron Radiation to Protein Crystallography: Preliminary Results". In: *Proceedings of the National Academy of Sciences* 73.1 (1976), pp. 128–132. DOI: 10.1073/pnas.73.1.128.
- [44] Z. Dauter, M. Jaskolski, and A. Wlodawer. "Impact of Synchrotron Radiation on Macromolecular Crystallography: A Personal View". In: *Journal of Synchrotron Radiation* 17.4 (2010), pp. 433–444. DOI: 10.1107/S0909049510011611.
- [45] R. L. Owen, J. Juanhuix, and M. Fuchs. "Current Advances in Synchrotron Radiation Instrumentation for MX Experiments". In: *Archives of Biochemistry and Biophysics* 602 (2016), pp. 21–31. DOI: 10.1016/j.abb.2016.03.021.
- [46] A. Kuller et al. "A Biologist's Guide to Synchrotron Facilities: The BioSync Web Resource." In: *Trends Biochem. Sci.* 27 (4 2002), pp. 213–215. DOI: 10.1016/S0968-0004(02)02056-x.
- [47] M. J. Gabanyi et al. "The Structural Biology Knowledgebase: A Portal to Protein Structures, Sequences, Functions, and Methods". In: *Journal of Structural and Functional Genomics* 12 (2 2011), pp. 45–54. DOI: 10.1007/s10969-011-9106-2.
- [48] S. Brockhauser, R. B. G. Ravelli, and A. A. McCarthy. "The Use of a Mini- $\kappa$  Goniometer Head in Macromolecular Crystallography Diffraction Experiments". In: *Acta Crystallographica Section D* 69.7 (2013), pp. 1241–1251. DOI: 10.1107/S0907444913003880.
- [49] S. Waltersperger et al. "PRIGo: a New Multi-Axis Goniometer for Macromolecular Crystallography". In: *Journal of Synchrotron Radiation* 22.4 (2015), pp. 895–900. DOI: 10.1107/S1600577515005354.
- [50] J. Aishima et al. "High-Speed Crystal Detection and Characterization Using a Fast-Readout Detector". In: *Acta Crystallographica Section D* 66.9 (2010), pp. 1032–1035. DOI: 10.1107/S0907444910028192.
- [51] V. Cherezov et al. "Rastering Strategy for Screening and Centring of Microcrystal Samples of Human Membrane Proteins with a Sub-10  $\mu\text{m}$  size X-ray Synchrotron Beam". In: *Journal of The Royal Society Interface* 6 (2009), S587–S597. DOI: 10.1098/rsif.2009.0142.focus.

- [52] J. Sanchez-Weatherby et al. "Improving Diffraction by Humidity Control: a Novel Device Compatible with X-ray Beamlines". In: *Acta Crystallographica Section D* 65.12 (2009), pp. 1237–1246. DOI: 10.1107/S0907444909037822.
- [53] T. Teng. "Mounting of Crystals for Macromolecular Crystallography in a Free-Standing Thin Film". In: *Journal of Applied Crystallography* 23.5 (1990), pp. 387–391. DOI: 10.1107/S0021889890005568.
- [54] P. Roedig et al. "Room-Temperature Macromolecular Crystallography Using a Micro-Patterned Silicon Chip with Minimal Background Scattering". In: *Journal of Applied Crystallography* 49.3 (2016), pp. 968–975. DOI: 10.1107/S1600576716006348.
- [55] D. P. DePonte et al. "Gas Dynamic Virtual Nozzle for Generation of Microscopic Droplet Streams". In: *Journal of Physics D: Applied Physics* 41.19 (2008), p. 195505. DOI: 10.1088/0022-3727/41/19/195505.
- [56] S. Botha et al. "Room-Temperature Serial Crystallography at Synchrotron X-ray Sources Using Slowly Flowing Free-Standing High-Viscosity Microstreams". In: *Acta Crystallographica Section D* 71.2 (2015), pp. 387–397. DOI: 10.1107/S1399004714026327.
