Serial and time-resolved crystallography of metalloproteins

Peter Smyth

A thesis submitted for the degree of Doctor of Philosophy in Biochemistry

School of Life Sciences

University of Essex

June 2024

Abstract

X-ray crystallography is an important biological research technique that has provided most of the structures in the Protein Data Bank. An inherent challenge of crystallography is that the X-rays which produce diffraction images also cause damage to the crystal, resulting in loss of diffraction quality as the absorbed dose increases. To minimise this effect, serial crystallography techniques have been developed to spread dose over many crystals, by collecting a small amount of data, usually only one diffraction image, from each. This is especially beneficial for metalloproteins since metal centres are particularly susceptible to site-specific radiation damage and are important to the function of many proteins. The benefits of serial crystallography include the ability to perform experiments at room temperature, where proteins are more likely to be in physiologically relevant conformations, rather than at the cryogenic temperatures usually used to limit radiation damage. Room temperature also allows increased diffusion and molecular motions, so reactions can take place within protein crystals. These can be investigated with time-resolved serial experiments, capturing structures of short-lived intermediate states of a protein undergoing a dynamic process. An important part of any time-resolved experiment is reaction initiation, including ensuring the reaction begins in the whole sample at the same time.

This thesis describes the development of a method for time-resolved serial crystallography using the fixed-target chip system available on beamline I24 at Diamond Light Source, which is also applicable to serial femtosecond crystallography at X-ray free electron laser (XFEL) sources. These experiments explore the structures of highly radiation-sensitive metalloproteins, including time-resolved studies of nitric oxide binding in the haem site using a UV laser pump-probe technique and a photocaged substrate molecule.

Acknowledgements

Firstly, I would like to thank my supervisors Prof Mike Hough, Dr Robin Owen, and Prof Jonathan Worrall, for their guidance during my PhD. I have learned more about crystallography in the last four years than I thought possible when I started.

I would like to thank Dr Hiroshi Sugimoto and Dr Takehiko Tosha, our collaborators in Japan, and the staff at SACLA, for their help with beamtime, especially collecting data from our crystals when we were unable to travel.

I am grateful to Marina and Kes, for making lab 5.08 an enjoyable place to work, and for the many lunches together. Thanks also to Hannah for the assistance with protein expression and crystallisation, and for keeping the lab running.

I am thankful to Anna for being a great friend and colleague. Thank you for answering innumerable questions on how the lab works, and for keeping me sane while battling issues with protein expression and while writing this thesis.

Thanks goes to Danny, Sam, Sofia, Ali, and anyone else associated with I24, for the assistance you have provided me over the years, and more importantly, for keeping the biscuit tin topped up! I am thankful for the company and friendly discussions with everybody at lunch, especially when cake is involved. My thanks also goes to anyone else I've met at Diamond and the Research Complex at Harwell.

I would also like to thank my friends and family for supporting me in Colchester and Didcot and giving me something to talk about besides work.

Finally, this project would not have been possible without the funding for the studentship provided by the University of Essex and Diamond Light Source.

Table of contents

List of figures

List of tables

Abbreviations

5c	5-coordinate
6c	6-coordinate
A_{280}	absorbance (280 nm)
A_{400}	absorbance (400 nm)
AxCP	*Alcaligenes xylosoxidans* cytochrome *c*'
CV	column volume
Cyt	cytochrome
DtpB	DyP type B
DyP	dye-decolourising peroxidase
FWHM	full width at half maximum
GDVN	gas dynamic virtual nozzle
GUI	graphical user interface
HARE	hit-and-return
HEC	hydroxyethyl cellulose
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
LAMA	liquid application method for time-resolved analyses
LCP	lipidic cubic phase
McCP-β	*Methylococcus capsulatus* (Bath) cytochrome *c*'-β
MES	2-(*N*-morpholino)ethanesulfonic acid
MPD	2-methyl-2,4-pentanediol
MR	molecular replacement
MSOX	multiple serial structures from one crystal
MSS	multiple serial structures from many crystals
MX	macromolecular crystallography
ND	neutral density
OAV	on-axis viewer
OD_{600}	optical density (600 nm)
PDB	Protein Data Bank
PEG	polyethylene glycol
RMSD	root mean square deviation
ROS	reactive oxygen species
RT	room temperature
SACLA	Spring-8 Angstrom Compact Free Electron Laser
SD	standard deviation
SF-ROX	serial femtosecond rotation crystallography
SFX	serial femtosecond crystallography
SSX	serial synchrotron crystallography
SX	serial crystallography
TR-	time-resolved
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
XFEL	X-ray free-electron laser
XRC	X-ray crystallography

1. Introduction

X-ray crystallography (XRC) is a major investigative technique that is used in many fields of biology, chemistry, and materials science. In biology, its use is for elucidating the structures of small organic molecules and macromolecules. XRC has been used to solve 85 % of structures in the Protein Data Bank (Research Collaboratory for Structural Bioinformatics, 2023a). An early example of XRC in the biosciences is that of John Kendrew and Max Perutz determining the structures of myoglobin and haemoglobin, both being oxygen-binding proteins found in most vertebrates (Kendrew *et al.*, 1958; Kendrew *et al.*, 1960; Perutz *et al.*, 1960). These were the first atomic-level structures of proteins, and enabled discoveries about molecular evolution. For their work, Kendrew and Perutz were awarded the 1962 Nobel Prize in Chemistry, the same year that Watson, Crick, and Wilkins won the Nobel Prize in Physiology or Medicine for their discovery of the structure of nucleic acids, relying on X-ray fibre diffraction data from Franklin and Gosling (Franklin and Gosling, 1953b; Watson and Crick, 1953; Franklin and Gosling, 1953a; Gosling, 1954). Crystallography is still highly relevant today, for example it has provided structures of many different GPCRs (G-protein coupled receptors) with and without antagonists bound, leading to clinical trials of several pharmaceutical compounds (Congreve *et al.*, 2020).

X-ray crystallography relies on X-rays being scattered by electrons in a molecule. In a crystal lattice, where molecules associate to form a repeating unit, scattering from the electrons in each unit interacts constructively to form Bragg diffraction (Bragg and Bragg, 1913). This results in Bragg spots that can be recorded on an area detector. The pattern of spots on the detector relates to the size and shape of the repeating unit of the crystal lattice. This is termed the unit cell and is the smallest section that repeats by translation across the crystal. By using the positions of spots and measuring their intensities, and solving the phase problem, the electron density within the unit cell can be determined (Rossmann and van Beek, 1999; Howard, 1996; Winn *et al.*, 2011). This is described in more detail in §1.1.

Synchrotrons are the most common X-ray sources used in macromolecular crystallography (MX), being used for most structures in the Protein Data Bank (Research Collaboratory for Structural Bioinformatics, 2023a).

Synchrotrons are used over standalone laboratory X-ray generators due to their much higher fluxes and smaller beamsizes, allowing faster data collection and use of smaller crystals (Als-Nielsen and McMorrow, 2011; Phillips *et al.*, 1976; Robinson, 2015). Synchrotron radiation is produced by accelerating electrons to relativistic speeds. These high energy electrons are then deflected by high magnetic fields in a polygonal accelerator or storage ring. Bending magnets deflect the beam at the corners between straight sections in the ring, producing radiation at a tangent to the beam. Most currently operating synchrotrons, including Diamond Light Source (UK), however, use insertion devices for their MX beamlines. An insertion device uses a series of magnets of alternating polarity along the beam to force electrons to take an oscillating rather than straight path, producing much more intense photon beams than bending magnets (Hwang *et al.*, 2011). Modern synchrotron light sources generally have ten or more beamlines where experiments take place, many of which are dedicated to MX. Another source of high intensity X-rays for MX is the X-ray free electron laser (XFEL), the first of which, the Linac Coherent Light Source (LCLS, USA), began operating in 2009 (Emma, 2009; Emma *et al.*, 2010). Unlike synchrotrons, XFELs produce light in extremely short pulses, on the order of tens of femtoseconds (Yabashi *et al.*, 2017; Galayda, 2018; Decking *et al.*, 2020). These extremely short pulses allow for data collection free of artefacts of beam-induced damage, a phenomenon called diffraction before destruction (Chapman *et al.*, 2014; Neutze *et al.*, 2000; Chapman *et al.*, 2011), described in detail in §1.2.

As the experimental demands of researchers advance, limitations of synchrotrons and traditional crystallographic techniques for biological research are being reached. These limitations are in part caused by sample degradation over the course of X-ray exposure. Serial crystallography (SX) is a technique which has developed over the last 15 years, building off older multi-crystal data collection techniques. Modern serial crystallography uses many crystals for the collection of one dataset, rather than the general convention of using a single crystal or small number (<10) of crystals per dataset (Schlichting, 2015; Garman and Schneider, 1997). Using many crystals prevents damage from building up with dose, as each crystal only receives a single short exposure (or small number of exposures) before being discarded. A complication of collecting only a single image per crystal (called a still) is that reflections are only partially captured, rather than the whole reflection being captured over frames in a rotation series. Capturing only partial reflections makes estimating

reflection intensities harder, discussed in §1.1. Another factor that needs considering in serial crystallography is sample delivery, see §1.4. Using a fresh crystal for each exposure means that thousands of crystals need to be delivered to the beam in rapid succession. Using standard pin mounts and sample robots would make data collection impractically slow, severely limiting the throughput of synchrotron beamlines (Cipriani *et al.*, 2006; Ohana *et al.*, 2004).

1.1 Crystallographic theory

Modern macromolecular crystallography as a method is made up of several related techniques (Martin-Garcia *et al.*, 2016). First, the purified protein of interest must be crystallised in a suitable form. Properties that make a crystal suitable for diffraction data collection include being the correct size for the intended experiment and being isomorphous, that is, having the same continuous lattice form across the whole crystal (Giordano *et al.*, 2012). These crystals need to be mounted on the beamline in a way that minimises damage due to dehydration, mechanical handling or cryocooling, as well as minimising background scatter caused by exposing material other than protein crystal to the beam (Cheng, 2020; Illava *et al.*, 2021). The crystals can then be exposed to the beam to collect high-quality diffraction data and obtain the best possible structure. There are many steps involved in processing diffraction data, described below, but automated data collection and processing pipelines are increasingly capable of saving researchers' time and allowing non-experts to solve high quality protein structures.

Diffraction of X-rays by atoms in crystals is dependent on their electrons. As an X-ray wave interacts with and excites electrons within its coherence length, the electrons emit X-rays of the same wavelength. The superimposition of the waves emitted from each electron results in a wave traveling at an angle to the incident wave, and with amplitude proportional to that of the incident wave. The same summation of waves applies with electrons in adjacent atoms, including within crystals (Rupp, 2010). At certain incident angles (θ), the difference in pathlength between waves scattered off atoms in adjacent layers of the crystal lattice is equal to a multiple of the wavelength (λ), described by the Bragg equation, $2d \cdot \sin\theta = n\lambda$. In this case, constructive interference occurs, and the amplitude of the scattered wave increases, as shown in Figure 1.1, where d is the

distance between lattice planes, θ is the angle of incidence, n is an integer, and λ is the wavelength (Bragg and Bragg, 1913).

Each reflection from a crystal includes the sum of the scattering contributions of every atom in the unit cell. Reflections occur from different lattice planes with changing X-ray incidence angle, as the lattice points cross the Ewald sphere. Diffraction intensity for a given reflection is spread into a larger spot on the detector, rather than a single point, due to mosaicity in the crystal, as well as variation in the angle and wavelength of incident photons. The intensity for a reflection is also spread over a range of incident angles, so measurements involve rotating the crystal in small angular increments much less than a degree, and collecting diffraction patterns at each one, known as fine phi slicing (Pflugrath, 1999; Warren *et al.*, 2016; Mueller *et al.*, 2012). Reflections recorded across multiple images are known as partial reflections.

Reflections at higher angles, further from the angle of incidence, contain higher resolution information about the scattering atoms in the unit cell, which can be used to obtain a more detailed electron density map. The

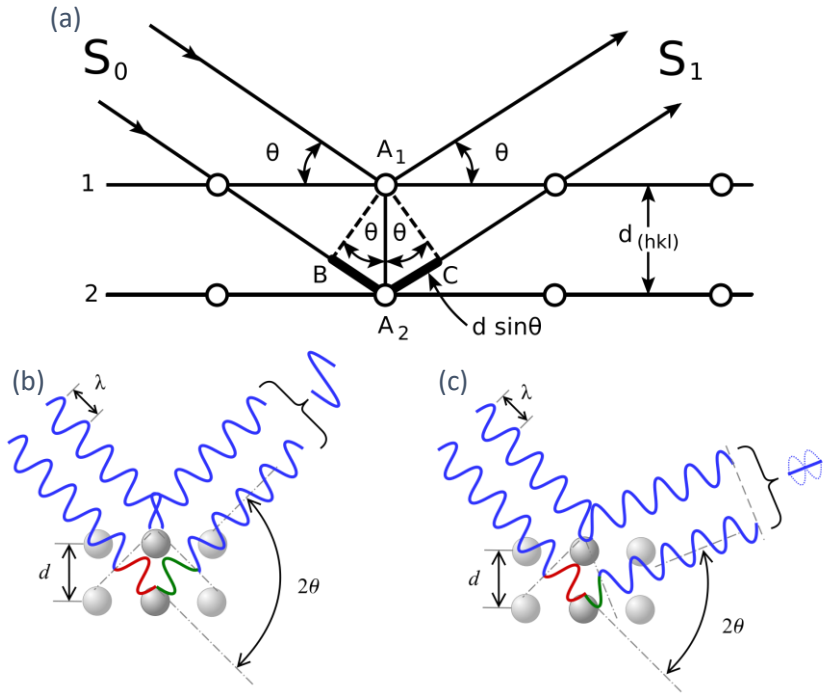

Figure 1.1: Bragg diffraction of waves from crystal lattice planes. (a) Diffraction that satisfies the Bragg condition, where the difference in path length of the two reflections is equal to a multiple of the wavelength, $2d \cdot sin\theta = n\lambda$. (b) Constructive interference occurs when the Bragg condition is satisfied, meaning the reflected waves are in phase, and (c) destructive interference occurs when difference in pathlength is not a multiple of the wavelength and the reflected waves are out of phase. Image credits: (a) ARMICRON, via Wikimedia Commons. (b) Cdang and Gregors, via Wikimedia commons, CC BY-SA 3.0.

higher resolution map enables more accuracy and confidence in fitting the protein molecule in the electron density. These higher angle reflections further from the incident angle naturally hit the detector further from the beam centre (which is usually the centre of the detector). To capture higher resolution data, the detector needs to be larger, or closer to the crystal. Detector distance is optimised to capture the highest resolution reflections from a given crystal, but while also using the whole detector area to increase precision in spot detection and measuring. Higher resolution reflections also have a lower intensity, so the signal to noise ratio drops further from the beam centre. Using a larger crystal increases diffraction intensity, improving identification of spots from high resolution reflections.

To begin interpreting the diffraction images, spot-finding algorithms are used to identify diffraction spots in the image. Spot-finding generally involves locating pixels with a high intensity, and comparing them to the intensity of nearby pixels to determine if they stand out above the background (Winter *et al.*, 2018). Indexing uses the positions of these spots to identify the space group and calculate unit cell dimensions. Once an image has been indexed, spot positions can be predicted, to ensure the measurement of all spots, not just those identified during initial spot-finding. Spot-finding and indexing of serial datasets with still diffraction patterns is similar to that for conventional rotation data collection.

The next step is integration, which measures the intensity of diffraction spots. This involves summing the counts from all pixels within the spot. To separate the intensity of the reflection from that of the underlying background intensity (for example from air scatter), local background subtraction is performed by examining the counts on pixels in a shoebox surrounding the spot. With fine phi slicing (Pflugrath, 1999{Mueller, 2012 #460)}, each image only contains a partial reflection, so the total intensity under the diffraction peak from adjacent images is summed (Van Beek *et al.*, 2012), known as 3D integration. In serial crystallography however, described in §1.3, often only one diffraction pattern is collected from each crystal, known as a still. This provides a challenge to integration, as a whole reflection is not recorded as a succession of partials over an angular range, but instead only a single partial measurement is made from the crystal. From a single partial measurement, there is no indication of which part of the peak is captured, therefore putting the intensity

measurements from different images onto common scale is difficult. This is compounded by the thin Ewald sphere shells when using monochromatic low divergence XFEL beams, which cause a narrow slice of each peak to be recorded on an image. To obtain reliable intensity measurements from partial reflections on stills, high multiplicity data is required, where each reflection is recorded multiple times. Using Monte Carlo methods allows an accurate value for the intensity of a reflection to be determined by averaging from many independent measurements from randomly oriented crystals (White *et al.*, 2012; Kirian *et al.*, 2010; Kirian *et al.*, 2011). Aiming for a multiplicity of 10 or more (each reflection recorded 10 times) is common practice in serial crystallography (Brewster *et al.*, 2019; Winter *et al.*, 2018), which equates to collecting many thousands of diffraction patterns, although the number of patterns required is dependent on the symmetry of the crystals.

Scaling involves putting the measured intensities on a common scale, accounting for factors that change during the experiment, including the illuminated volume of crystal, radiation damage, and changes in beam flux and wavelength (Evans, 2006; Beilsten-Edmands *et al.*, 2020). Scaling requires a scale factor for each reflection. Merging combines all the measurements of symmetry-equivalent reflections into a single intensity and error value and calculates metrics to express the quality of the data. Unfortunately, only the intensities of the reflections can be measured in crystallography, not the phases. The amplitudes of the structure factors are proportional to the square root of the spot intensities, $\sqrt{I} \propto |F_{hkl}|$ (Rupp, 2010; Taylor, 2010).

Data processing also involves estimation of errors in the data. for the intensities. Part of this involves the signal to noise ratio of the data, $I/\sigma(I)$, meaning the measured intensity of the peak compared to the background intensity under the peak, which indicates precision. A common measure merged data quality is the $CC_{1/2}$ value, which is the Pearson correlation coefficient comparing two halves of the dataset (Karplus and Diederichs, 2015; Karplus and Diederichs, 2012). Higher values indicate better correlation between two random halves of the data, and therefore higher precision of the data. These data quality statistics are typically calculated as a function of resolution, in order to identify a resolution cutoff for the data.

To calculate structure factors and build an electron density map, both the amplitude, F_{hkl}, and phase, ρ_{hkl}, are required (Taylor, 2010; Taylor, 2003), but only the amplitudes can be obtained from the recorded diffraction data (via intensity meaurements). This is termed the phase problem. There exist various methods for solving the phase problem to calculate the electron density from the structure factor amplitudes, and fit the known protein sequence to this density (Rossmann and van Beek, 1999). The most common method in macromolecular crystallography is molecular replacement (MR), using software such as MOLREP (Vagin and Teplyakov, 1997; Vagin and Teplyakov, 2010) or Phaser (McCoy *et al.*, 2007).

Molecular replacement relies on already having a structural model of the protein of interest (such as obtained from a different crystal form), or a homologous protein with greater than 40 % sequence identity, to use as a search model (Wlodawer, 2017), with ideally minimal structural difference to the target protein (<1.5 Å RMSD in atom positions). Patterson maps are generated from the search model and the experimental data, using the Patterson function which takes the structure factor amplitudes (which can be obtained from measured intensities) and set the phases of all reflections to zero. The search model is first rotated on three axes in reciprocal space, and scored based on the correlation of its Patterson map with that of the experimental data. After rotation, the search model is translated in three dimensions of reciprocal space. The best scoring rotation-translation is used to provide initial phases, which are combined with the experimental intensity measurements (Vagin and Teplyakov, 2010; Taylor, 2010; Rossmann and Blow, 1962). Recently, it has been shown that ab initio structure prediction tools such as Rosetta (Leman *et al.*, 2020) or AlphaFold (Jumper *et al.*, 2021) can provide search models for proteins where suitable experimental models do not exist (Simpkin *et al.*, 2019; Barbarin-Bocahu and Graille, 2022). Although AlphaFold can accurately predict the structures of most proteins in their apo state, it is weak in predicting ligand binding and providing information on protein dynamics, so serial and time-resolved crystallography are still of importance. The phase problem can also be solved experimentally using anomalous diffraction (Taylor, 2010; Rupp, 2010), although this is not discussed here.

Once a starting model has been fitted into the electron density, the model can be refined. Refinement involves making changes to positions of atoms and groups within the model to better fit the experimental electron density, while considering geometric constraints provided by prior chemical knowledge, which limit bond lengths and angles (Wlodawer, 2017; Vagin *et al.*, 2004). With higher resolution electron density maps to fit the model to, more weight can be applied to the experimental map, overriding the ideal geometry provided by restraints (Urzhumtsev and Lunin, 2019). Refinement can be performed in both real space, where refinement is against a $2F_o - F_c$ electron density map, such as in *Coot* (Emsley *et al.*, 2010), and in reciprocal space against F_o, as done by REFMAC5 (Murshudov *et al.*, 2011). With each round of refinement, the model improves its fit to the observed electron density. As the model improves, so the phases calculated from the model improve and become closer to the experimental phases (which were not measured). Therefore, the electron density can be recalculated with the new phases and more refinement performed in an iterative cycle.

As well as the positions of the protein atoms, other molecules can be added and refined within the model, including bound water molecules and ligands. These ligands can be relevant to the protein function, such as substrates in the active site, or be part of the crystallisation conditions, such as sulfate ions or glycerol. In addition to the positions of the atom, their occupancies can also be refined, along with the B-factors, which are described below. Occupancies, which are values in the range 0–1, represent the proportion of unit cells containing a given atom. For the peptide component of the model, this is usually assumed to be one for all atoms, except in the case of multiple conformers. Refinement is performed with the maximum likelihood method using software such as REFMAC5 (Murshudov *et al.*, 2011) and *Coot* (Emsley *et al.*, 2010).

Although crystals consist of molecules fixed in a lattice, some constrained movement of atoms oscillating around an equilibrium position is allowed (Carugo, 2018). Motion of an atom within a crystal attenuates its structure factor and appears in an electron density map as a spreading of the electron density for the atom and reduction of the peak density, to represent the variation in atom position over time and over the different unit cells in the crystal. Atomic B-factors (herein referred to simply as B-factors) are a measure of the

magnitude of displacement of a given atom from its average position, with units of Å². B-factors are calculated using the equation $B = 8\pi^2 u^2$, where u represents the mean amplitude, in Å, of displacement of the atom (Sun *et al.*, 2019). Higher values represent atoms with less well-defined positions (Trueblood *et al.*, 1996; Nelson, 2019). Atoms with the same local environment, including closely bonded in the same molecule, or adjacent in space, are likely to be similar in their disorder and amount of allowed thermal motion, so will have similar B-factors. Occupancy of an atom within the structure also influences B-factor. Where the specified occupancy of the ligand in the model is higher than the occupancy suggested by the experimental electron density, the B-factors for the ligand atoms increase to represent increased disorder and motion to account for the reduced electron density. This interaction between occupancy and B-factors is used to develop a method, detailed in §5.2.3, for determining occupancy of a nitric oxide ligand in protein structures.

The relationship between B-factor and occupancy has been used in quantifying radiation damage present in datasets by examining B-factors of aspartate and glutamate residues. These residues get decarboxylated as part of the site-specific damage induced by the X-ray beam at lower doses than where global damage becomes evident by loss of diffraction resolution (see §1.2). Where sidechain carboxyl groups on these residues are modelled at 100 % occupancy, but the experimental electron density indicates some of them are lost, their B-factors increase. Comparing these B-factors to those for neighbouring atoms allows detection of site-specific damage in a structure and is used to calculate the B_{Damage} and B_{net} metrics (Shelley and Garman, 2022; Gerstel *et al.*, 2015).

Errors in the experimental data manifest as noise in the electron density map after data reduction. When viewing maps during model building and refinement, the displayed contour level can be changed to hide low level signal likely to be noise. Maps are contoured by RMSD (root mean square deviation) of the points from the average electron density value of the map. Typical contour levels are 1σ for $2F_{obs} - F_{calc}$ and ±3σ for $F_{obs} - F_{calc}$ maps. To quantify how well the refined model fits the experimental data, the structure factor amplitudes, $F_{calc}\ and\ F_{obs}$, are compared using the R value (also known as R_{work}), which is calculated using $\sum|F_{obs} - F_{calc}|/\sum F_{obs}$. A properly refined model should have an R value less than 20 %. The R_{free} value is an

equivalent measure calculated from a subset (usually 5 %) of reflections not used in refinement, and protects against overfitting of the model to the experimental data. The R

1.2 X-ray damage to biological samples

One of the main complicating factors in macromolecular crystallography is damage to crystals caused by exposure to the X-ray beam. Of the incident radiation hitting the crystal, approximately 98 % passes through without interacting, and only 0.15 % undergoes elastic scattering to provide useful diffraction information. The remaining photons are involved in damage processes within the crystal. Inelastic Compton scattering occurs with 0.15 % of the incident photons, meaning they transfer some of their energy to electrons in the crystal. The remaining 1.7 % of photons are absorbed via the photoelectric effect, with their energy being transferred to an inner shell electron, causing its ejection from the atom, shown in Figure 1.2a. These ejected electrons, called photoelectrons, are responsible for more than 90 % of energy absorbed by the crystal, and travel until they interact with other atoms in the crystal and cause ionisation damage, as described below. Each photoelectron has the potential to cause approximately 500 further ionisation events as secondary damage (Nave and Garman, 2005; Nass, 2019), shown in Figure 1.2b.

Beam damage to crystals means data quality rapidly decreases as exposure accumulates, with the dose being measured in Grays (Gy), equal to 1 joule of absorbed energy per kilogram of matter (1 $J \cdot kg^{-1}$). X-ray exposure drives two types of damage processes in protein crystals global damage to the crystal, and specific molecular damage to the protein molecules. Global damage is evident in loss of diffraction power and increased B-factors. These are brought about by unit cell changes are the result of increased disorder and displacement of atoms within the unit cell (Carugo, 2018; Trueblood *et al.*, 1996). The decrease in crystalline disorder results in a reduction of the maximum resolution of collected reflections (Burmeister, 2000; Holton, 2009). The rate of accumulation of global damage is quantified by the absorbed dose at which diffraction power has reduced by half ($D_{1/2}$). This limit was calculated from electron diffraction data to be 2×10^7 Gy (Henderson, 1990), and is known as the Henderson limit. The limit, known as the Garman limit, has been experimentally determined

with X-ray data at 100 K to be 4.3×10^7 Gy (Owen *et al.*, 2006), or approximately 1×10^6 Gy at room temperature.

Site-specific damage occurs at much lower doses than global crystal damage and induces structural and chemical changes in the crystal (de la Mora *et al.*, 2020; Pfanzagl *et al.*, 2020). With specific damage, biological integrity of structures obtained by X-ray diffraction can be compromised, as the structure of the protein changes while it is being exposed to the beam (Garman, 2010). Such damage occurs at much lower doses, on the order of 10^4–10^5 Gy at both cryogenic and room temperatures (Gotthard *et al.*, 2019), for reasons not yet fully understood. The same research that came to the 4.3×10^7 Gy limit at cryogenic temperatures also suggests a lower dose limit for data collection of 3.0×10^7 Gy, corresponding with $D_{\ln 2}$, rather than $D_{1/2}$ to minimise the manifestation of site-specific damage (Owen *et al.*, 2006).

Chemical damage commonly observed includes reduction of disulfide bonds and cleavage of certain amino acid side chains such as carboxyl groups, but importantly charged residues can get reduced, for example redox centres with flavins or metal ions also get reduced (Burmeister, 2000; Carugo and Carugo, 2005; Yano *et al.*, 2005; Owen *et al.*, 2011; Bhattacharyya *et al.*, 2020; Bowman *et al.*, 2016; Becker *et al.*, 2010). Some of the X-ray radiation is absorbed by water present in the crystal, forming oxoniumyl ions (H_2O^+) and hundreds of solvated electrons per absorbed photon (e^-, scheme 1). These solvated electrons are short-lived and undergo various reactions with water and oxygen to produce more reactive oxygen species (ROS) such as superoxide ($O_2{\cdot}^-$, scheme 2) and hydrogen peroxide (H_2O_2, scheme 3).

1. $H_2O + h\nu \rightarrow H_2O^+ + e^-$
2. $O_2 + e^- \rightarrow O_2{\cdot}^-$
3. $O_2{\cdot}^- + e^- + 2H^+ \rightarrow H_2O_2$

These ROS, and the solvated electrons themselves, react with the protein molecules to cause damage (Cadenas, 1989; Novo and Parola, 2008). In metalloproteins, these effects occur more rapidly at the active site where the metal is bound, which are also the areas of greatest biological interest and importance to study. Redox centres in proteins often contain transition metal ions such as iron and copper (see §1.7 for detailed

examples), which accept the solvated electrons and become reduced (Horrell *et al.*, 2018). For metalloproteins with Fe^{3+} ions (including haem proteins studied here), a dose limit of 40 kGy has been determined, corresponding to photoreduction of 50 % of the iron sites in the crystal (Pfanzagl *et al.*, 2020).

When avoiding site-specific damage in structures is of high importance, room temperature data collection can be advantageous. At cryogenic temperatures, the rate of specific damage, such as disulfide bond cleavage or metal centre reduction, are largely decoupled from the rate of global damage (and loss of diffraction) by a factor of 100 or more times (Gotthard *et al.*, 2019). This decoupling means that structural and biochemical changes to the protein take place long before the crystal would be discarded, or the data truncated due to loss of diffraction. Therefore, effects of specific damage are frequently present in structures obtained at 100 K, which affects how well they represent the native state of the protein and can hinder interpretation of the function of the protein in its biological context. At room temperature, although both global and site-specific damage are likely to occur at lower doses than cryogenic temperatures, the rates of the two forms of damage are likely to be closer to each other. For example, in data collected from lysozyme at 100 K, the rates of global damage and disulfide bond cleavage differ by a factor of 21, with site-specific damage observed at a dose 21 times lower than global damage. At room temperature (293 K), this is reduced to a factor of 1.3, showing almost complete coupling of the rates of the two types of damage (Gotthard *et al.*, 2019). This means that specific damage in structures collected at room temperature can more easily be tracked by monitoring loss of diffraction. Later images with a higher dose and lower diffraction resolution can be truncated before merging, therefore removing images where specific damage is also likely to be present. Although monitoring loss of diffraction reduces the amount of specific damage seen, specific changes can also be tracked during data collection with other techniques where available, such as using UV-Vis, Raman, and X-ray absorbance and emission spectroscopies for validating metal centre reduction (Hough *et al.*, 2008; Kekilli *et al.*, 2017a) and disulfide bond cleavage (Weik *et al.*, 2002).

Data collection temperature can influence the dose limit by the dose rate effect, which is only present at room temperature. At cryogenic temperatures, dose rate (measured in $MGy \cdot s^{-1}$) has little effect on beam damage

(Leiros *et al.*, 2006; Sliz *et al.*, 2003; Owen *et al.*, 2006; Owen *et al.*, 2012), but at room temperature, there is a positive correlation between the dose rate and the dose limit ($D_{1/2}$), for example a four-fold increase in $D_{1/2}$ was seen when dose rate was increased by 60 % (Southworth-Davies *et al.*, 2007). This means better data comes from faster data collection with higher flux (Warkentin *et al.*, 2017). Another set of experiments found a 58 % increase in $D_{1/2}$ at room temperature when going from a dose rate of 0.36 Mgy·s^{-1} to 40.3 Mgy·s^{-1} (de la Mora *et al.*, 2020). This dose rate effect is proposed to be due to the build-up of various radicals (such as hydroxyl radicals, •OH), generated by X-ray photoelectric absorption and subsequent relaxations, which form at higher concentrations with higher dose rates (Hill and Smith, 1994). At higher concentrations, these radicals are more likely to recombine with each other, rather than interact with the protein. The dose rate effect is also partially observed through us of shorter exposure times with higher flux rates, so data is collected before the radicals have had time to fully diffuse through the crystal (Owen *et al.*, 2012). At cryogenic temperatures, the radicals are trapped, unable to diffuse through the crystal, therefore dose rate has little effect (Zakurdaeva

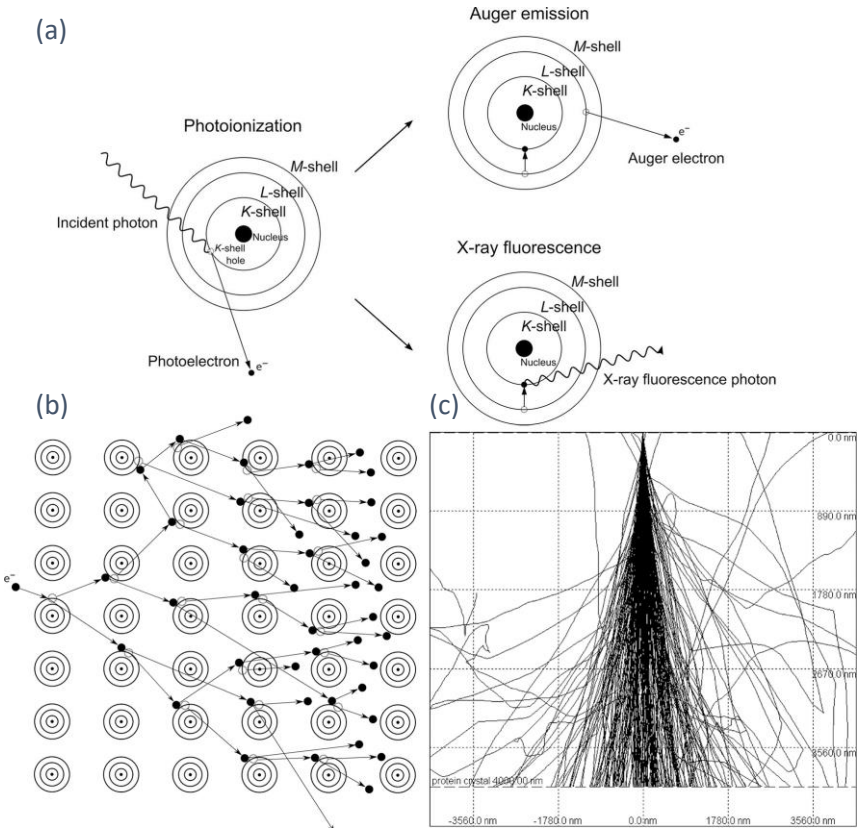

Figure 1.2: (a) Diagram showing the formation of photoelectrons (left), and the subsequent damage processes that can occur downstream. (b) A cascade of ionisation events resulting from a single photoelectron. Each photoelectron can cause another 500 ionisation events. (c) Simulated paths of photoelectrons formed from 30 keV X-rays absorbed 4 µm from the rear surface of a protein crystal. (a, b) Reproduced from Nave and Hill (2005), (c) reproduced from (Conrad *et al.*, 2015), both with permission of the IUCr.

et al., 2005; Brivati *et al.*, 1969). One advantage cryogenic data collection has over room temperature is that radical species released within cryocooled crystals are confined to their local environment, therefore damaging only the exposed volume of crystal and a small distance outside it. This allows data collection from several regions of a crystal independent of damage to other regions. At room temperature, the radicals are free to diffuse throughout and damage the whole crystal.

The development of microfocus beamlines at synchrotrons in the early 2000s enabled use of smaller crystals in MX data collection. Examples include X06SA at Swiss Light Source (SLS, Switzerland), with a 25 × 5 µm beam, ID232 at the European Synchrotron Radiation Facility (ESRF, France), with a 7.5 × 5 µm beam (Flot *et al.*, 2010; Pohl *et al.*, 2006) and I24 at Diamond Light Source (DLS, UK) with a 7 × 7 µm beam. Being able to use smaller crystals is advantageous because smaller crystals are often easier to grow than larger ones, reducing the amount of time and precious protein consumed in optimising crystal growth conditions (see §1.5). Unfortunately, with smaller crystals, less data can be collected from each of them, since absorbed dose is a function of mass. To collect enough data for a high-quality structure from smaller crystals, often multiple crystals are therefore required. Additionally, microfocus beams used for data collection from smaller crystals contain the same number of photons as a wider beam, but spread over a smaller area, so the absorbed dose of the illuminated crystal volume increases. A microfocus beam is used with the aim of illuminating only the crystal, reducing the amount of extraneous matter in the beam, which would contribute to background scatter and reduce signal to noise, making weaker diffraction information harder to process. Microfocus beamlines are also useful for illuminating only the highest quality parts of a larger crystal, avoiding inhomogeneous areas. This can be done by low dose pre-screening of crystals for diffraction prior to full data collection (Bowler *et al.*, 2010; Zhang *et al.*, 2006; Evans *et al.*, 2011). For the purposes of this thesis, microcrystals are defined as crystals less than 20 µm across (Cusack *et al.*, 1998; Martin-Garcia *et al.*, 2016), although larger crystals can also be used in multi-crystal serial experiments.

Another benefit of using microcrystals is the potential to exploit photoelectron escape. Photoelectron escape is where photoelectrons reach the edge of the crystal before being absorbed and causing further damage. The

path length of a photoelectron is dependent on its energy and the material it is travelling through. In a protein crystal, a typical path length for a photoelectron generated from a 25 keV beam is ~10 µm (LaVerne and Pimblott, 1997; Nave and Hill, 2005). With a given beam energy and crystal composition, the proportion of photoelectrons escaping is dependent on the crystal size. In microcrystal diffraction experiments, the crystal dimensions are similar to the vector distance of the photoelectrons, so a significant proportion of photoelectrons can escape the crystal without being absorbed (Storm *et al.*, 2020; Storm *et al.*, 2021). Results of a simulation are displayed in Figure 1.2c, where a 30 keV beam hits a 4 µm thick crystal. Similar simulations of 2.5 µm crystals with a 25 Å unit cell have shown that moving from 12 keV to 25 keV X-rays increases crystal lifetime by seven times (Cowan and Nave, 2008). The reduced quantum efficiency of silicon detectors at higher energies offsets the benefits of increased photoelectron escape on reducing dose absorption (Murray *et al.*, 2004; Dickerson and Garman, 2019). For example the DECTRIS PILATUS detector, with 450 µm thick silicon sensors has a quantum efficiency of 98 % at 10 keV, but only 33 % at 20 keV (Donath *et al.*, 2013). Recently developed cadmium telluride (CdTe) detectors with 1 mm sensor thickness have a near 100 % quantum efficiency up to 26 keV (Zambon *et al.*, 2018), so are better for high energy data collection. The recent addition of a DECTRIS EIGER2 9M CdTe at beamline I24 allows high energy data collection to be exploited for microfocus and serial crystallography.

1.3 X-ray free-electron lasers and serial crystallography

X-ray free-electron lasers (XFELs) are capable of much higher peak brilliance than synchrotrons, in terms of the rate of photons delivered to the sample, approximately 10^9 times higher (Emma *et al.*, 2010). Peak brilliances on the order of 10^{27} photons/s/mm/mrad²/0.1% bandwidth have allowed, in the last 10 years, data to be acquired from a crystal in shorter periods, on the femtosecond timescale. XFELs are different from synchrotrons in that they are linear, with electron bunches being used once, rather than making many laps of a storage ring. They consist of an injector, a long (approximately 1 km) electron accelerator, followed by a series of undulators, of similar design to those at a synchrotron but with a combined length of approximately 100 m, and finally several beamlines positioned downstream (Galayda, 2018; Weise and Decking, 2017; Hara *et al.*, 2016). XFELs produce a pulsed beam, usually at tens of Hertz (Yabashi *et al.*, 2017), although newer

facilities such as the European XFEL and LCLS-II (under commissioning) operate at megahertz repetition rates (Galayda, 2018; Mancuso *et al.*, 2019). This contrasts with synchrotrons, which produce a near-continuous beam, so exposure times and image collection rates are dependent on the specific source being used. At a synchrotron, the electrons are grouped into bunches to travel around the ring, thereby producing a pulsed beam. For example, at Diamond Light Source, the most commonly used fill pattern of electron bunches includes a 'no-beam' period of 72 ns during each repeating period of 1872 ns. This no-beam period is on a much faster timescale than the millisecond exposure times per frame for MX data collection and therefore for most experiments (including all experiments carried out as part of this thesis) the synchrotron beam can be considered continuous. XFEL pulses typically last tens of femtoseconds, and a single pulse contains approximately the same number of photons as a synchrotron beamline delivers in a second. The brief but very intense nature of the femtosecond pulses produced by XFELs destroys crystals but prevents beam induced site-specific damage from affecting collected diffraction patterns. This is due to data being collected on a timescale shorter than that over which the damage processes take place, a technique referred to as diffraction before destruction (Barty *et al.*, 2012; Boutet *et al.*, 2012; Chapman *et al.*, 2006; Chapman *et al.*, 2011; Neutze *et al.*, 2000). This is true for very short exposure times of 10 fs, but some damage may appear at exposure times of tens of femtoseconds (Chapman *et al.*, 2014; Nass *et al.*, 2015).