- [57] P. Nogly et al. "Lipidic Cubic Phase Serial Millisecond Crystallography Using Synchrotron Radiation". In: *IUCrJ* 2.2 (2015), pp. 168–176. DOI: 10.1107/S2052252514026487.
- [58] P. Berntsen et al. "The Serial Millisecond Crystallography Instrument at the Australian Synchrotron Incorporating the "Lipidico" Injector". In: *Review of Scientific Instruments* 90.8 (2019), p. 085110. DOI: 10.1063/1.5104298.
- [59] P. Berntsen et al. "Lipidico Injection Protocol for Serial Crystallography Measurements at the Australian Synchrotron". In: *JoVE* 163 (2020). DOI: doi:10.3791/61650.
- [60] S. Sui et al. "A Capillary-Based Microfluidic Device Enables Primary High-Throughput Room-Temperature Crystallographic Screening". In: *Journal of Applied Crystallography* 54.4 (2021), pp. 1034–1046. DOI: 10.1107/S1600576721004155.
- [61] C. Blasetti and S. Deiuri. *Way For Light: The Catalogue of European Lightsources*. 2015. URL: <https://www.wayforlight.eu/en/>.
- [62] Y. S. Derbenev, A. M. Kondratenko, and E. L. Saldin. "On the Possibility of Using a Free Electron Laser for Polarization of Electrons in Storage Rings". In: *Nuclear Instruments and Methods in Physics Research* 193.3 (1982), pp. 415–421. DOI: 10.1016/0029-554X(82)90233-6.
- [63] J. B. Murphy and C. Pellegrini. "Free Electron Lasers for the XUV Spectral Region". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 237.1 (1985), pp. 159–167. DOI: 10.1016/0168-9002(85)90344-4.

- [64] P. Emma et al. "First Lasing and Operation of an Angstrom-Wavelength Free-Electron Laser". In: *Nature Photonics* 4 (9 2010), pp. 641–647. DOI: 10.1038/nphoton.2010.176.
- [65] J. M. Holton and K. A. Frankel. "The Minimum Crystal Size Needed for a Complete Diffraction Data Set". In: *Acta Cryst. D* 66 (2010), pp. 393–408. DOI: 10.1107/S0907444910007262.
- [66] O. B. Zeldin, M. Gerstel, and E. F. Garman. "RADDPOSE-3D: Time- and Space-Resolved Modelling of Dose in Macromolecular Crystallography". In: *Journal of Applied Crystallography* 46 (2013), pp. 1225–1230. DOI: 10.1107/S0021889813011461.
- [67] C. S. Bury et al. "Estimate Your Dose: RADDPOSE-3D". In: *Protein Science* 27 (2018), pp. 217–228. DOI: <https://doi.org/10.1002/pro.3302>.
- [68] J. W. Murray, E. F. Garman, and R. B. G. Ravelli. "X-ray Absorption by Macromolecular Crystals: The Effects of Wavelength and Crystal Composition on Absorbed Dose". In: *Journal of Applied Crystallography* 37.4 (2004), pp. 513–522. DOI: 10.1107/S0021889804010660.
- [69] K. S. Paithankar, R. L. Owen, and E. F. Garman. "Absorbed Dose Calculations for Macromolecular Crystals: Improvements to RADDPOSE". In: *Journal of Synchrotron Radiation* 16 (2009), pp. 152–162. DOI: 10.1107/S0909049508040430.
- [70] K. S. Paithankar and E. F. Garman. "Know Your Dose: RADDPOSE". In: *Acta Crystallographica Section D* 66 (2010), pp. 381–388. DOI: 10.1107/S0907444910006724.
- [71] J. M. Holton. "A Beginner's Guide to Radiation Damage". In: *Journal of Synchrotron Radiation* 16.2 (2009), pp. 133–142. DOI: 10.1107/S0909049509004361.