This data collection differs to synchrotrons, where exposure times are in the region of 10 ms per frame, and even at room temperature, crystals can survive for several images. Use of XFELs for serial crystallography is called serial femtosecond crystallography (SFX), whereas at synchrotrons it is called serial synchrotron crystallography (SSX). The destruction of crystals within the XFEL beam necessitates serial crystallography methods, where each crystal is used for a single (still) diffraction image, and a fresh crystal is provided for each exposure, described in §1.4.

As SX generally uses a fresh crystal for each exposure, beam damage to the crystals is of lesser importance. This means that SX can be performed with the crystals at room temperature, as cryogenic temperatures are not required to slow damage processes. As well as the practical advantages of not having to keep crystals cool

on the beamline or spend time optimising cryoprotection (to prevent increased mosaicity and other cooling-induced defects), collecting structures of biomolecules at room temperature rather than 100 K is of more physiological relevance. The increased molecular motion at room temperature also allows for reactions to take place in the crystals (Burnley *et al.*, 2012; Ramakrishnan *et al.*, 2021), which can be investigated in time-resolved SX (TR-SX), see §1.6.4. These benefits are traded off against the major reduction in data that can be collected from one crystal before unacceptable crystal damage.

1.4 Serial sample delivery

Because of the high numbers of crystals used in many SX experiments, rather than the one or two large crystals used in conventional crystallography, mounting magnetic pins with loops onto a goniometer on a beamline is generally not suitable. The throughput of exchanging pins, even with a robotic sample changer, is far too low to allow collection of data from thousands of crystals (Martin-Garcia *et al.*, 2016; Schlichting, 2015). Therefore, alternative methods must be sought to rapidly deliver random oriented microcrystals successively to the X-ray beam.

Sample delivery methods must be optimised in such a way as to reduce the mechanical and physical stress put on the crystals. Protein crystals only have a small number of intermolecular bonds relative to their volume, which allows flexibility of molecules within the crystal, and for the molecules to still perform their binding and catalytic roles. This also makes the crystals sensitive to mechanical stress, changes in mother liquor composition, temperature fluctuation, and dehydration (Schlichting, 2015; Cheng, 2020). Another aspect to consider with delivery methods is the pulsed nature of X-ray beams at XFEL sources, and whether all crystals would be exposed to X-rays. Crystals that do not get exposed to the beam are wasted, therefore increasing the amount of sample required to get a structure. In the first 10 years of serial crystallography, several methods have been used successfully, each suited to different sample and experiment types, and the repetition rates of different X-ray sources (different XFEL facilities provide different pulse lengths and frequencies, sometimes with pulse grouping). These techniques can mostly be grouped into three categories: liquid jets, viscous extrusion, and fixed targets.

1.4.1 Liquid jets

Liquid jets involve firing crystals in their mother liquor out of a nozzle and through the X-ray beam, powered by a high-performance liquid chromatography (HPLC) pump or gas pressure. Liquid jets consume large amounts of sample, as only 1 in 10 000 crystals in the fast-flowing jet typically interact with the X-ray beam (DePonte *et al.*, 2008; Shapiro *et al.*, 2008), assuming an XFEL pulsing at tens of Hertz. A common type of injector is the gas dynamic virtual nozzle (GDVN), with a minimum flow rate of 10 $\mu L \cdot min^{-1}$. This consists of a sample nozzle inside the tapered tip of a larger gas capillary, so the sample is released coaxially into a gas stream. The gas keeps the sample jet stable while focusing it down to a smaller diameter, approximately 50 µm, to reduce sample consumption over a basic jet (Calvey *et al.*, 2016; Oberthuer *et al.*, 2017). A double flow focus nozzle (DFFN) relies on the same principle, but with an extra liquid solvent layer between the sample and the gas (Calvey *et al.*, 2016; Olmos *et al.*, 2018). Liquid jets can be adapted to perform time-resolved experiments by either light activation shortly before X-ray exposure, or by mixing of the crystal slurry with a reactant solution upstream of the nozzle (Schmidt, 2013). Because of the high jet speed (approximately 10 $m \cdot s^{-1}$), liquid jets are only able to access very short timepoints in a reaction, on the order of nanosecond or microseconds (Aquila *et al.*, 2012; Standfuss, 2017).

1.4.2 High viscosity extrusion

High viscosity extruders are like liquid jets, in that they produce a stream of crystals from a nozzle, and this travels through the X-ray beam. If the crystal slurry has a higher viscosity, like toothpaste, a stable stream can be formed at a much lower flow rate (Botha *et al.*, 2015). Lower flow rates mean reduced sample consumption and enable longer timepoints for TR-SX than liquid jets, as the crystals remain in the stream for longer, allowing more time between reaction initiation and diffraction data collection (Weierstall *et al.*, 2014). Extruders use a HPLC pump delivering water to drive a piston plunger connected to a sample reservoir, which amplifies the pressure and extrudes the viscous medium at a much lower flow rate than the water (Conrad *et al.*, 2015; Sugahara *et al.*, 2017). High viscosity extruders often still use a coaxial gas stream to steady the extrusion and

keep it aligned with the beam. The sample reservoir sits directly behind the short (50 mm) capillary in the extruder to reduce wastage of sample left in the capillary.

Media such as hydroxyethyl cellulose (HEC) or agarose can be added to the crystal mother liquor, with demonstrated flow rates of 0.45 µL·min^{-1} and 0.16 µL·min^{-1} respectively (Conrad *et al.*, 2015; Sugahara *et al.*, 2017). High viscosity is also achievable by growing crystals in lipidic cubic phase (LCP), known as *in meso* crystallisation, which is a common method for crystallisation of membrane proteins (Weierstall *et al.*, 2014; Aherne *et al.*, 2012). Viscous media has also been prepared from lard (Nam, 2020). Unfortunately, the high viscosity media surrounding the crystals scatters X-rays more than air, resulting in higher background counts on the detector. The correct choice of medium is important for maintaining crystal integrity, while still allowing a suitably low flow rate, and limiting background scatter (Park and Nam, 2023).

1.4.3 Fixed targets

Serial crystallography can also make use of solid mounts that hold crystals in fixed positions. Because the crystals are stably mounted, not transiently passing through the beam as a jet or extrusion, there are no wasted crystals that are supplied between beam pulses as with liquid jet delivery methods. Fixed targets are designed to support crystals with minimal mother liquor and or other material in the beam, causing only minimal background scattering. For further reduction of background scattering, many fixed target designs incorporate apertures that allow the beam to pass through. The crystal slurry is applied to the fixed target and crystals settle at these apertures (Ebrahim *et al.*, 2019a; Ebrahim *et al.*, 2019b; Oghbaey *et al.*, 2016; Sherrell

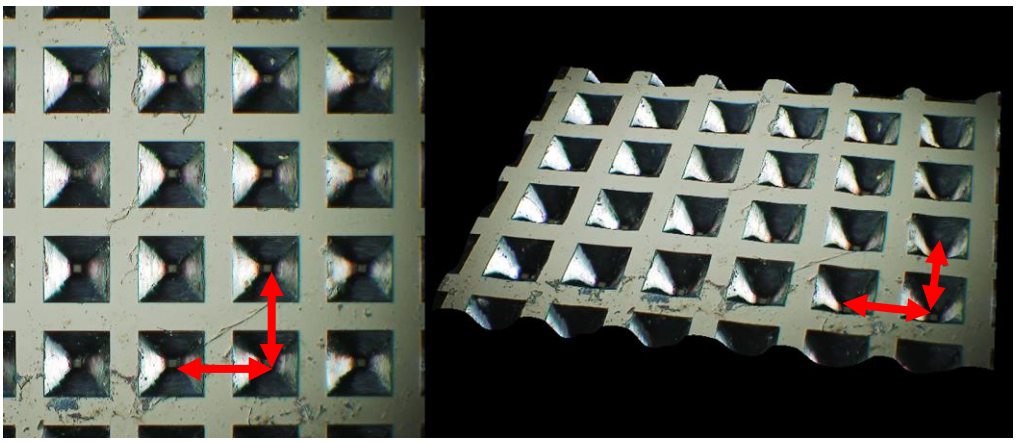

Figure 1.3: Micrographs of the wells of a fixed target chip. The tapered square wells with 10–30 µm apertures can be seen (left), with a 3D model showing the taper of the wells (right). Each well is spaced 125 µm on centres (red arrows). Images produced using a Hirox HRX-01 microscope.

et al., 2015; Horrell *et al.*, 2021), as shown in Figure 1.3. Because a much higher proportion of the crystals get exposed to the beam and used, the sample requirement is much lower than for other delivery methods. Crystals are moved into and out of the X-ray beam by translation of (the fixed target, usually on a precise motorised translation stage (Ebrahim *et al.*, 2019a; Ebrahim *et al.*, 2019b; Sherrell *et al.*, 2015). Moving the stage allows different crystals on the target, or different volumes of a larger crystal, to be positioned in the X-ray beam consecutively. As the crystals are held in place on the mount, there is no time limit to how long a crystal can remain in the path of the beam, nor on how many times it can be brought back into the beam. Removing these limitations enables additional experimental possibilities, including longer timepoints for time-resolved experiments. Multiple or prolonged exposures of one crystal can be used for investigation of X-ray-induced damage and X-ray-initiated reactions (Doak *et al.*, 2018; Lee *et al.*, 2019), described in §1.2 and §1.6.6.

Some fixed target designs involve applying crystal slurry across a surface that is transparent to X-rays (such as a polyimide or Mylar® film), with crystals distributed in unknown positions on the surface (Murray *et al.*, 2015; Lee *et al.*, 2020), while others use a linear slot to confine crystals into known positions in one dimension (Hunter *et al.*, 2014; Park and Nam, 2023). Data collection positions on the fixed target can be spaced arbitrarily since crystals are not confined to specific points on the surface. The only requirement is that data collection locations are spaced far enough apart to prevent beam damage from X-ray exposure at one position

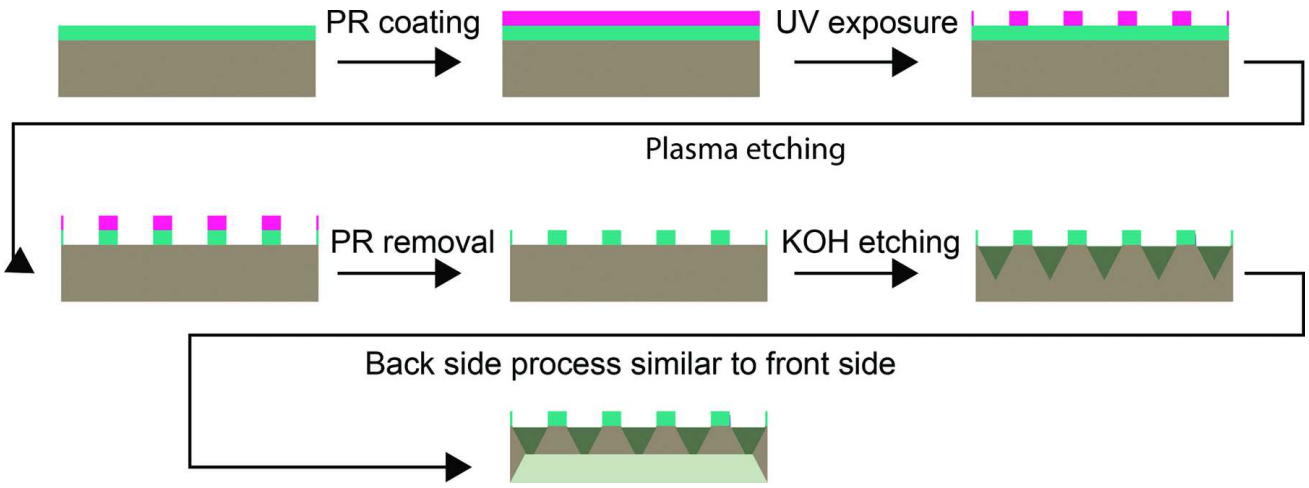

Figure 1.4: Manufacturing of a fixed-target chip. A silicon wafer is coated in silicon nitride (SiN, green) and photoresist (magenta), then the pattern of features is made by UV exposure. Plasma etching is used to transfer the pattern onto the SiN, after which the photoresist is removed. The SiN forms the mask for wet etching of the silicon layer with potassium hydroxide. The process is repeated for the back side of the chip. Reproduced from Oghbaey *et al.* (2016) with permission of the IUCr.

impacting- the next, especially when using an XFEL (Hunter *et al.*, 2014). Crystals can be kept hydrated on the surface by embedding in oil or LCP (Li *et al.*, 2018), or by sandwiching the crystals between thin films (Doak *et al.*, 2018).

Other fixed targets trap crystals in predefined positions, with the fixed target rastered between these positions. An example of this is the tapered apertures (wells) in the fixed target silicon chips developed at Diamond Light Source, Hamburg and Toronto (Sherrell *et al.*, 2015; Owen *et al.*, 2017; Mueller *et al.*, 2015). Having crystals held at known positions on a fixed target increases hit rates over other fixed targets, with more crystals being exposed to the beam and having diffraction data collected. The fixed-target approach used at beamline I24 at Diamond involves loading crystals into apertures in a square silicon chip approximately 30 mm in size (Oghbaey *et al.*, 2016; Sherrell *et al.*, 2015; Mueller *et al.*, 2015). Chips are made with a standard silicon wafer etching technique, see Figure 1.4. This approach greatly lowers sample requirement, and provides higher hit rates, reducing the amount of beamtime needed for data collection.

Several silicon chip designs exist, with different arrangements of wells and fiducials for alignment on the beamline. Wells on chips are arranged in rows and columns within groups known as blocks. An early chip design was 28 mm square, with a 9 × 9 arrangement of 12 × 12 well blocks, with five wells of each block

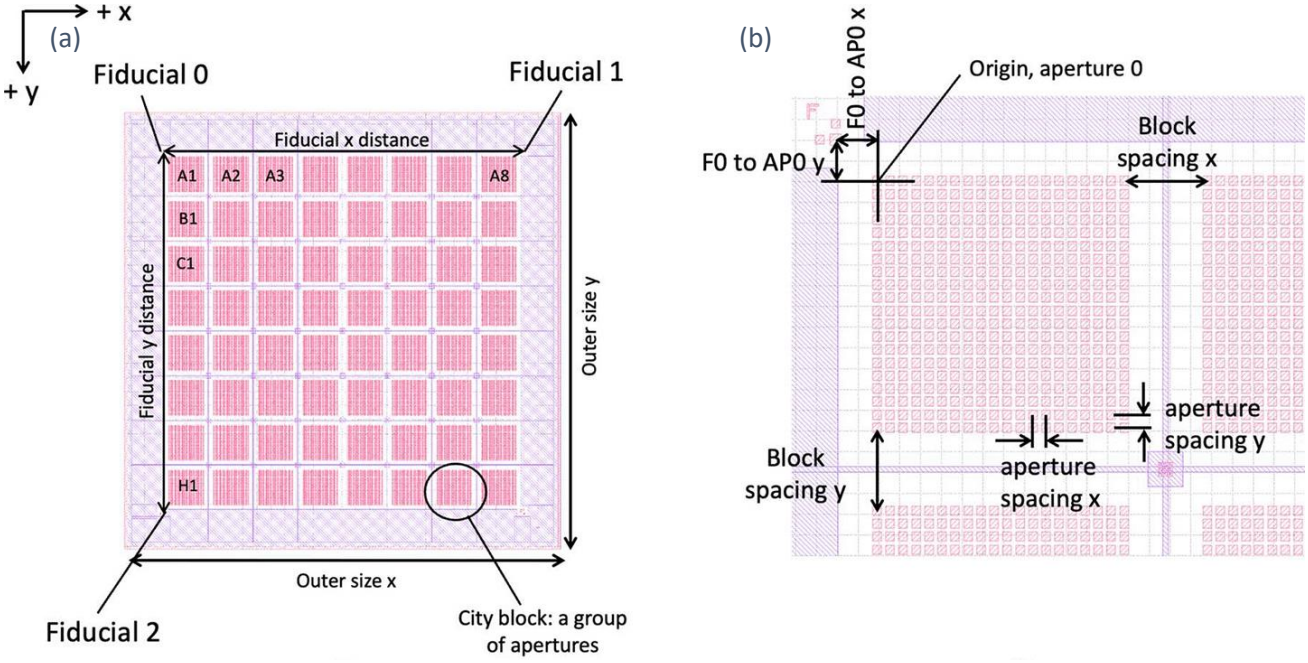

Figure 1.5: The layout of an Oxford chip, with (a) a grid of 64 blocks numbered A1–H8, along with fiducial markers in each corner for aligning the chip. A closeup of one corner (b) shows one block containing a 20 × 20 grid of 400 tapered apertures, spaced 125 µm between centres.

replaced with hash-shaped (#) fiducials, leaving 11 259 total useable wells (Oghbaey et al., 2016). The Oxford chip shown in Figure 1.3 and Figure 1.5 uses a 30 mm square layout, with an 8 × 8 layout of 20 x 20 compartments (25 600 wells total), with fiducials in the corners of the chip, outside the grid wells (Ebrahim et al., 2019a; Owen et al., 2017; Horrell et al., 2021) area. The hit-and-return (HARE) chip is approximately the same size as the Oxford chip, but comprises 6 × 6 compartments of 24 × 24 wells (20 736 total). The different chip layouts affect the time delays available for TR-SX experiments (Mehrabi et al., 2020; Schulz et al., 2018).

Chips can be loaded by pipetting a volume of crystal slurry onto the chip and spreading it across the surface. Applying a weak vacuum to the underside then draws the slurry down into the tapered apertures, shown in Figure 1.3, where the crystals are trapped and unable to pass through the chips. The excess mother liquor passes through the chip into a waste flask, with the process shown in Figure 1.6a. Care must be taken to ensure the apertures in the chip are smaller than the crystal dimensions, otherwise the crystals will be lost through the chip with the mother liquor. An even distribution of crystals across the chip results in improved chip loading, with more wells with one crystal in them, rather than multiple loading near the centre of the chip. Crystal distribution is made easier if the chips are glow-discharged prior to loading, as making the surface hydrophobic allows slurry to be spread more easily across the surface, similar to preparation of cryo-EM grids (Drulyte *et al.*, 2018). Because crystals are distributed more easily, a lesser volume of slurry is needed to load the chip. With this method, a 30 mm square chip with 25 600 wells can be loaded with 100–200 µL slurry, resulting in a typical hit rate of 30 %, that is approximately 7700 indexed images per 100 µL of slurry (Mehrabi *et al.*, 2020; Murray *et al.*, 2015; Oghbaey *et al.*, 2016; Horrell *et al.*, 2021).

To avoid dehydration of crystals during loading, loading can take place in a high humidity environment. After loading, chips are often sandwiched between two layers of thin (2.5–10 µm) film, preventing evaporation of the liquid surrounding the crystals (Hunter *et al.*, 2014). These films are specifically chosen for their X-ray transparency to reduce background scatter (see §3.9). Mylar®, made from biaxially-oriented polyethylene terephthalate, and Kapton® (polyimide), are two commonly used films. For spectroscopy of crystals in the chip, or when using light to initiate reactions in TR-SX, transparency to visible and UV light is required. When using

crystals that are sensitive to oxygen, for example redox proteins, permeability of films to oxygen needs considering. Alternatively, crystals can be embedded in a cryoprotectant oil such as Parabar 10312 (formerly Paratone-N), leaving a small amount of mother liquor around each microcrystal (Zarrine-Afsar *et al.*, 2012; Horrell *et al.*, 2021). Because of their viscosity, crystals in cryoprotectant must be manually spread onto the chip, rather than making use of the above-described vacuum method. Cryoprotectants can also be protective against oxygen reaching the crystals, with the added benefit of enabling in vacuo measurements to be taken.

UV-Vis spectroscopy of crystals can be performed shortly before or after crystallography data collection, in order to identify oxidation states of redox centres within proteins, or to identify ligand binding states, and ensure crystal structures are collected from the proteins in relevant states. Optical spectroscopic monitoring of protein redox state has been extensively demonstrated for single crystal experiments (Kekilli *et al.*, 2014; Owen *et al.*, 2011; Kekilli *et al.*, 2017a; Beitlich *et al.*, 2007). Identifying redox states before and after data collection can be used to monitor radiation damage within the crystals. In addition to monitoring crystal states, when performing serial crystallography, more efficient use of beamtime could be made by identification of crystal positions within the chip using offline spectroscopy prior to X-ray exposure. This map of crystal positions could be used to skip diffraction data collection from unoccupied wells, saving precious beamtime,

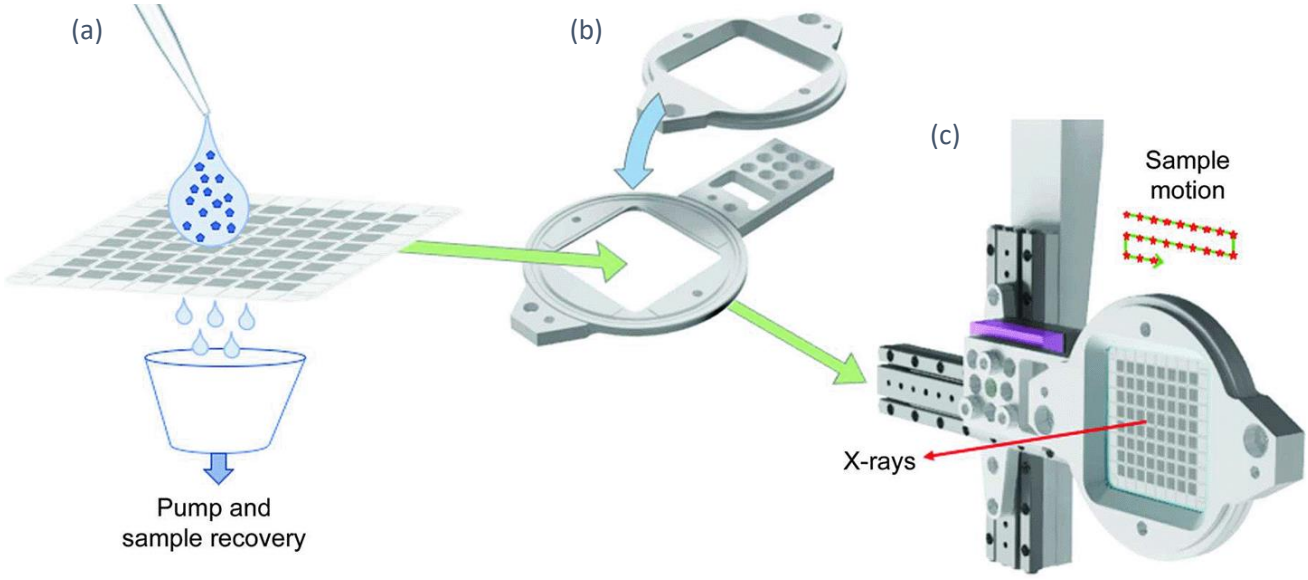

Figure 1.6: Schematic of the process of loading a fixed target chip. Crystal slurry is pipetted onto a chip (a), and the excess mother liquor is removed through the apertures using weak vacuum pressure. After assembling in a two-part holder (b), the chip is mounted on a motion stage (c). Crystals on the chip are exposed to the beam in turn by moving the chip between positions. Reproduced from Ebrahim *et al.* (2019a) with permission of the IUCr.

although this is yet to be implemented on a beamline (Oghbaey *et al.*, 2016). The same spectroscopy setup could be used to monitor redox states of crystals within chips.

1.4.4 Additional sample delivery methods

Flows of crystal suspension can be manipulated by enclosing them within small tubing in microfluidic devices, such as a glass capillary placed across the X-ray beam (Monteiro *et al.*, 2020; Lyubimov *et al.*, 2015). Use of a capillary removes complications with getting a stable jet from a nozzle at the expense of increased X-ray background scatter. Microfluidics also allows for rapid mixing of crystals with ligands shortly before exposure to the X-ray beam, enabling TR-SX (Monteiro *et al.*, 2020). Ultrasound waves can be used to levitate objects in air, without them touching any solid support. This has also been investigated as a sample delivery mechanism, suspending picolitre drops of crystal slurry in air (Priego-Capote and de Castro, 2006; Watanabe *et al.*, 2018; Morris *et al.*, 2019).

All the sample delivery methods described in §1.4 rely on a ready supply of thousands of protein crystals, which places very different requirements on crystallisation than conventional rotation crystallography. Where conventional methods usually demand a single large crystal, serial methods use microcrystals, commonly no bigger than 20 µm. The production of suitable crystals for serial experiments is described below in §1.5.

1.5 Sample production for serial crystallography

As described above, macromolecular crystallography has conventionally used a small number (less than 20) of large crystals mounted in loops (Cipriani *et al.*, 2006; Ohana *et al.*, 2004), rotated within the beam (Rodgers, 1994). Although large crystals (on the order of 100 µm) are easier to manipulate individually, and their optimisation is well-established, there are several advantages to using smaller crystals, or even microcrystals (20 µm), for crystallography experiments. One such advantage is smaller crystals better allowing light to penetrate to their centres, and for ligands to diffuse to their centres faster in time-resolved studies.

Crystalising a protein relies on manipulation of the crystallisation phase space, shown in Figure 1.7, by changing various parameters. The parameters include composition and concentration of precipitant solution, protein concentration, pH, and temperature (Asherie, 2004; Schlichting, 2015). When the concentration of a

protein exceeds the limit of its solubility, it begins to precipitate, by forming either amorphous precipitate or hopefully an ordered crystal. There are two stages to crystal formation; nucleation, where molecules spontaneously associate to start crystallising; and crystal growth, where crystals increase in size as more molecules are recruited at their surfaces (Chayen, 2004). To produce crystals, the composition of the crystallisation cocktail must fall within the nucleation zone in Figure 1.7.

The most common method of protein crystallisation is vapour diffusion, using either sitting or hanging drops. This involves mixing a small volume, on the order of single microlitres or less, with a similar volume of precipitant solution, then placing it in a sealed well with a much larger (500 µL) reservoir of precipitant solution. Hanging drops adhere to the underside of a coverslip, whereas sitting drops occupy a small well above the reservoir. Vapour diffusion refers to the way in which the solvent vapour, trapped in the well, will diffuse from the crystallisation drop to the reservoir solution to reach an equilibrium of precipitant concentration between the two (D'Arcy *et al.*, 2014). Reducing the volume of the crystallisation drop gradually increases both protein and precipitant concentration from undersaturation to a point where the nucleation zone is reached, indicated by the relevant line (dot dash) in Figure 1.7. In the nucleation zone, the protein

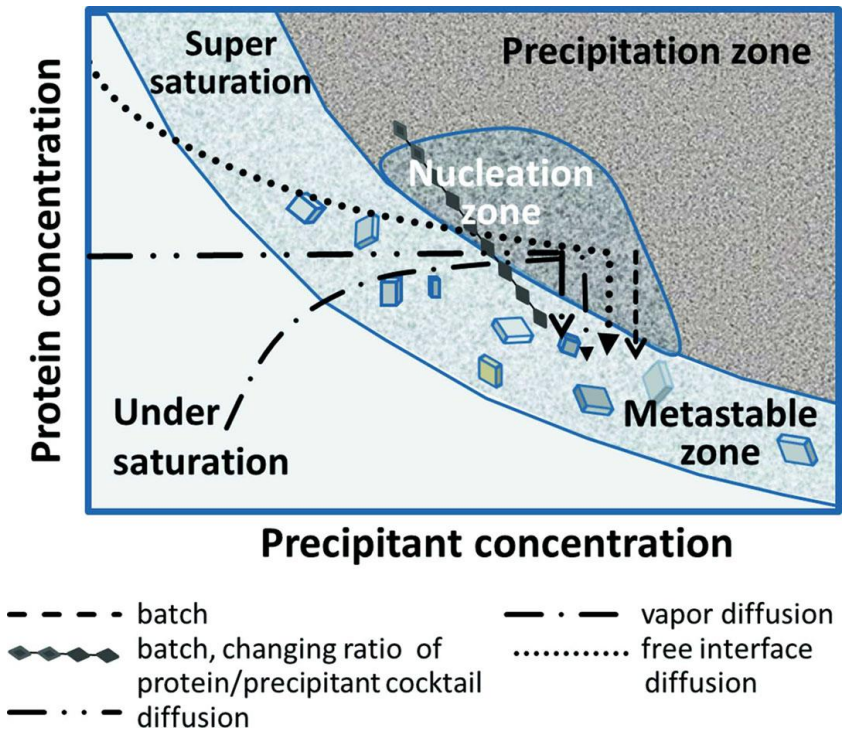

Figure 1.7: Diagram of the protein crystallisation phase space, showing effects of changing precipitant and protein concentration. Other factors that can be changed are temperature, pH, and additive concentrations. Reproduced from Schlichting (2015) with permission of the IUCr.

concentration exceeds its solubility limit and molecules spontaneously associate into crystals, which are more thermodynamically favourable than high concentration solutions. As crystal nuclei form, the protein concentration decreases, and the crystallisation drop enters the metastable zone. This is where crystals continue to grow by recruiting more protein molecules to their surfaces, but no new nuclei form. When aiming to grow a few large crystals, conditions should be chosen to spend little time in the nucleation zone, so fewer nuclei form. With fewer nuclei, more molecules can be recruited to each to form larger crystals.

Part of the challenge in growing crystals is controlling the amount of nucleation that occurs. To produce large crystals, the ideal starting conditions allow for the formation of few nuclei, so each crystal can grow larger. To produce microcrystals, more nucleation is desired, so many smaller crystals are produced. If conditions are known that produce crystals of any size, the crystals can be used as seeds. Seeds can be added to a crystallisation cocktail, inducing crystal formation without having to enter the nucleation zone of phase space (D'Arcy *et al.*, 2014; Kupitz *et al.*, 2014b; Stohrer *et al.*, 2021). Seeding can simplify finding ideal conditions and reduce the potential for excess nucleation. Any crystals of a protein can be used as seeds, including large crystals that have been crushed. Crystal crushing is either done using a glass rod with a hemispherical end, or by vortexing the crystals in a tube with a glass, steel, or PTFE bead. Seed crystals are often too small to identify and count, but the seed stock can be added to a crystallisation cocktail in various volumes to find the concentration of seeds which produces the desired amount of nucleation. High volumes of seed stock introduce more nuclei, promoting the growth of many microcrystals, whereas less seed stock means fewer, and potentially larger, crystals grow.

1.6 Serial experiment possibilities

1.6.1 Room temperature low dose crystallography

Conventional rotation crystallography usually takes place at cryogenic temperatures to slow damage to the crystal (de la Mora *et al.*, 2020; Owen *et al.*, 2006), as described in §1.2, although this is not necessary for serial crystallography where only one diffraction pattern is collected per crystal, and damage can only build up over the course of a single exposure (Owen *et al.*, 2017). Structures collected at room temperature are more likely

to represent the conformation of the protein at physiological temperatures, and therefore are of more physiological relevance than those at cryogenic temperatures (Fischer *et al.*, 2015; Sharma *et al.*, 2023; Ebrahim *et al.*, 2022). Room temperature data collection is also required for time-resolved crystallography, as at cryogenic temperatures proteins would not undergo conformational motion and reactions (Burnley *et al.*, 2012).

1.6.2 Damage-free structures at XFELs

As discussed in §1.2, X-ray exposure drives damage processes in protein crystals, and all protein structures solved at synchrotrons contain artefacts caused by beam-induced damage. The only way of solving structures completely free of X-ray damage is to collect data at XFELs, or with neutron diffraction (Blakeley *et al.*, 2008; Meilleur *et al.*, 2006) or NMR (Bax and Clore, 2019). Structures collected using XFELs have been known as zero-dose structures, as despite the doses being very high, the data is collected before any of the dose has been able to damage the crystals. Neutron diffraction for damage-free structures is limited by beamtime availability, the requirement for very large single crystals, and limits on unit cell dimensions, so is not discussed further here.

1.6.3 Serial femtosecond rotation crystallography (SF-ROX)

A technique exists for collecting rotation datasets at XFELs in a similar way to conventional crystallography with synchrotron sources. As intense XFEL radiation greatly damages or destroys the illuminated area of a crystal upon a single exposure, serial femtosecond rotation crystallography (SF-ROX) has been developed (Hirata *et al.*, 2014). Between collection of consecutive diffraction images, as well as being rotated, the crystal is also translated in the beam, exposing a fresh volume of the crystal each time. Collecting images at different known angles from the same crystal can aid data processing, and allows all reflections to be recorded systematically, not just by randomly sampling with stills. Taking advantage of the images collected at known angles depends on the data processing software. Currently, *DIALS* (Winter *et al.*, 2018) and *CrystFEL* (White *et al.*, 2016) do not have specific SF-ROX modules, so XDS is usually used for processing the data (Kabsch, 2010). Data collection at XFELs can also be applied to multiple crystals mounted on a micromesh pin. The distance by

which the crystal is translated is based on the zone which is damaged around the previously exposed volume. To maximise the data collected from one crystal, cryogenic conditions are used to prevent migration of radicals in the crystal, therefore slowing beam damage. Unfortunately, the cryocooling also prevents substrate migration, and therefore preventing SF-ROX from being used to study dynamics, except for states that can be captured by cryotrapping (Suga *et al.*, 2019). SF-ROX has been shown to produce damage-free structures of macromolecules (Suga *et al.*, 2015; Tosha *et al.*, 2017). This method would be suitable if conditions to grow microcrystals of a protein cannot be found, but large single crystals are available.

1.6.4 Time-resolved serial crystallography (TR-SX)

Because of the data collection time involved with obtaining diffraction images from a crystal at multiple angles, traditionally crystallography can only be done with macromolecules in static or very long-lived states, where the data collection time is shorter than the lifetime of the intermediate state being probed (Wulff *et al.*, 1997; Moffat, 1997). Alternatively, reversible reactions can be repeated several times on the same crystal, collecting a small amount of data each time. This has been used to investigate light-initiated carbon monoxide dissociation from myoglobin with a timepoint as short as 100 ps (Schotte *et al.*, 2004; Schotte *et al.*, 2003), made possible using Laue crystallography.

The Laue technique reduces the required X-ray exposure time per image by using a high flux pulse of polychromatic X-ray beam to illuminate the sample (Ren and Moffat, 1995; Moffat, 1997). As a polychromatic beam is used the Bragg condition is satisfied for many reflections simultaneously. Because of the high number of reflections on the same image, adjacent spots in each image lie very close together. Data integration becomes more complex because of potential spot overlap and the wide X-ray bandwidth used (Ren and Moffat, 1995). To avoid this spot overlap, having isomorphous crystals is important, as spots remain small with low crystal mosaicity. Another early example of time-resolved crystallography is glycogen phosphorylase *b*, with various experiments showing presence of enzyme-substrate complexes and slow conversion to a product-bound state over several hours (Hajdu *et al.*, 1987), such as GTP hydrolysis with Ha-Ras p21 (Schlichting *et al.*, 1990). By merging wedges collected from several crystals (or from a reversible reaction in a single crystal), the

required duration of continuous data collection is reduced. With the increased flux of synchrotron sources over the last 35 years, the same reaction could be studied today with much shorter exposures and at shorter timepoints (Schmidt, 2020; Geremia *et al.*, 2006; Schmidt, 2013).

The restriction imposed by minimum data collection time for time-resolved crystallography precludes its use to investigate protein dynamics of fast and irreversible reactions, other than what can be inferred from static structures. Serial crystallography circumvents this by collecting only a small amount of data from each crystal, which is done in a very short period, on the order of milliseconds at synchrotrons, and femtoseconds at XFELs (Chapman *et al.*, 2011; Stellato *et al.*, 2014). Because only one short exposure is collected from each crystal, the molecules within the crystal need not be in a steady state, but exposures can capture diffraction of an intermediate within a reaction (Pearson and Mehrabi, 2020). This is called time-resolved serial crystallography (TR-SX), and TR-SFX or TR-SSX specifically at XFELs and synchrotrons respectively.

Reactions can often take place within protein crystals because of their relatively high (compared to small molecule crystals) solvent content of approximately 50 %, and their somewhat retained conformational flexibility. The high solvent content allows for diffusion of substrates and products through the crystal via solvent channels, including into and out of active sites. The ability of substrates to access active sites within crystals is dependent on the way the crystals in question form, and where the solvent channels lie in relation to the active sites. Crystal packing may also limit certain conformational motions that are necessary for substrate binding and catalysis, resulting in either the crystallised protein not being catalytically active, or the crystals disintegrating upon reaction initiation (Hajdu *et al.*, 1987; Schmidt, 2020). Where reactions do take place inside crystals, the rate of reaction is greatly reduced compared to the protein in solution, in part affected by the conditions the crystals are grown in, as viscous precipitants reduce diffusion rates. Reduced rates of reaction are often advantageous for time-resolved crystallography experiments, as the prolonged lifetimes of intermediate states makes them easier to capture by X-ray diffraction.

1.6.5 Initiating reactions for time-resolved serial crystallography

Reaction initiation is a crucial step in any time-resolved biological experiment. Reactions need to be started at a controlled time, and for ensemble techniques such as crystallography, reaction must occur in the whole sample (whole crystal) at the same time (Grünbein *et al.*, 2020). Since ensemble techniques collect an average from all molecules in the sample, the molecules must be in the same state, or one of a few states, to provide clear information, just as a musical ensemble must play a piece in time with each other to produce a clear and pleasant sound. If the molecules within the crystal are not identical, such as flexible regions of a protein, or exhibiting different intermediate states in a chemical reaction, the resulting electron density map is blurred to represent this ensemble of states. This map can be harder or impossible to model a complete structure into, and the multiple conformations can make interpretation of the reaction mechanism harder (De Zitter *et al.*, 2022). For serial crystallography, optimal reaction initiation means concurrent triggering in the whole crystal and using the same delay time for all crystals. The precision required for reaction initiation depends on the timescale of the reaction being studied. If the timepoints are all relatively long-lived, then the permissible variability of reaction triggering times across the sample is greater.

When the required time precision in reaction initiation is lower, the obvious method of triggering is mixing of an aqueous reagent or ligand with the protein crystal by soaking, such as mounting a single crystal in a flow cell on a goniometer (Stoddard and Farber, 1995). The main disadvantage of mixing for time-resolved experiments is the time taken for the ligand to diffuse into the crystal and reach all protein molecules. Rates of diffusion are much lower in crystals than in solution, so the time to reach equilibrium will be much greater, especially for large crystals with long distances from the surface to the centre (Gouet *et al.*, 1996). As a result, the reaction will begin sooner in the unit cells closer to the surface. Incomplete mixing can result in mixed populations of molecules within the crystal, including partial occupancy of ligands, greatly complicating structure determination from the diffraction data, and not providing clear information on intermediates within the reaction (Schmidt, 2013; Beyerlein *et al.*, 2017). This limits time resolution to timepoints much slower than the diffusion time, ruling out study of many molecular interactions. For conventional single-crystal data collection, diffusion times on large crystals can be on the order of a minute, and has been used for studying

slow reactions, such as the phosphorylation reaction of glycogen phosphorylase *b* (Schmidt, 2013). With microcrystals, diffusion times are much less than for large single crystals, but still limit the resolution of time-resolved diffraction experiments to the order of single milliseconds (Olmos *et al.*, 2018; Ishigami *et al.*, 2019; Kupitz *et al.*, 2016; Mehrabi *et al.*, 2019). Mixing methods have been successfully employed in several TR-SX experiments (Schulz *et al.*, 2022), including by mixing crystals with a substrate shortly before injection via a jet (see §1.4.1), termed mix-and-inject serial crystallography (Stagno *et al.*, 2017; Dasgupta *et al.*, 2019; Olmos *et al.*, 2018). Another mixing technique uses a droplet injector to put picolitre drops of ligand solution onto individual wells of a fixed-target chip (see 1.4.3), each of which contains one crystal (Mehrabi *et al.*, 2020; Glerup, 2022). Because the amount of mother liquor surrounding the crystal is small, the ligand can diffuse into it and bind active sites rapidly (Shilova *et al.*, 2022).

As well as mixing, it is also possible to initiate reactions in crystals using light. Unlike diffusing substrates into crystals, light reaches the whole crystal in a negligible period, allowing synchronous reaction progression throughout the crystal (Grünbein *et al.*, 2020). This is a form of pump-probe experiment, as one pulse of UV or visible light initiates a reaction in the crystal (forms the pump), and another pulse (X-ray beam) interrogates (probes) the sample. Although light activation is almost instantaneous across the crystal, intensity of light can vary massively due to the high optical density of the crystal, even across moderately sized crystals (Monteiro *et al.*, 2021). This has consequences for investigating single-photon reactions, as even with the correct total number of photons, the molecules near the front face of the crystal will absorb far more photons than those further back, which may not absorb any (Grünbein *et al.*, 2020). If multiple photons are absorbed by a chromophore in a short space of time, it can trigger changes outside of the usual reaction pathway seen in nature, and therefore provide misleading information about protein function and dynamics. Absorbance effects also apply to delivery media in liquid jets and viscous extruders, which can also absorb light, as well as the round profile of the jet or extrusion acting as a lens to focus the light, leading to areas of higher and lower intensity (Weinert *et al.*, 2019; Monteiro *et al.*, 2021).