- [72] W. P. Burmeister. "Structural Changes in a Cryo-Cooled Protein Crystal Owing to Radiation Damage". In: *Acta Crystallographica Section D* 56.3 (2000), pp. 328–341. DOI: 10.1107/S0907444999016261.
- [73] R. B. G. Ravelli and S. M. McSweeney. "The 'Fingerprint' That X-rays Can Leave on Structures". In: *Structure* 8 (2000), pp. 315–328. DOI: 10.1016/S0969-2126(00)00109-X.
- [74] E. F. Garman and M. Weik. "Radiation Damage in Macromolecular Crystallography". In: *Synchrotron Radiation News* 28 (2015), pp. 15–19. DOI: 10.1080/08940886.2015.1101322.
- [75] D. Lieschner et al. "Radiation Decay of Thaumatin Crystals at Three X-ray Energies". In: *Acta Crystallographica Section D* 71.4 (2015), pp. 772–778. DOI: 10.1107/S1399004715001030.
- [76] M. A. Warkentin et al. "Lifetimes and Spatio-Temporal Response of Protein Crystals in Intense X-ray Microbeams". In: *IUCrJ* 4 (2017), pp. 785–794. DOI: 10.1107/S2052252517013495.

- [77] C. Nave and E. F. Garman. "Towards an Understanding of Radiation Damage in Cryocooled Macromolecular Crystals". In: *Journal of Synchrotron Radiation* 12.3 (2005), pp. 257–260. DOI: 10.1107/S0909049505007132.
- [78] T.-Y. Teng and K. Moffat. "Primary Radiation Damage of Protein Crystals by an Intense Synchrotron X-ray Beam". In: *Journal of Synchrotron Radiation* 7.5 (2000), pp. 313–317. DOI: 10.1107/S0909049500008694.
- [79] T.-Y. Teng and K. Moffat. "Radiation Damage of Protein Crystals at Cryogenic Temperatures Between 40K and 150K". In: *Journal of Synchrotron Radiation* 9.4 (2002), pp. 198–201. DOI: 10.1107/S0909049502008579.
- [80] E. F. Garman. *Personal communication*. Sept. 2020.
- [81] J. S. Fraser et al. "Accessing Protein Conformational Ensembles Using Room-Temperature X-ray Crystallography". In: *Proceedings of the National Academy of Sciences* 108.39 (2011), pp. 16247–16252. DOI: 10.1073/pnas.1111325108.
- [82] E. De la Mora et al. "Radiation Damage and Dose Limits in Serial Synchrotron Crystallography at Cryo- and Room Temperatures". In: 117.8 (2020), pp. 4142–4151. DOI: 10.1073/pnas.1821522117.
- [83] C. C. F. Blake and D. C. Phillips. In: *Biological Effects of Ionizing Radiation at the Molecular Level*. International Atomic Energy Agency, Vienna. 1962, pp. 183–191.
- [84] W. A. Hendrickson. "Radiation Damage in Protein Crystallography". In: *Journal of Molecular Biology* 106.3 (1976), pp. 889–893. DOI: 10.1016/0022-2836(76)90271-0.
- [85] J. Sygusch and M. Allaire. "Sequential Radiation Damage in Protein Crystallography". In: *Acta Crystallographica Section A* 44.4 (1988), pp. 443–448. DOI: 10.1107/S0108767388001394.
- [86] H. Atakisi et al. "Resolution and Dose Dependence of Radiation Damage in Biomolecular Systems". In: *IUCrJ* 6.6 (2019), pp. 1040–1053. DOI: 10.1107/S2052252519008777.
- [87] M. R. Howells et al. "An Assessment of the Resolution Limitation Due to Radiation-Damage in X-ray Diffraction Microscopy". In: *Journal of Electron Spectroscopy and Related Phenomena* 170 (2009), pp. 4–12. DOI: 10.1016/j.elspec.2008.10.008.
- [88] J. Cosier and A. M. Glazer. "A Nitrogen-Gas-Stream Cryostat for General X-ray Diffraction Studies". In: *Journal of Applied Crystallography* 19.2 (1986), pp. 105–107. DOI: 10.1107/S0021889886089835.
- [89] E. G. Allan et al. "To Scavenge or not to Scavenge, that is STILL the Question". In: *Journal of Synchrotron Radiation* 20.1 (2013), pp. 23–36. DOI: 10.1107/S0909049512046237.
- [90] J. Murray and E. Garman. "Investigation of Possible Free-Radical Scavengers and Metrics for Radiation Damage in Protein Cryocrystallography". In: *Journal of Synchrotron Radiation* 9.6 (2002), pp. 347–354. DOI: 10.1107/S0909049502014632.

- [91] B. Kauffmann et al. "How to Avoid Premature Decay of Your Macromolecular Crystal: A Quick Soak for Long Life". In: *Structure* 14 (7 2006), pp. 1099–1105. DOI: 10.1016/j.str.2006.05.015.
- [92] R. J. Southworth-Davies and E. F. Garman. "Radioprotectant Screening for Cryocrystallography". In: *Journal of Synchrotron Radiation* 14.1 (2007), pp. 73–83. DOI: 10.1107/S0909049506044177.
- [93] E. Nowak et al. "To Scavenge or not to Scavenge: That is the Question". In: *Acta Crystallographica Section D* 65.9 (2009), pp. 1004–1006. DOI: 10.1107/S0907444909026821.
- [94] A. I. Barker et al. "Room-Temperature Scavengers for Macromolecular Crystallography: Increased Lifetimes and Modified Dose Dependence of the Intensity Decay". In: *Journal of Synchrotron Radiation* 16.2 (2009), pp. 205–216. DOI: 10.1107/S0909049509003343.
- [95] S. Macedo et al. "Can Soaked-in Scavengers Protect Metalloprotein Active Sites From Reduction During Data Collection?" In: *Journal of Synchrotron Radiation* 16.2 (2009), pp. 191–204. DOI: 10.1107/S0909049509003331.
- [96] E. De la Mora, I. Carmichael, and E. F. Garman. "Effective Scavenging at Cryotemperatures: Further Increasing the Dose Tolerance of Protein Crystals". In: *Journal of Synchrotron Radiation* 18.3 (2011), pp. 346–357. DOI: 10.1107/S0909049511007163.
- [97] J. Kmetko et al. "Can Radiation Damage to Protein Crystals be Reduced Using Small-Molecule Compounds?" In: *Acta Crystallographica Section D* 67.10 (2011), pp. 881–893. DOI: 10.1107/S0907444911032835.
- [98] R. Neutze et al. "Potential for Biomolecular Imaging with Femtosecond X-ray Pulses". In: *Nature* 406 (6797 2000), pp. 752–757. DOI: 10.1038/35021099.
- [99] A. Doerr. "Diffraction Before Destruction". In: *Nature Methods* 8 (4 2011). DOI: 10.1038/nmeth0411-283.
- [100] H. N. Chapman, C. Caleman, and N. Timneanu. "Diffraction Before Destruction". In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 369 (1647 2014), p. 2130313. DOI: 10.1098/rstb.2013.0313.
- [101] N. J. Carron. *An Introduction to the Passage of Energetic Particles through Matter*. Taylor & Francis, 2007.
- [102] S. Tanuma, C. J. Powell, and D. R. Penn. "Calculations of Electron Inelastic Mean Free Paths. IX. Data for 41 Elemental Solids Over the 50 eV to 30 keV Range". In: *Surf. Interface Anal.* 43 (2011), pp. 689–713. DOI: <https://doi.org/10.1002/sia.3522>.
- [103] H. Shinotsuka et al. "Calculations of Electron Inelastic Mean Free Paths. XI. Data for Liquid Water for Energies From 50 eV to 30 keV". In: *Surf. Interface Anal.* 49 (2017), pp. 238–252. DOI: <https://doi.org/10.1002/sia.6123>.