Some proteins and protein complexes are naturally photoactive, meaning they absorb light and undergo detectable changes in response. Examples include photoreceptors such as rhodopsin (Nango *et al.*, 2016; Standfuss, 2017), or photosystems from bacteria or plants (Kupitz *et al.*, 2014a). Light has been used to activate reactions in phytochromes, which photoconvert between two stable forms that absorb different wavelengths of light (Burgie *et al.*, 2020). In these phytochrome experiments at 150 K, only partial photoconversion was observed, so light activation and synchrotron data collection could be employed to capture various intermediate states. If similar partial interconversion occurs in other light-activated reactions, cryotrapping methods could be used combined with room temperature SFX to determine structures of other light-sensitive proteins.

For reactions that are not naturally photoactivated, use of photocages can be considered. These photolabile compounds can be soaked into crystals in the dark, and when lysed by a pulse of light, release a smaller chemical species within the crystals. Photocages can be synthesised to release a reaction substrate, which becomes available for interaction with the protein in a period of microseconds after receiving the pulse of light (Schlichting *et al.*, 1990; Schlichting *et al.*, 1989; Monteiro *et al.*, 2021). Caged substrate photolysis has been used in experiments with conventional macrocrystals, especially before the advent of modern serial techniques. Both single crystal (Sakaguchi *et al.*, 2016) and time-resolved serial (Tosha *et al.*, 2017) experiments have been performed using N,N′-bis-(carboxymethyl)-N,N′-dinitroso-p-phenylenediamine, a

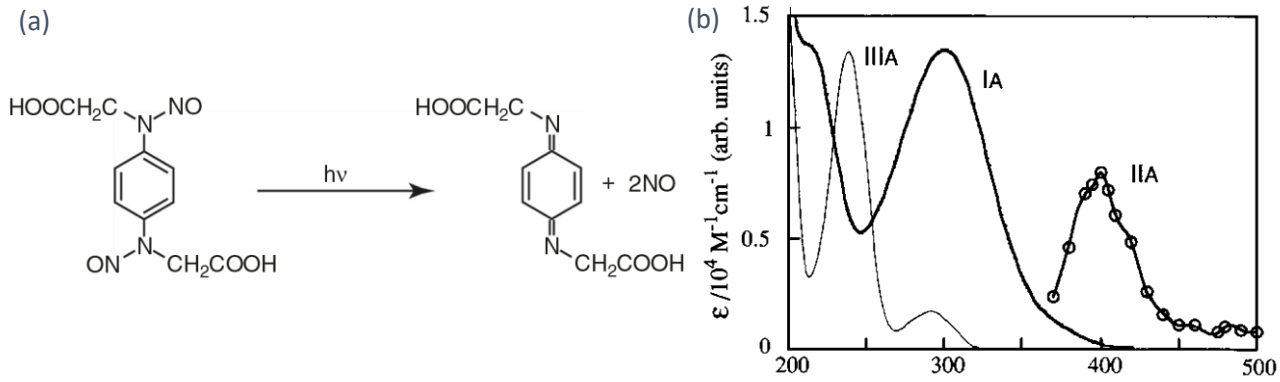

Figure 1.8: (a) Photolytic release of two NO radicals by the photocage N,N′-bis-(carboxymethyl)-N,N′-dinitroso-p-phenylenediamine. (b) Absorbance spectra from the photocage. I_A represents the absorbance before photolysis and features a 300 nm peak, II_A is an unstable intermediate after release of a single NO molecule, and III_A is the product shown in (a). (b) reproduced from Namiki *et al.* (1997) with permission, copyright American Chemical Society.

photocage that releases NO upon UV illumination centred at 300 nm (Namiki *et al.*, 1997), shown in Figure 1.8. In an SFX experiment performed by Tosha *et al.* (2017), the NO photocage was soaked into the protein crystals as a solution in mother liquor, and also dissolved as part of the hydroxyethyl cellulose (HEC) viscous medium (Sugahara *et al.*, 2017) used in the viscous extruder (Weierstall *et al.*, 2014). The experiment studies the haem-containing nitric oxide reductase P450nor from the fungus *Fusarium oxysporum*, and captured an intermediate with NO bound to the distal side of the haem, where it would undergo reduction to form an NHO^- or NHOH intermediate in the presence of NADH (Shoun *et al.*, 2012), and have suggested a mechanism for reaction specificity in the enzyme. This NO photocage differs from most other available photocages, in that the photolysis product of biological relevance, the NO, is significantly smaller than the *N*,*N*'-bis-(carboxymethyl)-*p*-phenylenediamine caging group. Most other caging groups are approximately equal to or are smaller than the biologically active molecule they release (Monteiro *et al.*, 2021).

Various caging groups have been developed that can be applied to different active molecules. Photocage design is non-trivial, and multiple factors affect the usefulness of a caged compound (Monteiro *et al.*, 2021). The ideal photocage is easily dissolved in the crystallisation mother liquor to a high concentration, usually either high concentrations of inorganic salts such as ammonium sulfate, or the organic solvent polyethylene glycol (Krauss *et al.*, 2011). Solubility is important as the cage must be present in at least the concentration of the protein in the crystals, to release the caged reaction substrate in a sufficient amount to initiate the reaction in a high proportion of unit cells. The uncaging reaction upon light initiation should be fast enough to provide all the substrate within a shorter period than turnover of the reaction in the crystals, to capture transient states within the reaction with high occupancy. The extinction coefficient of the cage is also important, as high absorption of the excitation wavelength prevents light from reaching the far side or centre of the crystal, so the reaction is only initiated in a proportion of unit cells.

Photocages are also available that bind to the protein to prevent reaction initiation, rather than sequestering a substrate. These often take the form of a protecting group bound to a side chain of a catalytically active amino acid in the active site, preventing turnover until the residue is uncaged by light (Deiters *et al.*, 2006; Wu

et al., 2004; Mangubat-Medina and Ball, 2021). Photocaged amino acids have been developed and applied to proteins for assays in live cells (McCray and Trentham, 1989), where light can easily penetrate the cell and uncage the enzyme in its natural environment, but have also been applied to TR-SX studies (Smith and Wilson, 2022). Artificial photosensitisation of proteins can be extended beyond protection of a specific active group with compounds such as photoswitches. The azobenzene photoswitch undergoes photoisomerization between cis and trans states upon illumination with green or UV light, and can be used to force conformational changes in protein secondary or tertiary structure (Beharry and Woolley, 2011; Sadovski *et al.*, 2009).

1.6.6 Investigating X-ray-induced damage to crystals

As well as trying to produce structures that are free from X-ray damage, the damage to crystals can be investigated. This -is important because the X-ray beam interacts with the crystals to produce photoelectrons, which react with the protein molecules and cause chemical changes (Carugo and Carugo, 2005; Garman, 2010), as described in §1.2. When damage is not considered, these changes can lead to misinterpretation of the protein structure, as the damaged state captured by crystallography is different from the natural state of the protein. Understanding damage processes in crystals can lead to better identification of these artefacts in the resulting structures and can lead to better data collection methods that reduce the amount of damage present in the structures.

An alternative form of pump-probe crystallography experiment uses X-rays as both the pump and the probe. Pump pulses initiate various X-ray-driven processes, and are followed by probe pulses used to collect

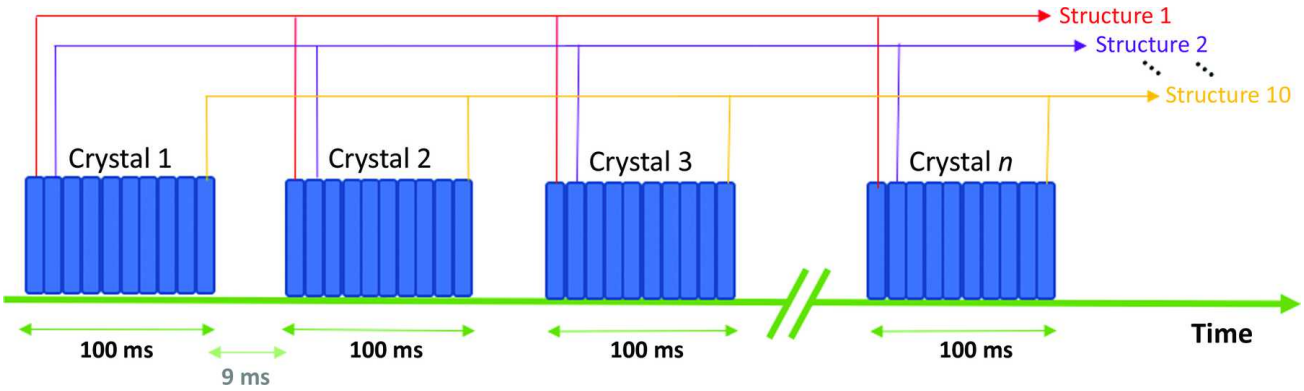

Figure 1.9: Representation of multiple serial structure (MSS) data collection from many crystals, possible with fixed target sample delivery. Consecutive 10 ms exposures are collected from one crystal at a time, then corresponding images of the same dose from each crystal are merged. Reproduced from Ebrahim *et al.* (2019a) with permission of the IUCr.

diffraction patterns (Ebrahim *et al.*, 2019b). Importantly, the exposures either have a time delay between them or are sequential (when using an almost continuous X-ray beam at a synchrotron), to allow the X-ray-induced changes to occur between exposures. In a technique called multiple serial structures from one crystal (MSOX), each X-ray exposure performs both pump and probe roles, enabling tens of structures to be captured sequentially as X-ray-driven changes progress in a crystal (Horrell *et al.*, 2016; Inoue *et al.*, 2016; Hough *et al.*, 2017). MSOX can produce a dose series of structures, where each consecutive structure has more beam damage than the last. From this the progression of damage processes within a protein crystal can be followed.

Multiple serial structures from many crystals (MSS) is an X-ray pump, X-ray probe technique that uses multiple crystals, and collects successive diffraction patterns from each crystal during a continuous X-ray exposure. Data is combined from the same respective exposure of each crystal, to produce a series of structures of crystals that have received increasing doses (Ebrahim *et al.*, 2019a), as shown in Figure 1.9. This technique can be used to track changes in crystals over increasing dose. Application of the technique to a copper nitrite reductase protein (crystallised in space group $P2_13$) has discovered a conversion between two polymorphs with unit cell lengths 1.4 Å apart by tracking changes with increasing dose up to 100 kGy. MSS has also been used to track changes at a radiation-sensitive haem site over increasing dose, to extrapolate a zero-dose structure (Ebrahim *et al.*, 2019b; Ebrahim *et al.*, 2019a).

1.7 Target proteins for serial crystallography

As described previously in §1.2, metalloproteins are particularly sensitive to site-specific X-ray damage, and metals are found in 40 % of enzymes with known structure (Andreini *et al.*, 2008). UV-Vis spectroscopy can be used to monitor photoreduction of certain metalloproteins (such as those with Fe^{3+} or Cu^{2+} ions), as recently demonstrated by Pfanzagl *et al.* (2020) and Lučić *et al.* (2020a). Therefore, metalloprotein studies benefit from the low dose and damage-free capabilities of serial crystallography, and make a good test case for new serial methods (Hough and Owen, 2021; Bowman *et al.*, 2016). A common cofactor found in metalloprotein is the haem group (Poulos, 2014), which consists of an iron ion inside a porphyrin ring, as shown in Figure 1.10. Haem proteins have a wide range of functions in ligand binding and catalysis, including facilitating redox reactions

(Lučić *et al.*, 2020b; Lučić *et al.*, 2021), and selective binding of small gas molecules such as nitric oxide and carbon monoxide (Lawson *et al.*, 2000; Manole *et al.*, 2015). The many roles of the haem cofactor are largely due to the iron at its centre. The octahedral coordination geometry of iron means it can be incorporated as part of the haem by the delocalised electrons of four pyrrole nitrogen atoms which lie in the plane of the porphyrin ring. This leaves the two axial coordination sites free for the iron to bind the protein or to provide a ligand binding function. Iron is a transition metal, meaning it has multiple oxidation states, most commonly ferrous (Fe^{2+}) and ferric (Fe^{3+}) iron, although ferryl (Fe^{4+}) iron is also found in haem proteins. These oxidation states mean that iron can act as an electron acceptor or donor within the protein, including in redox reactions.

Haem is bound to the protein by coordination with an axial residue as a ligand on the proximal side, most commonly histidine or cysteine, with polar interactions between its two propionic acid side groups and the protein providing additional stabilisation (Liu *et al.*, 2014). Some types of haem also involve covalent attachment, including haem C, which is covalently linked via two thioester bonds from its vinyl groups to two cysteine residues as part of a CXXCH motif (with X representing any amino acid) in the protein (Bowman and Bren, 2008). Covalent attachment of the haem aids stability of the protein, as demonstrated by mutagenesis of the relevant cysteine residues to convert a *c*-type to a *b*-type cytochrome (Tomlinson and Ferguson, 2000).

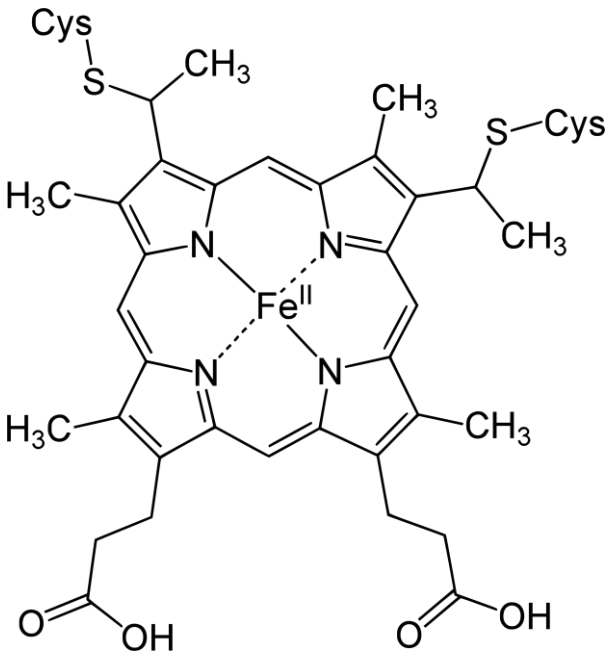

Figure 1.10: Haem C, shown covalently bound via two thioester links to the two cysteine residues in the CXXCH haem binding motif.

Less frequently, other covalent bonds form, such as the lysine crosslink in cytochromes P460 (Adams *et al.*, 2019; Bergmann and Hooper, 2003).

The nitrogen cycle involves several metalloproteins that bind nitrogen-containing species or catalyse reactions involving them. These proteins are of interest because of their roles in transferring nitrogen between the soil and the atmosphere. Nitrogenous compounds in the soil are necessary for crop growth and meeting global food demands (Cameron *et al.*, 2013; Stein, 2011). Nitric oxide, nitrogen dioxide, and nitrous oxide (NO, NO_2, and N_2O) are air pollutants released from combustion of fossil fuels (including road transport) and in industrial processes (Sillman, 1999). NO_2 has negative health effects via inflammation of the airways (Hamra *et al.*, 2015; Jiang *et al.*, 2019) additionally to being a precursor in ground-level ozone production (Hargreaves *et al.*, 1992). N_2O is a greenhouse gas that depletes the ozone layer in the upper atmosphere (Lasek and Lajnert, 2022; Winiwarter and Klimont, 2011). As well as being produced anthropogenically, the main source of N_2O is release from soil and bodies of water by conversion of nitrate and nitrite (NO_3^- and NO_2^-) by denitrifying bacteria (Butterbach-Bahl *et al.*, 2013). In addition to the wider effects of nitrogen species on the climate, NO is also a highly important biological signalling molecule, including in humans (Farah *et al.*, 2018; Heinrich *et al.*, 2013). Mammalian nitric oxide synthases contain haem groups (Minhas *et al.*, 2020) and have similarities to the proteins discussed below.

NO is a radical, meaning it has an unpaired valance electron, in this case in a 2π orbital, and can be considered to have 2.5 bonds between the nitrogen and oxygen atoms (Pacher *et al.*, 2007). When NO binds a transition metal, such as binding iron in a haem protein, it forms a metal nitrosyl complex. Bonding to transition metal cations occurs via the nitrogen due to its negative dipole. The bond angle between the metal and the NO can be either linear or bent, depending on the metal and its other ligands. For the Fe–NO complex formed when NO binds haem proteins, the angle is usually bent due to π backbonding (Lehnert *et al.*, 2010). A unique behaviour of NO binding to haem proteins not seen with CO and O_2 is a lowering of the binding strength of the other axial ligand (Cooper, 1999), commonly histidine. This can lead to dissociation of the other ligand, as described for the cytochrome *c*' from *Alcaligenes xylosixidans* below.

As described in §1.6.5, there exists a photocage that releases NO upon UV illumination (Namiki *et al.*, 1997), which has already been used successfully in a TR-SFX experiment (Tosha *et al.*, 2017). As well as the P450nor, cytochromes *c*′ (cyts *c*′) and dye decolourising peroxidase type B (DtpB) are also haem proteins that are known or expected to bind NO (Adams *et al.*, 2019; Lučić *et al.*, 2020b), which means time-resolved crystallography using the photocage is also possible with these proteins.

1.7.1 Cytochromes *c*′

Cytochromes are ubiquitous redox-active haem proteins that have functions in many different organisms and biological processes, including respiration, photosynthesis, and substrate oxidation (Adams *et al.*, 2019). Their roles include electron transfer, catalysis, and gas binding. The macrocyclic tetrapyrrole ring surrounding the iron has delocalised electrons that are important to the function of the proteins. Cytochromes are broadly classified depending on the form of the haem binding motif (Bowman and Bren, 2008; Allen *et al.*, 2005). The iron is coordinated by the four pyrrole nitrogen atoms in the porphyrin ring, and directly by the epsilon nitrogen atom of a histidine residue, a feature common to 80 % of *b*- and *c*-type haem proteins (Bowman and Bren, 2008; Li *et al.*, 2011). The side of the haem that coordinates histidine is called the proximal face (Hough and Andrew, 2015).

Cytochromes *c*′ are a group of cyts *c* that contain a 5-coordinate (5c) haem (pyrrole nitrogens and proximal histidine) in their resting states, implying a ligand binding function of their haem (Barbieri *et al.*, 2008; Tahirov *et al.*, 1996; Manole *et al.*, 2015; Cross *et al.*, 2001). Most cyts *c*′ discovered to date are primarily alpha-helical, with a four-helix bundle structure (Hough and Andrew, 2015). The *cycP* gene that encodes many cyts *c*′ is conserved across a range of methanotrophic, photosynthetic, and denitrifying bacteria, implying the proteins have a range of functions (Kekilli, 2015). One of the identified features of cyts *c*′ is the ability to bind diatomic gases such as CO and NO (Antonyuk *et al.*, 2011; Adams *et al.*, 2023), either on the vacant distal side of the haem, forming a 6-coordinate (6c) iron species (Barbieri *et al.*, 2008; Mayburd and Kassner, 2002; Manole *et al.*, 2015; Lawson *et al.*, 2000), or by displacing the histidine in the proximal position, leaving a 5c iron as the final ligand-bound state (Andrew *et al.*, 2016; Kekilli *et al.*, 2014), which takes place via a mechanism described

below. One such example of a protein that binds NO on the proximal side is AxCP, the cyt c' from *Alcaligenes xylosoxidans* (Kekilli *et al.*, 2017b; Pixton *et al.*, 2009).