- [104] R. Sanishvili et al. "Radiation Damage in Protein Crystals is Reduced With a Micron-Sized X-ray Beam". In: *Proceedings of the National Academy of Sciences* 108.15 (2011), pp. 6127–6132. DOI: 10.1073/pnas.1017701108.
- [105] P. Hovington, D. Drouin, and R. Gauvin. "CASINO: A New Monte Carlo Code in C Language for Electron Beam Interaction –Part I: Description of the Program". In: *Scanning* 19 (1997), pp. 1–14. DOI: <https://doi.org/10.1002/sca.4950190101>.
- [106] D. Drouin, P. Hovington, and R. Gauvin. "CASINO: A New Monte Carlo Code in C Language for Electron Beam Interaction –Part II: Tabulated Values of the Mott Cross Section". In: *Scanning* 19 (1997), pp. 20–28. DOI: <https://doi.org/10.1002/sca.4950190103>.
- [107] D. Drouin et al. "CASINO V2.42 – A Fast and Easy-to-use Modeling Tool for Scanning Electron Microscopy and Microanalysis Users". In: *Scanning* 29 (2007), pp. 92–101. DOI: <https://doi.org/10.1002/sca.20000>.
- [108] E. A. Stern et al. "Reducing Radiation Damage in Macromolecular Crystals at Synchrotron Sources". In: *Acta Crystallographica Section D* 65.4 (2009), pp. 366–374. DOI: 10.1107/S090744490900540X.
- [109] Y. Z. Finfrook et al. "Spatial Dependence and Mitigation of Radiation Damage by a Line-Focus Mini-Beam". In: *Acta Crystallographica Section D* 66 (2010), pp. 1287–1294. DOI: 10.1107/S0907444910036875.
- [110] Y. Z. Finfrook et al. "Mitigation of X-ray Damage in Macromolecular Crystallography by Submicrometre Line Focusing". In: *Acta Crystallographica Section D* 69 (2013), pp. 1463–1469. DOI: 10.1107/S0907444913009335.
- [111] H. D. Coughlan. "Protein Crystals and Radiation Damage Studied Using Bragg Coherent Diffractive Imaging". PhD. La Trobe University, 2017.
- [112] J. L. Laclare. "Target Specifications and Performance of the ESRF Source". In: *Journal of Synchrotron Radiation* 1.1 (Oct. 1994), pp. 12–18. DOI: 10.1107/S0909049594006564.
- [113] E. T. White et al. "The Density of a Protein Crystal". In: *Powder Technology* 179.1 (2007), pp. 55–58. DOI: <https://doi.org/10.1016/j.powtec.2006.11.014>.
- [114] M.J. Berger et al. *XCOM: Photon Cross Section Database (version 1.5)*. National Institute of Standards and Technology. 2010. DOI: 10.2172/6016002. URL: <http://physics.nist.gov/xcom>.
- [115] M. Eriksson, J. F. van der Veen, and C. Quitmann. "Diffraction-Limited Storage Rings – A Window to the Science of Tomorrow". In: *Journal of Synchrotron Radiation* 21.5 (2014), pp. 837–842. DOI: 10.1107/S1600577514019286.
- [116] R. Hettel. "DLSR Design and Plans: An International Overview". In: *Journal of Synchrotron Radiation* 21.5 (2014), pp. 843–855. DOI: 10.1107/S1600577514011515.

- [117] P. Sliz, S. C. Harrison, and G. Rosenbaum. "How Does Radiation Damage in Protein Crystals Depend on X-Ray Dose?" In: *Structure* 11.1 (2003), pp. 13–19. DOI: 10.1016/S0969-2126(02)00910-3.
- [118] H.-K. S. Leiros et al. "Is Radiation Damage Dependent on the Dose Rate Used During Macromolecular Crystallography Data Collection?" In: *Acta Crystallographica Section D* 62.2 (2006), pp. 125–132. DOI: 10.1107/S0907444905033627.