Although the majority of cyts *c*' identified to date are four-helix bundles, a small number of examples with beta-sheets have been identified (Pearson *et al.*, 2007; Bergmann *et al.*, 2000; Liew *et al.*, 2019). This new class of cyts has been designated cyt *c*'-β after their secondary structure, as opposed to the more common alpha-helical type, now designated cyt *c*'-α. These cyts c'-β contain the same fold as cyts P460 (Adams *et al.*, 2019; Bergmann *et al.*, 2000), but the lysine residue, which in P460 is crosslinked to the haem and responsible for the characteristic Soret absorbance peak centred at 460 nm, is mutated, usually to phenylalanine. An example of this similarity is shown in *Methylococcus capsulatus* (Bath), with the structures of its cyt *c*'-β (McCP-β) and cyt P460 (McP460) being shown in Figure 1.11 (Arciero *et al.*, 1993; Adams *et al.*, 2019). Lys 78 in McP460 is

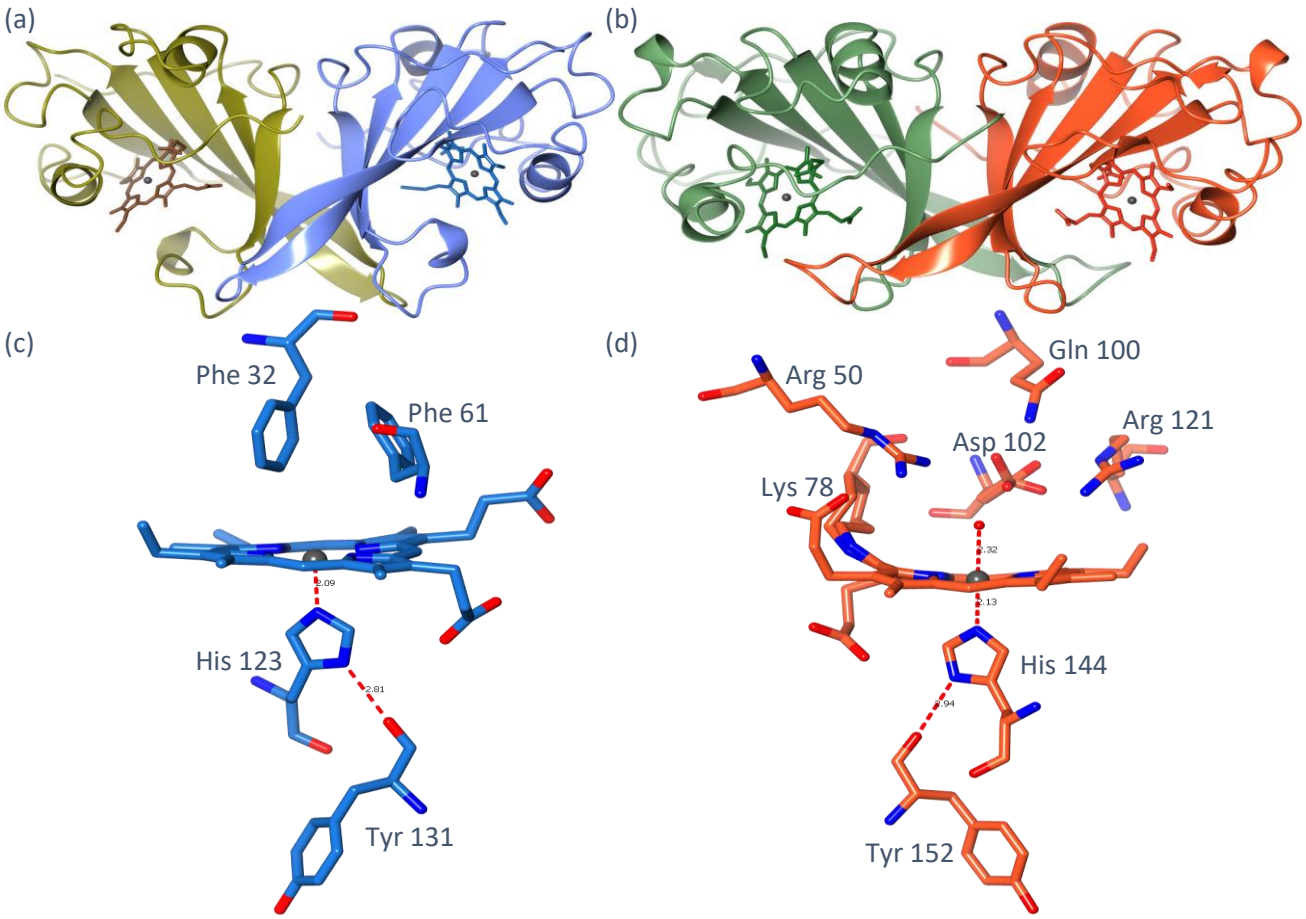

Figure 1.11: Comparison of cytochrome *c*'-β (a, c) and cytochrome P460 (b, d) from *Methylococcus capsulatus* (Bath). (a, b) The overall structure of the two homodimers, with the mostly β-sheet structure visible in both proteins. (c, d) The haem sites, which show similarity in a proximal His and Tyr (below haem), but have different distal sites, including a crosslink between the haem and Lys78 in cyt P460 (d). P460 has water bound in the distal pocket, unlike McCP-β. Data collected at 100 K (Adams *et al.*, 2019), PDB: 6hih (cyt c'-β), 6hiu (cyt P460).

mutated to Phe 61 in McCP-β, but both proteins share the tyrosine (Tyr 152 or 131) that coordinates the proximal histidine. Haem P460 is also found in the enzyme hydroxylamine oxidoreductase, which is also part of the nitrogen reduction pathway (Adams *et al.*, 2023), converting hydroxylamine to nitric oxide (Caranto and Lancaster, 2017), which is subsequently oxidised abiotically to nitrite ($NH_2OH \rightarrow NO \rightarrow NO_2^-$).

Both cyts *c*' given as examples above are suitable for TR-SX experiments because of the obvious differences in the active site structure upon NO binding. In McCP-β, the NO simply binds the vacant distal site of the 5c iron (Adams *et al.*, 2023; Russell *et al.*, 2013) to form a 6c complex, so the presence of the ligand is clearly visible in the electron density using an omit map ($F_{obs} - F_{calc}$ difference map) after molecular replacement. For some haem proteins, including McP460 (Figure 1.11d), the native state of the protein includes a water molecule bound in the haem site (6c iron), which gets displaced upon ligand binding (Kekilli, 2015; Kekilli *et al.*, 2017b; Hough *et al.*, 2011; Lawson *et al.*, 2000). This results in a much smaller change in electron density than the ligand binding an empty site, as the change is only from one non-hydrogen atom to two, rather than two non-hydrogen atoms binding an empty site, as is the case for McCP-β. A clear electron density difference aids development of TR-SX methods by simplifying interpretation of data processed in near real-time, to provide definitive feedback as to whether ligand binding has successfully been triggered in the crystals. This fast feedback cycle is vital to a successful beamtime, allowing adjustments to the experimental setup and reaction initiation parameters to be made within the same beamtime shift. See §3 for more detail on development of TR-SX methods.

Figure 1.12: Mechanism of binding of nitric oxide to *Alcaligenes xylosoxidans* cytochrome *c*' (AxCP). An NO molecule (red) binds the distal side of the haem, causing dissociation of the proximal histidine residue. A second NO molecule (green) can then bind to the proximal side of the haem, causing dissociation of the first NO molecule (red). Reproduced from Kekilli *et al.* (2017b).

In AxCP, where the final state of NO binding is a 5c distal species, binding occurs via multiple steps in a distal to proximal mechanism (Kekilli *et al.*, 2014), and only occurs with the protein in its reduced (ferrous, Fe^{2+}) state. This involves several intermediates that can be distinguished between by their electron density maps. The first step is binding of NO to the vacant distal site, similarly to McCP-β, shown by k_{6on} in the first step of Figure 1.12, with the NO indicated in red. This causes dissociation of the proximal His 120 residue from the iron (k_{-His}), which is displaced away from the haem and allows a second NO (green) to bind the proximal site of the haem (k_1) to form a 6c dinitrosyl intermediate. The first (distal) NO molecule then dissociates (k_2) to leave the final 5c proximal NO-bound product.

These intermediates provide target states to capture with TR-SX, that so far have not been captured by crystallography in the wild type protein. Unlike an enzyme that returns to a resting state after turnover, this protein with a separate endpoint can provide indication that the NO-binding has occurred, to differentiate between errors in reaction triggering, and incorrect timepoints missing the short-lived intermediate states. Several mutants have been produced that alter the NO-binding properties, including mutating the Leu 16 residue that constricts the distal pocket. L16A, L16V, and L16I mutants have a less constricted distal site, preventing the distal NO from causing His 120 dissociation, and halting the reaction at the 6c distal form (Lawson *et al.*, 2000). Crystal structures have been deposited for the final wild type proximal NO-bound state (Kekilli *et al.*, 2017b), and the distal 6c NO-bound state for several mutants, however no true intermediates, that being a state that is not an endpoint, have been captured with crystallography.

1.7.2 Dye-decolourising peroxidase type B (DtpB)

The dye-decolourising peroxidases (DyPs) are a family of *b*-type haem peroxidases with the ability to degrade various synthetic dyes (Kim and Shoda, 1999; Petrus *et al.*, 2016). Like most haem proteins, including cyts *c*', DyPs also feature a conserved proximal histidine residue coordinated to the haem iron. A member of the B-type subfamily of DyPs, named DtpB, has been identified from the Gram-positive soil bacterium *Streptomyces lividans* (Chater, 2016; Lučić *et al.*, 2020b), which unusually for a DyP, has a resting state distal pocket that is dry (free from water), visible in Figure 1.13b. Therefore, if DtpB binds NO, this will occur at the

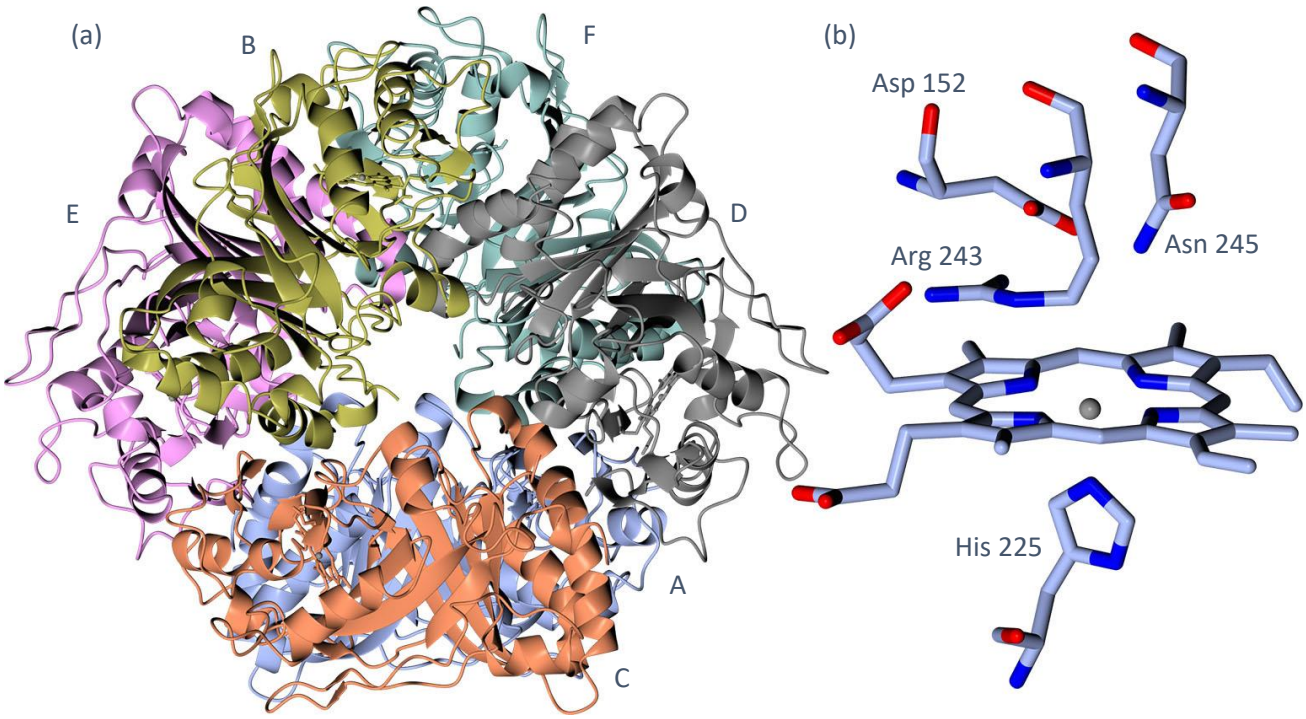

Figure 1.13: SFX structure of DtpB in the resting state, including (a) a ribbon view of all six chains, and (b) the haem site of chain A. PDB: 6yrj (Lučić *et al.*, 2020b).

vacant distal haem site, like in McCP-β, rather than via displacement of a water molecule. Whereas McCP-β and AxCP form homodimers (Adams *et al.*, 2019; Hough and Andrew, 2015), DtpB forms a homohexamer, shown in Figure 1.13a, meaning there are three times the number of haem sites present in the structure (Rožman, 2022). The additional haem sites provide more information on ligand binding from one structure. Together, these three proteins make a suitable collection of targets for use in development or TR-SX methods using a photocage that releases NO, providing variety in structure, dimerization state, and binding mechanism.

1.8 Objectives

Time-resolved serial crystallography is a powerful structural biology technique that provides insight into processes within protein crystals, rather than just the structure of the protein in a static condition. The field of TR-SX is relatively young, and there are still open questions regarding optimal sample delivery and methods of initiating reactions within crystals (Cheng, 2020; Schulz *et al.*, 2022; Brändén and Neutze, 2021). Here, development of a method for conducting TR-SX studies in fixed-target chips at beamline I24 and SACLA is described (§3), including initiating reactions with light via photolysis of a photocaged NO molecule (Namiki *et al.*, 1997). The effect of laser energy on the release of the photocaged ligand is explored via the resulting

ligand-bound structures collected, described in §4.2.2–4.2.3 and §5.2.1. McCP-β, AxCP, and DtpB, studied here, are all radiation-sensitive NO-binding haem proteins of bacterial origin, described in §4–5.

2. Methods

2.1 Expression and purification of McCP-β

Expression and purification of *Methylococcus capsulatus* cytochrome *c*'-β was performed as previously described (Adams *et al.*, 2019). The *cytS* gene encoding McCP-β (GenBank accession number AAD39218.1) was synthesised in a pBluescript SK(+) plasmid (Epoch Life sciences) with restriction sites EcoR1 and Xba1 (Michel *et al.*, 1995). The gene was excised and ligated into the expression plasmid pMMB503EH (Arslan *et al.*, 1998; Schulz *et al.*, 1999), with *E. coli* XL-1 Blue cells (Stratagene) used for cloning. The pMMB503EH plasmid with the *cytS* gene was transformed into *E. coli* BL21(DE3) cells which already contained the pEC86 plasmid (GenBank MH328007.1), shown in Figure 2.1, containing the *c*-type haem maturation genes *ccmABCDEFGH* (Barbieri *et al.*, 2008; Kekilli, 2015; Adams *et al.*, 2019). Unless otherwise stated, all culture growth used Miller's lysogeny broth (LB-Miller, Melford) at 37 °C, shaking at 180 rpm, or on LB agar (Fisher) plates at 37 °C.

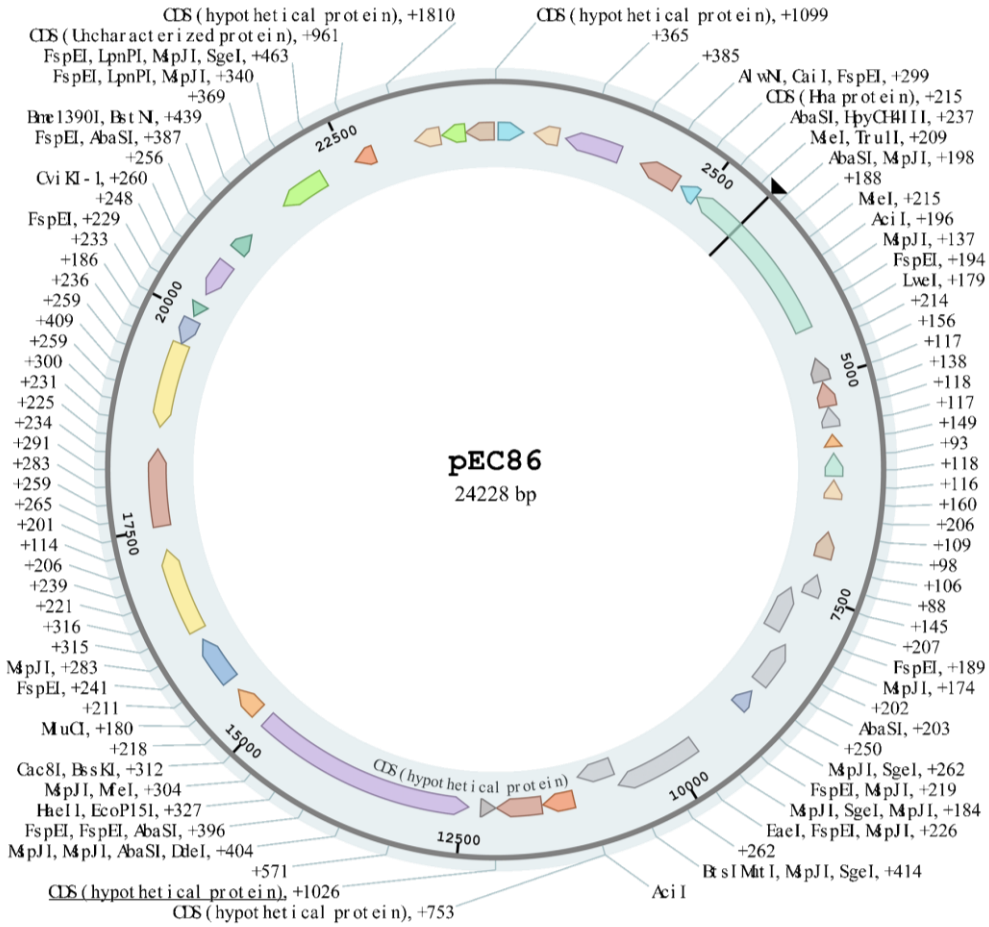

Figure 2.1: Map of the pEC86 plasmid containing the *ccmABCDEFGH* haem maturation genes, which are constitutively expressed with a *tet* promotor, not under oxygen control. These genes are essential for correct expression of cytochrome *c* proteins under aerobic conditions.

Transformation was performed using a standard heat shock protocol. 2 µL of plasmid at approximately 50 ng·µL^{-1} was gently mixed with 100 µL of competent cells and incubated on ice for 30 min. The cells were heat shocked for 45 s at 42 °C, then allowed to recover on ice for 2 min on ice. 800 µL of LB was added and the cells were incubated at 37 °C with shaking for 75 min. 100 µL was spread on an LB agar plate with appropriate antibiotics and incubated overnight at 37 °C.

Chemically competent cells containing the pEC86 plasmid were prepared by transformation as described above. One colony was picked and grown in a 10 mL overnight culture. 2 mL of this was subcultured into 50 mL of LB and incubated until an OD_{600} of 0.6–1.0 was reached. The culture was incubated on ice for 20 min, then centrifuged at 3700 × g for 10 min. All centrifugation was performed at 4 °C. Cells were resuspended in 10 mL of chilled 0.1 M $CaCl_2$ (Sigma) and incubated on ice for 30 min, then centrifuged at 1500 × g, for 10 min. Finally, the cells were resuspended in 2 mL of 0.1 M $CaCl_2$ with 15 % glycerol (Fisher), and 100 µL aliquots were flash frozen in liquid nitrogen. These competent cells were used for transformation with the pMMB503EH plasmid containing the *cytS* gene.

Transformants were selected for using streptomycin (Fisher) for the pMMB503EH plasmid and chloramphenicol (Fisher) for the pEC86 plasmid, at final concentrations of 50 µg·mL^{-1} and 20 µg·mL^{-1} respectively in all growth media. A colony was picked from an LB agar plate and cultured overnight in 10 mL of LB. This was subcultured into 100 mL of LB overnight under the same conditions. Terrific broth was prepared by dissolving 24 g·L^{-1} yeast extract (Oxoid), 12 g·L^{-1} tryptone (Oxoid), and 4 mL glycerol in 850 mL water and autoclaving. Once cool, 100 mL of sterile filtered solution containing 0.72 M K_2HPO_4 (Thermo) and 0.17 M KH_2PO_4 (Thermo) was added. 50 mL of starter culture was added to a final volume of 1 L of terrific broth and incubated at 37 °C, 180 rpm. Expression was induced with a final concentration of 119 µg·mL^{-1} Isopropyl-β-D-thiogalactoside (IPTG, Melford) at mid-log phase (OD_{600} = 0.4–0.6). After five hours, 30 µg·mL^{-1} (final concentration) ferriprotoporphyrin IX chloride (Acros Organics) was added as a haem precursor, followed by a metal ion mix (final concentrations 2 µM Ni^{2+}, 2 µM Co^{2+}, 10 µM Zn^{2+}, 10 µM Mn^{2+} and 50 µM Fe^{3+}) 24 h after induction (Adams *et al.*, 2019). The iron is necessary for the haem in the cytochrome, and the other

metals are required cofactors of the *c*-type haem maturation proteins, CcmABCDEFGH, expressed by the pEC86 plasmid.

After a further 96 h, cells were harvested by centrifugation at 3000 × g, for 20 min. Cells were resuspended in chilled 20 mM Tris-HCl pH 8 (Fisher), approximately 30 mL per litre of culture, and lysed by three passages through an Emulsiflex C5 cell disruptor (Avestin) at 80 MPa. After centrifugation at 26 500 × g for 30 min, the supernatant was dialysed for 24 h against H_2O, changing the water every 4 h or overnight. A 113 mL diethylaminoethyl-Sepharose (DEAE-Sepharose) Fast Flow anion exchange column (Cytiva) was equilibrated with 2 column volumes (CV) of 200 mM Tris-HCl pH 8, followed by 2 CV of 20 mM Tris-HCl pH 8. The sample was centrifuged again at 26 500 × g for 20 min, prior to loading onto the column at 2.5 $mL \cdot min^{-1}$. After loading, the column was washed with 2 CV H_2O and 2 CV 20 mM Tris-HCl pH 8, then protein was eluted in Tris with a 5–50 mM NaCl (Sigma) gradient over 4 CV. Binding of the protein to the column was very weak, but this was improved by dialysis into water prior to loading, and the subsequent wash with water. Purity was assessed by comparing the ratio of A_{280} to A_{400}, with a ratio of less than five indicating purity. Further purification was performed with a HiLoad Superdex 75 16/600 size exclusion chromatography column (Cytiva) equilibrated with 2 CV of 20 mM Tris-HCL pH 8.

2.2 Crystallisation of McCP-β

Purified protein was buffer exchanged into 100 mM HEPES pH 7.5 using a PD-10 desalting column (Cytiva). The protein was concentrated to 40 $mg \cdot mL^{-1}$ in a spin-concentrator (Merck Millipore), using the Soret band extinction coefficient of $\varepsilon_{400} \approx 70\ mM^{-1} \cdot cm^{-1}$ to calculate concentration (Adams *et al.*, 2019). Microcrystallisation was performed under batch conditions in microcentrifuge tubes. 250 µL of protein solution was mixed by pipette with an equal volume of crystallisation solution consisting of 68–70 % (v/v) polyethylene glycol 550, 100 mM MES pH 6.5, 10 mM $ZnSO_4$. Cubic crystals 30 µm across grew within 24 h incubated at 18 °C, shown in Figure 2.2a. Crystal sizes were measured using a Hirox HRX-01 microscope with HR-5000E lens. A 1 µL drop of microcrystal suspension was added to a glass slide, with a coverslip applied to flatten the drop into the plane of focus and reduce its rate of evaporation. To measure crystal concentration,

an improved Neubauer haemocytometer (Hawksley) was used. The crystals in four 1 × 1 x 0.1 mm squares were counted, and the number was multiplied by 2500 to get a concentration in crystals per millilitre. The ideal crystal concentration for loading a fixed target chip is 10^7 crystals per millilitre. When crystals grew at too low a concentration, they were sedimented using time stored upright, or by gentle centrifugation, and the required volume of mother liquor was removed to increase the concentration.

2.3 Expression and purification of AxCP

Expression and purification were performed as previously described (Yoshimura *et al.*, 1986; Barbieri *et al.*, 2008; Kekilli, 2015). The *cycP* gene coding for *Alcaligenes xylosoxidans* cytochrome *c*′ in the pET-26b(+) expression plasmid was transformed into *E. coli* BL21(DE3) cells that had also been transformed with the pEC86 plasmid, as with expression of McCP-β (§2.1). Transformants were selected for using final

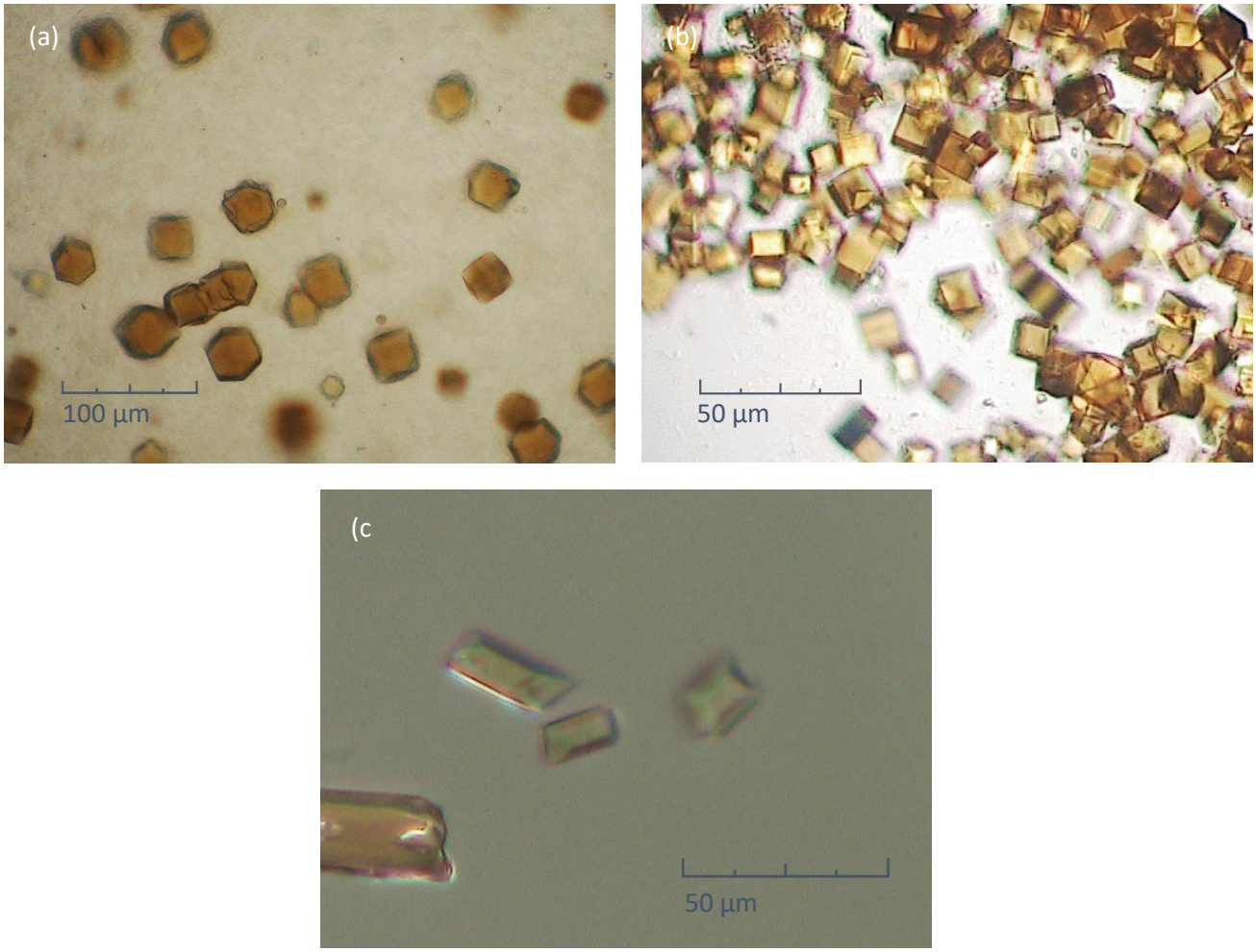

Figure 2.2: Microcrystals in mother liquor. (a) McCP-β, with approximate dimensions of 30 × 30 × 30 µm. (b) AxCP, with approximate dimensions of 15 × 15 × 15 µm. (c) DtpB, with approximate dimensions of 15 × 10 × 10 µm. Micrographs captured with Hirox HRX-01 microscope.

concentrations of 50 µg·mL^{-1} kanamycin (Fisher) and 20 µg·mL^{-1} chloramphenicol. Expression followed the same protocol as McCP-β, but upon induction, when OD_{600} reached 0.4–0.8, the temperature was lowered to 25 °C. Cells were harvested, lysed, and dialysed as with McCP-β, but 40 mM MES pH 6 was used for resuspension. The dialysed sample was loaded onto a CM-Sepharose Fast Flow cation exchange column (Sigma) equilibrated with 4 CV of 40 mM MES pH 6, then washed with 2 CV of water and 2 CV of MES buffer. The protein was eluted with a gradient of 25–150 mM NaCl in MES.

2.4 Crystallisation of AxCP

Purified protein was buffer exchanged into 50 mM HEPES pH 7.5 and concentrated to 40 mg·mL^{-1}, using the Soret band extinction coefficient of $\varepsilon_{400} \approx 80$ mM^{-1}·cm^{-1} (Barbieri *et al.*, 2008). Microcrystallisation was performed using batch conditions like those described for McCP-β above (see §2.2). Protein solution was pipette mixed with an equal volume of crystallisation solution consisting of 70 % (v/v) polyethylene glycol 550, 100 mM MES pH 6.5, 10 mM $ZnSO_4$. Crystals 15 µm across grew within 24 h incubated at 18 °C, see Figure 2.2b.

2.5 Expression, purification, and crystallisation of DtpB

DtpB microcrystals were provided by Dr Marina Rožman, produced as per the published protocol (Lučić *et al.*, 2020b; Rožman, 2022). The pET-28a(+) plasmid containing the DtpB gene was transformed into *E. coli* BL21(DE3) as described above and selected for using 20 µg·mL^{-1} kanamycin. Expression was performed with 1.4 L of LB in a 2 L conical flask, shaking at 180 rpm at 37 °C. When cell density reached OD_{600}=1.2, expression was induced with 0.5 mM IPTG, and cultures were supplemented with final concentrations of 250 µM 5-aminolevulinic acid as a haem precursor, and 100 µM iron citrate as an iron source. Carbon monoxide was bubbled through each culture for 30 s before sealing the flask, and expression continued at 30 °C, shaking at 100 rpm overnight (15 h). Cells were harvested by centrifugation at 10 000 x g for 10 min, then resuspended in 50 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole.

After cell lysis with an Emulsiflex C5 cell disruptor as described above, the lysate was clarified by centrifugation at 22 000 x g, for 30 min at 4 °C. A 5 mL HisTrap column (Cytiva) was equilibrated with lysis buffer, then the

clarified lysate was loaded onto the column. Protein was eluted with a linear gradient of 0–500 mM imidazole, and fractions to pool were chosen based on visual observation of the red colour of the haem. The pooled fractions were concentrated using a Vivaspin 5 kDa centrifugal protein concentrator, then further purified using a Sephadex 200 size exclusion column (Cytiva), which was equilibrated and eluted with 50 mM sodium acetate pH 5.0, 150 mM NaCl. Pure fractions were identified using SDS-PAGE, and purified protein was stored at −20 °C.

Microcrystallisation was performed by mixing equal 100 µL volumes of 6–10 mg·mL^{-1} protein solution and a precipitant solution consisting of 150 mM HEPES pH 7.5, 150 mM $MgCl_2$, 20 % PEG 4000. Crystals grew in a microcentrifuge tube after 24–48 h, shown in Figure 2.2c.

2.6 Fixed-target serial crystallography

Serial synchrotron crystallography (SSX) experiments were performed at beamline I24, Diamond Light Source (UK), using a fixed-target chip system developed in-house (Horrell *et al.*, 2021). The fixed-target chips are made from a 30 × 30 mm square silicon wafer (see §1.4.3). A chip has a grid of 50 µm deep square wells, which taper to 8–30 µm apertures to trap protein crystals, shown in Figure 2.3. The layout of the wells on the chip is shown in Figure 2.5a, with 64 blocks that each contain 400 wells in a 20 × 20 µm grid. For each experiment, a chip was chosen with an appropriate aperture size for the size of crystals in the batch, to trap the crystals while still allowing excess liquid to be removed. Prior to loading, each chip was glow discharged for 25 s in a 39 Pa vacuum with a current of 15 mA, to make the surface hydrophilic and increase surface wetting by adding a net negative charge to the surface. The glow discharging process encourages crystal slurry to flow across the chip, similar to grid preparation for cryo-electron microscopy (Drulyte *et al.*, 2018). This ensures even distribution

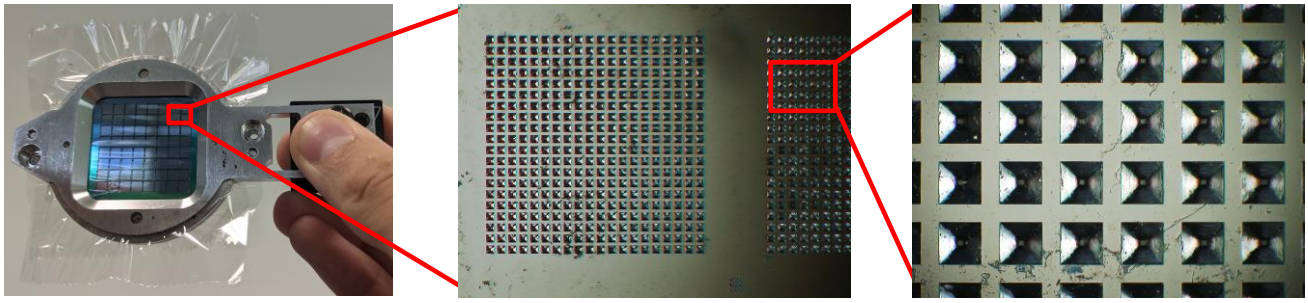

Figure 2.3: Detail of wells in an Oxford fixed target chip. The chip consists of 64 blocks, each containing a 20 × 20 grid of tapered wells, spaced 125 µm apart on centres.

of crystals, reducing multiple lattice events where two crystals share a well and have their diffraction patterns collected on the same image. Loading was performed in a humidity tent (Solo Containment) at >70 % relative humidity, to prevent crystals drying out during loading. A micropipette was used to apply 180 µL of crystal slurry to the chip and gently spread it across the surface to cover the whole grid of wells. A weak vacuum was then applied to the rear of the chip to draw the crystals into the tapered wells and remove excess mother liquor through the apertures into a waste flask. As well as the glow discharging, careful control of the crystal concentration and the volume of slurry applied are important for ensuring even crystal loading and reduction of multiple lattice events. The optimal crystal concentration is approximately 10^7 crystals·mL^{-1} (Horrell *et al.*, 2021). Although multiple lattices can be detected and separately indexed during processing with *DIALS*, the processing workflow for serial data does not yet have a way to separate images with multiple lattices and exclude them from further analysis when performing laser-initiated experiments (Winter *et al.*, 2018). Multiple lattice events usually account for approximately 10 % of indexed frames.

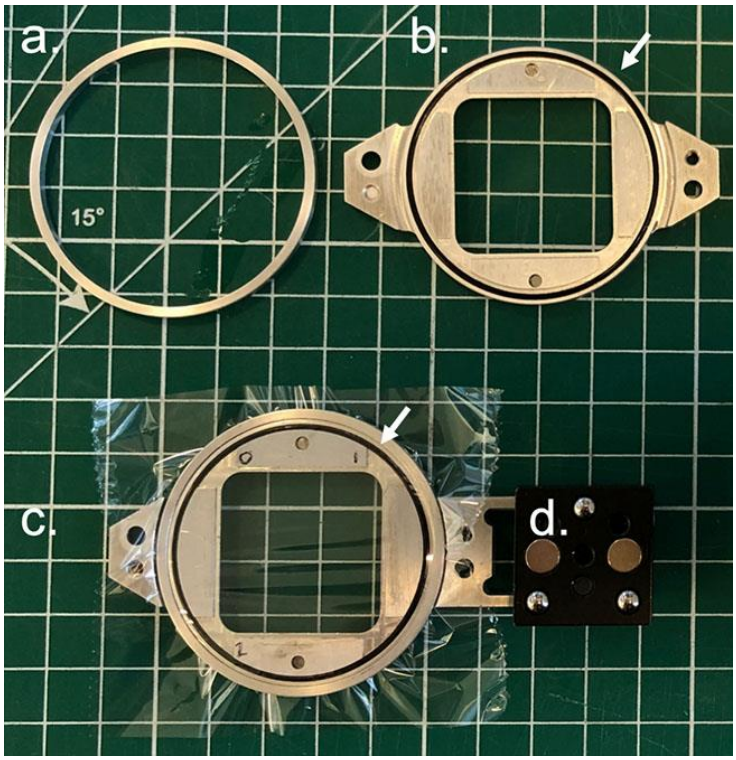

Figure 2.4: A fixed target holder. The left panel shows the parts of a chip holder on a 10 mm grid. A thin film on each side seals the chip to prevent dehydration. The ring (a) holds the film in place on each half of the holder (b), shown assembled in (c). The larger half of the holder has a kinematic mount (d), enabling easy mounting on the motion stage. Reproduced from (Horrell *et al.*, 2021).

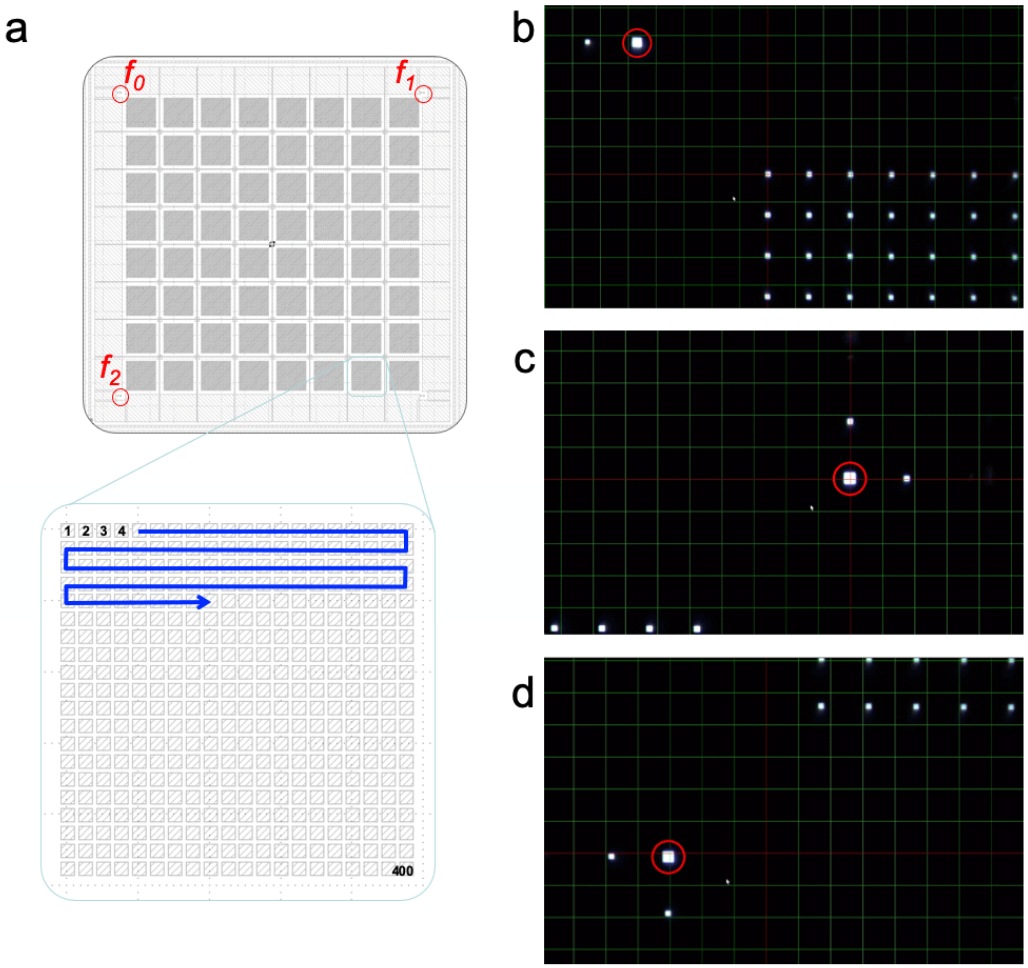

Figure 2.5: (a) The layout of wells on an Oxford silicon fixed target chip. The upper panel shows the 64 blocks of the chip and the locations of the three fiducials (red) used to align the chip. The lower panel shows the 20 × 20 grid of 400 wells in a single block, with the movement of the chip for data collection indicated with the blue arrow. (b–d) The fiducials ($f_{0–2}$) in the corners of the chip, as seen through the OAV (red circles added). The backlight shines through the apertures to make them visible in the OAV image. Reproduced from Horrell *et al.* (2021).