- [119] R. L. Owen, E. Rudiño-Piñera, and E. F. Garman. "Experimental Determination of the Radiation Dose Limit for Cryocooled Protein Crystals". In: *Proceedings of the National Academy of Sciences* 103.13 (2006), pp. 4912–4917. DOI: 10.1073/pnas.0600973103.
- [120] Symposium on the Biological Effects of Ionizing Radiation at the Molecular Level (1962: Brno). *Biological Effects of Ionizing Radiation at the Molecular Level*. International Atomic Energy Agency, 1962.
- [121] R. J. Southworth-Davies et al. "Observation of Decreased Radiation Damage at Higher Dose Rates in Room Temperature Protein Crystallography". In: *Structure* 15 (12 2007), pp. 1531–1541. DOI: 10.1016/j.str.2007.10.013.
- [122] E. F. Garman and M. Weik. "X-ray Radiation Damage to Biological Macromolecules: Further Insights". In: *Journal of Synchrotron Radiation* 24 (2017), pp. 1–6. DOI: 10.1107/S160057751602018X.
- [123] T. Donath et al. "Characterization of the PILATUS Photon-Counting Pixel Detector for X-ray Energies From 1.75 keV to 60 keV". In: *Journal of Physics: Conference Series* 425.6 (2013), p. 062001. DOI: 10.1088/1742-6596/425/6/062001.
- [124] P. Zambon et al. "Spectral Response Characterization of CdTe Sensors of Different Pixel Size with the IBEX ASIC". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 892 (2018), pp. 106–113. DOI: 10.1016/j.nima.2018.03.006.
- [125] M. O. Wiedorn et al. "Megahertz Serial Crystallography". In: *Nature Communications* 9 (1 2018), p. 4025. DOI: 10.1038/s41467-018-06156-7.
- [126] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri. "NIH Image to ImageJ: 25 years of Image Analysis". In: *Nature Methods* 9 (7 2012), pp. 671–675. DOI: 10.1038/nmeth.2089.
- [127] C. T. Rueden et al. "ImageJ2: ImageJ for the Next Generation of Scientific Image Data". In: *BMC Bioinformatics* 18 (1 2017), p. 529. DOI: 10.1186/s12859-017-1934-z.
- [128] J. Schindelin et al. "Fiji: An Open-Source Platform for Biological-Image Analysis". In: *Nature Methods* 9 (7 2012), pp. 676–682. DOI: 10.1038/nmeth.2019.
- [129] Gatan Inc. *Digital Micrograph*. 2021. URL: <https://www.gatan.com/installation-instructions>.

- [130] A. Burkhardt et al. "Status of the Ccrystallography Beamlines at PETRA III". In: *The European Physical Journal Plus* 131 (3 2016), p. 56. DOI: 10.1140/epjp/i2016-16056-0.
- [131] Dectris AG. *ALBULA Software: Data Visualization and Image Manipulation*. 2021. URL: <https://www.dectris.com/detectors/albula-software/>.
- [132] T. A. White et al. "CrystFEL: A Software Suite for Snapshot Serial Crystallography". In: *Journal of Applied Crystallography* 45.2 (2012), pp. 335–341. DOI: 10.1107/S0021889812002312.
- [133] M. Fuchs et al. "NSLS-II Biomedical Beamlines for Macromolecular Crystallography, FMX and AMX, and for X-ray Scattering, LIX: Current Developments". In: *Journal of Physics: Conference Series* 493 (Feb. 2014). DOI: 10.1088/1742-6596/493/1/012021.
- [134] M. R. Fuchs et al. "NSLS-II Biomedical Beamlines for Micro-Crystallography, FMX, and for Highly Automated Crystallography, AMX: New Opportunities for Advanced Data Collection". In: *AIP Conference Proceedings* 1741.1 (2016), p. 030006. DOI: 10.1063/1.4952829.
- [135] F. Willeke. "Commissioning of NSLS-II". In: (2015). URL: <https://www.osti.gov/biblio/1193203>.
- [136] E. F. Garman. *Personal communication*. Aug. 2019.
- [137] The pandas development team. *Pandas*. Version 1.1.0. 2020. DOI: 10.5281/zenodo.3509134.
- [138] W. McKinney. "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by S. van der Walt and J. Millman. 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.
- [139] J. L. Dickerson and E. F. Garman. "Doses for Experiments with Microbeams and Microcrystals: Monte Carlo Simulations in RADDPOSE-3D". In: *Protein Science* 30.1 (2021), pp. 8–19. DOI: <https://doi.org/10.1002/pro.3922>.
- [140] S. L. S. Storm et al. "Measuring Energy-Dependent Photoelectron Escape in Microcrystals". In: *IUCrJ* 7 (2020), pp. 129–135. DOI: 10.1107/S2052252519016178.
- [141] G. P. Bourenkov and A. N. Popov. "Optimization of Data Collection Taking Radiation Damage into Account". In: *Acta Crystallographica Section D* 66 (2010), pp. 409–419. DOI: 10.1107/S0907444909054961.
- [142] M. A. Warkentin and R. E. Thorne. "Glass Transition in Thaumatin Crystals Revealed Through Temperature-Dependent Radiation-Sensitivity Measurements". In: *Act Crystallographica Section D* 66 (2010), pp. 1092–1100. DOI: doi:10.1107/S0907444910035523.

- [143] C. Nave et al. "Imperfection and Radiation Damage in Protein Crystals Studied with Coherent Radiation". In: *Journal of Synchrotron Radiation* 23 (2016), pp. 228–237. DOI: 10.1107/S1600577515019700.
- [144] A. Rohatgi. *Webplotdigitizer: Version 4.5*. 2021. URL: <https://automeris.io/WebPlotDigitizer>.
- [145] J. Kmetko et al. "Quantifying X-ray Radiation Damage in Protein Crystals at Cryogenic Temperatures". In: *Acta Crystallographica Section D* 62.9 (2006), pp. 1030–1038. DOI: 10.1107/S0907444906023869.
- [146] M. Weik et al. "Specific Chemical and Structural Damage to Proteins Produced by Synchrotron Radiation". In: 97.2 (2000), pp. 623–628. DOI: 10.1073/pnas.97.2.623.
- [147] M. Weik et al. "Solvent Behaviour in Flash-Cooled Protein Crystals at Cryogenic Temperatures". In: *Acta Crystallographica Section D* 57.4 (2001), pp. 566–573. DOI: 10.1107/S0907444901001196.
- [148] R. L. Owen et al. "Outrunning Free Radicals in Room-Temperature Macromolecular Crystallography". In: *Acta Crystallographica Section D* 68.7 (2012), pp. 810–818. DOI: 10.1107/S0907444912012553.
- [149] R. L. Owen et al. "Exploiting Fast Detectors to Enter a New Dimension in Room Temperature Crystallography". In: *Acta Crystallographica Section D* 70.5 (2014), pp. 1248–1256. DOI: 10.1107/S1399004714005379.
- [150] M. Warkentin et al. "Global Radiation Damage at 300 and 260K With Dose Rates Approaching 1 MGy s<sup>-1</sup>". In: *Acta Crystallographica Section D* 68.2 (2012), pp. 124–133. DOI: 10.1107/S0907444911052085.
- [151] M. Hadian-Jazi et al. "Analysis of Multi-Hit Crystals in Serial Synchrotron Crystallography Experiments Using High-Viscosity Injectors". In: *Crystals* 11.1 (2021). DOI: 10.3390/cryst11010049.
- [152] N. Asherie et al. "Effects of Protein Purity and Precipitant Stereochemistry on the Crystallization of Thaumatin". In: *Crystal Growth & Design* 8 (12 2008), pp. 4200–4207. DOI: 10.1021/cg800616q.
- [153] G. Guo et al. "Sample manipulation and data assembly for robust microcrystal synchrotron crystallography". In: *IUCrJ* 5.3 (2018), pp. 238–246. DOI: 10.1107/S2052252518005389.