The loaded chip was placed into a machined aluminium holder, shown in Figure 2.3 and Figure 2.4, sealed on each side by a 6 µm layer of Mylar® film (biaxially oriented polyethylene terephthalate, BoPET, SPEX SamplePrep), to prevent evaporation and crystals drying prior to and during data collection. For time-resolved experiments, the front layer of Mylar® (facing the beam and on-axis viewer) was replaced by a 12 µm layer of EVAL™ (Kuraray), an alternative film made of ethylene-vinyl alcohol copolymer (EVOH). EVAL™ has greater UV transmission to allow pump laser penetration, but with the trade-off of slightly increased X-ray scattering, resulting in increased background on the detector. Thin films are compared in detail in §3.9.

The loaded chips in their holders were mounted onto the kinetic stage on the beamline using magnetic mounts (Thorlabs) for fast changeover. A graphical user interface (GUI) was used to align the wells on the chips, shown

in Figure 2.6. Translation in three dimensions allows visual positioning of fiducials in three corners of the chips with the beam using the on-axis viewer (OAV) microscope as shown in Figure 2.5b–d. These known positions are saved to form the basis for a coordinate system transform to allow each well to be accessed reliably, see the "Co-ordinate system setup" section, Figure 2.6b. A drop-down menu (c), specifies the type of chip to be used, from a list with preprogramed information on the layout of wells and fiducials. For all experiments

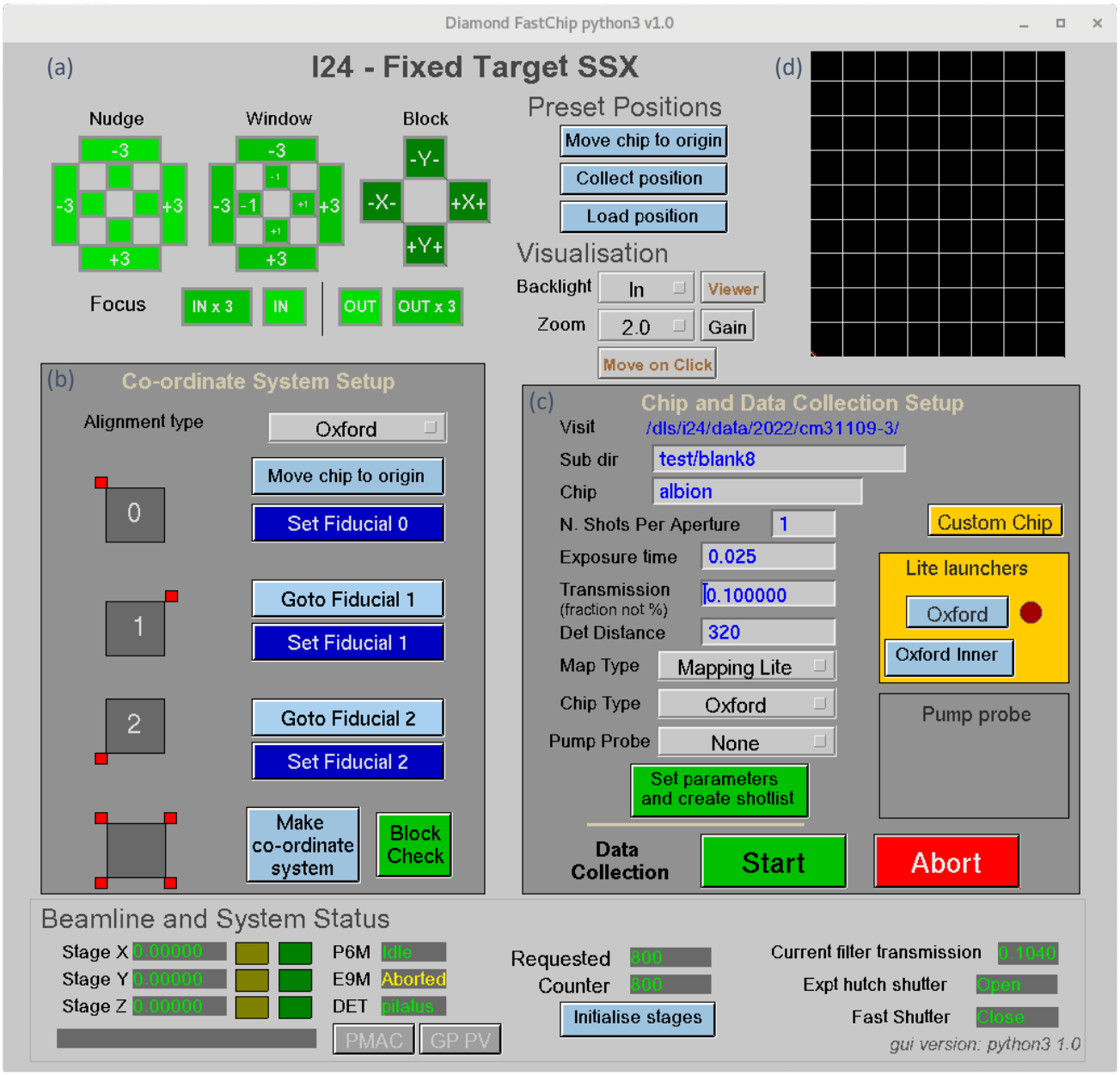

Figure 2.6: Graphical user interface (GUI) for controlling fixed target SSX data collection at beamline I24. The upper left area (a) is for moving the chip stage in three dimensions to align fiducials (shown in Figure 2.5) with the beam using the OAV. The co-ordinate system setup area (b) is used to save positions once aligned, and to create a co-ordinate system based off fiducials in the corners of each chip. The chip and data collection setup section (c) specifies the type of chip (and such the layout of wells), and as well as the sample details to be saved. During data collection, progress across the chip is displayed in (d).

described here, the Oxford chip was used (Figure 1.5). The GUI was also used for setting data collection parameters. SSX experiments at I24 used an energy of 12.4 keV and a DECTRIS PILATUS3 6M, or an energy of 20 keV and a DECTRIS EIGER2 X 9M CdTe. The beam size for most SSX experiments was 20 × 20 µm (full width half maximum, FWHM), with some being collected using an 8 × 8 µm beam.

The nitric oxide-bound structure of McCP-β was collected by pre-soaking crystals with an NO donor compound. This compound, 1-(hydroxy-NNO-azoxy)-L-proline, known as PROLI NONOate (Cayman Chemical), dissociates into proline and two NO molecules in a pH-dependent manner (Saavedra *et al.*, 1996). The donor was dissolved at 1 mg·mL^{-1} in 10 mM NaOH, then 30 µL of donor solution was soaked into 180 µL of crystal slurry for 10 min. The resulting drop in pH upon mixing with the larger volume of mother liquor released NO in solution to bind the haem sites within the crystals. The soaked crystals were then loaded into a chip in the usual manner described above.

2.7 Time-resolved serial crystallography

A more detailed explanation of time-resolved data collection, including method development at I24, is given in §3. Time-resolved NO binding experiments used the sodium salt of N,N'-bis-(carboxymethyl)-N,N'-dinitroso-p-phenylenediamine, a photocage that releases NO upon UV illumination (Namiki *et al.*, 1997). The photocage was synthesised by a collaborator, Dr Sinan Battah, following the protocol shown in Figure 2.7.

A solution was made to match the contents of the crystal mother liquor, containing the same final concentrations of precipitant, buffers, and additives, and the photocage was dissolved in it to 10 mg·mL^{-1}. This was to enable the photocage to be added to the crystals at a known concentration, without damaging the

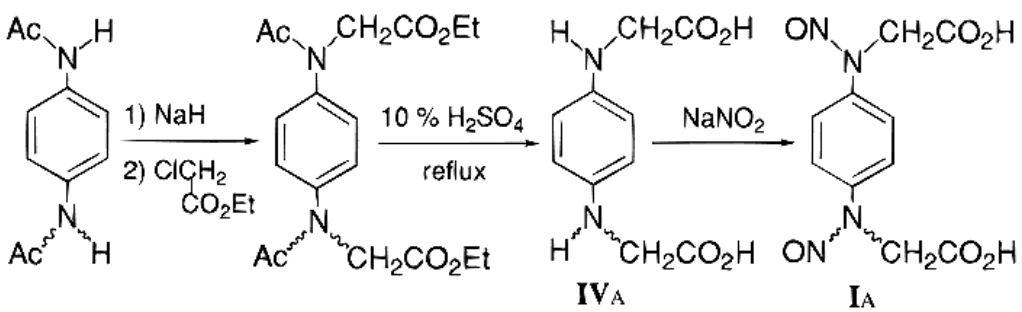

Figure 2.7: Synthesis of N,N'-bis-(carboxymethyl)-N,N'-dinitroso-p-phenylenediamine, reproduced from Namiki *et al.* (1997) with permission of the American Chemical Society.

crystals by otherwise changing the properties of the solution they were suspended in. Crystals were centrifuged briefly at 2 000 × g or sedimented with time, and the mother liquor supernatant was removed. Photocage solution was added to reach a final volume of 200 µL, and crystals were soaked for 2 min to allow diffusion of the cage into the crystals. Photocage preparation and subsequent chip loading were performed under red light, and samples were transferred in an opaque box to a darkened beamline for data collection, to prevent premature caged NO release.

Reaction initiation was performed using a 308 nm wavelength femtosecond UV laser aligned to the X-ray beam at the interaction region. Experiments aimed to optimise parameters for laser flash photolysis of the photocage within the protein crystals in the fixed-target chips. A more detailed description of these experiments and the method for laser initiation is available in §3. The energy used to initiate the reactions was controlled by either changing the pulse rate of the laser, or by adding neutral density (ND) filters to the beampath. The energy delivered from the laser was in the range 0.19–1600 µJ, which was imparted to the sample area in a 50–60 µm (FWHM) focal spot over 5–10 ms, with an energy density of 0.097–570 nJ·µm^{-2}. Laser power was verified prior to data collection using a laser power meter (ThorLabs) at the sample position, with a knife edge scan to determine focal spot size. The thermal laser power sensor works by absorbing the laser at a surface to generate heat, and sensing the temperature gradient between the surface and a reference surface of a heatsink by means of a series of thermopiles. A digital console reads the change in voltage and converts it into a power value in Watts.

A two-visit approach was used for data collection, where four rows of wells on the chip were illuminated with the pump laser, before returning to the same wells to collect diffraction data 1.24 s or 1.44 s later. This allows for a much shorter data collection time than if staying at one location for the whole delay period (Schulz *et al.*, 2018), as described in §3.10.

2.8 Viscous extruder serial crystallography

Hydroxyethyl cellulose (HEC) was used as a viscous medium for extrusion of crystal slurry into the X-ray beam (Sugahara *et al.*, 2017). The viscous medium was prepared a day in advance, involving dissolution of 25 % (w/v)

of HEC (VWR) in the relevant mother liquors in disposable syringes. The extra mother liquor was prepared by mixing precipitant solutions with the protein buffers in the same way as performing batch crystallisation, but without the protein. 40 µL of HEC medium was transferred to a 50 µL gas-tight syringe (Hamilton) via a syringe coupler (Conrad *et al.*, 2015). The crystal slurry was centrifuged briefly at 2000 × g to pellet the crystals, and excess mother liquor was removed. 10 µL of concentrated crystal slurry was transferred to a second 50 µL gas-tight syringe using a micropipette with gel-loading tip. The two gas-tight syringes were connected using a syringe coupler (SPT Labtech), as shown in Figure 2.7, and the contents were transferred back and forth until the crystals were evenly dispersed throughout the medium, approximately 100 times, with a final HEC concentration of 20 % (w/v). Figure 2.8 shows McCP-β crystals embedded in HEC during mixing, with a final volume of 50 µL. A syringe needle was used to transfer the prepared crystals into the sample reservoir of the extruder (Figure 2.9). The delivery capillary assembly was screwed into the front of the reservoir, which was then attached to the front of the plunger and connected to the gas line.

The extruder shown in Figure 2.10 was mounted vertically to the horizontal goniometer (designed for holding 96-well plates) on beamline I24, pointing downwards into the X-ray interaction region (Figure 2.11). The water side of the extruder was connected to a HPLC pump to control sample flow, and the sheath gas line was connected to a regulated supply of helium at 300 kPa to provide a flow rate sufficient to aid the vertical alignment of the extrusion (Weierstall *et al.*, 2014). The nozzle was aligned with the interaction region by

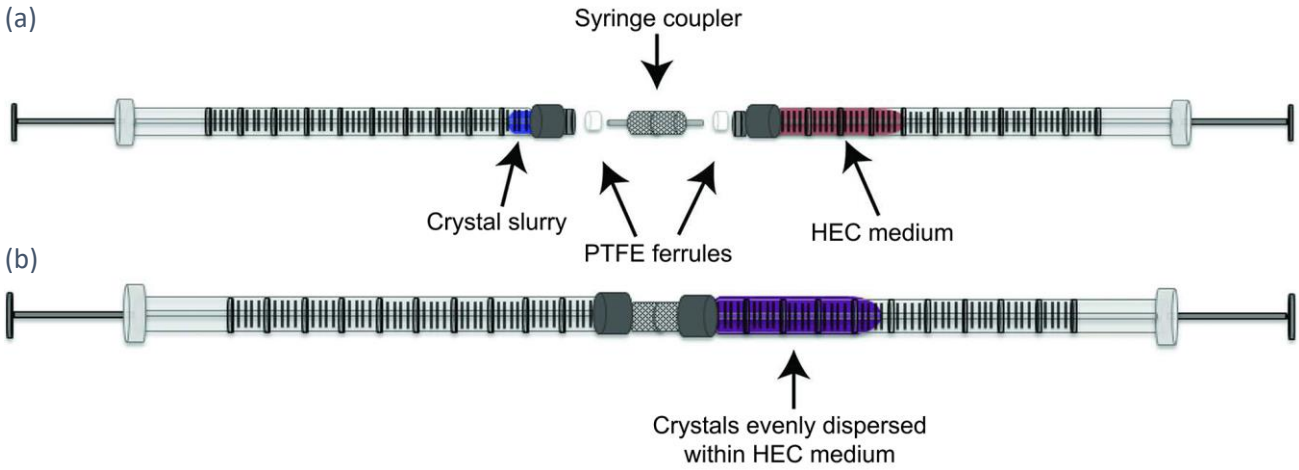

Figure 2.8: Mixing of protein crystals into hydroxyethyl cellulose (HEC) medium for viscous extruder crystallography. Two gas-tight glass syringes were connected using a syringe coupler (a), and concentrated crystal slurry was mixed with the viscous medium (b) by transferring back and forth, approximately 100 times, until homogeneity is reached. Modified from Conrad *et al.* (2015) with permission of the IUCr.

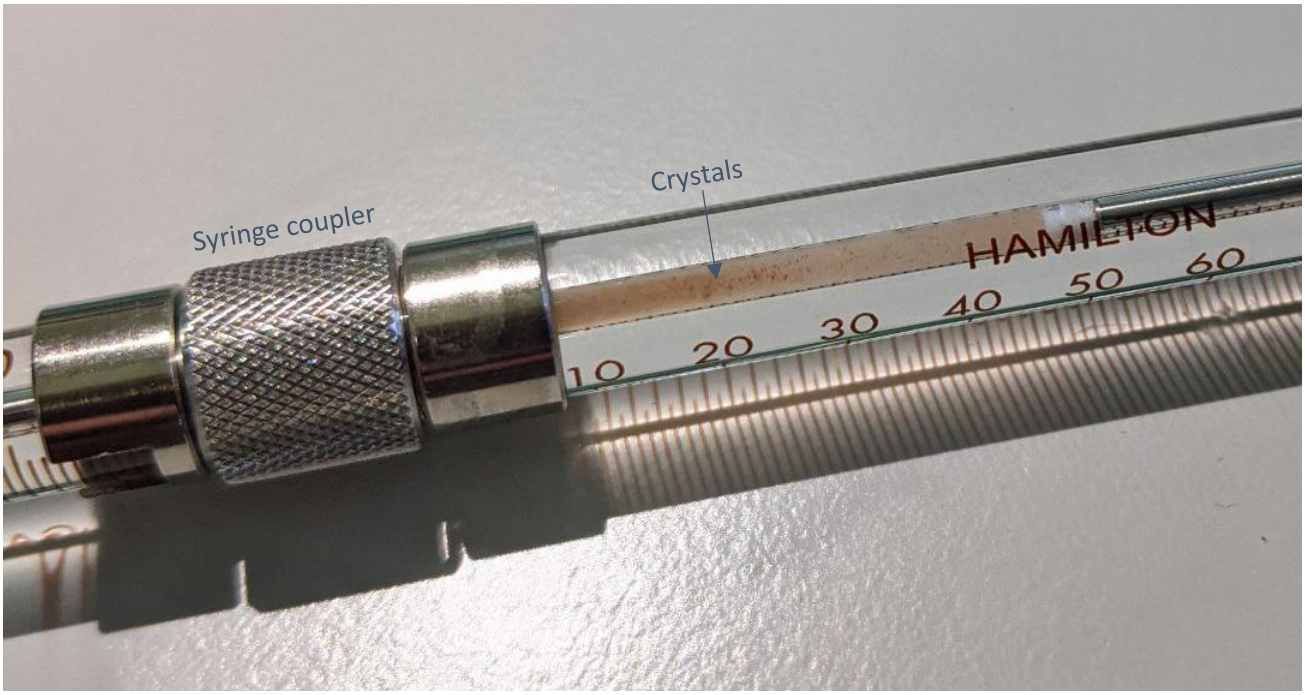

Figure 2.9: McCP-β crystals during mixing with hydroxyethyl cellulose medium in a gas-tight syringe (Hamilton). Crystals can be seen embedded in the medium in the syringe lumen. The volume markings on the syringe indicate microlitres.

translation of the horizontal goniometer, and a wide nozzle connected to a vacuum pump was similarly positioned underneath (blue in Figure 2.10) to catch the extruded medium after it had been exposed to the beam. The HPLC pump was set to run at a fixed flow rate of 13 µL·min^{-1}, which, due to the pressure multiplying effect of the piston in the extruder (see §1.4.2), corresponded to a viscous medium flow rate of 380 nL·min^{-1}. The flow rate was adjusted to optimise the hit rate (number of recorded images containing protein diffraction) and the sample consumption. At SACLA, the extruder was installed as part of the DAPHNIS (Diverse application platform for hard X-ray diffraction in SACLA) system on BL2 (Tono *et al.*, 2015; ono *et al.*, 2013).

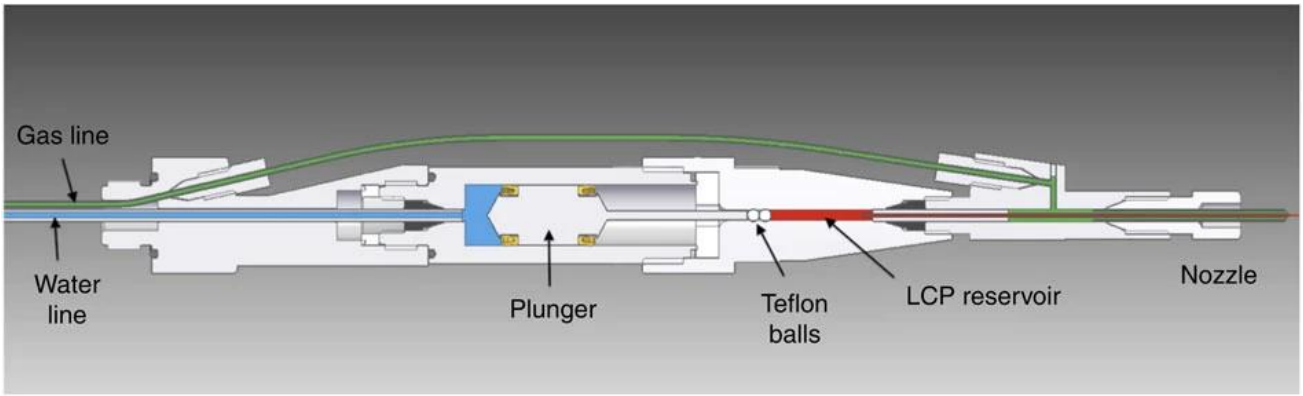

Figure 2.10: Diagram of the viscous extruder used at I24. An HPLC pump attached to the water line (blue) pushes the plunger, which pushes against the Teflon balls that seal the reservoir containing the viscous medium (red) with the crystals. The medium is extruded from the nozzle (right) and stabilised by the concentric flow of helium gas (green) surrounding it. The extruder is shown in a horizontal position, but during use points downwards into the X-ray beam. Reproduced from Weierstall *et al.* (2014).

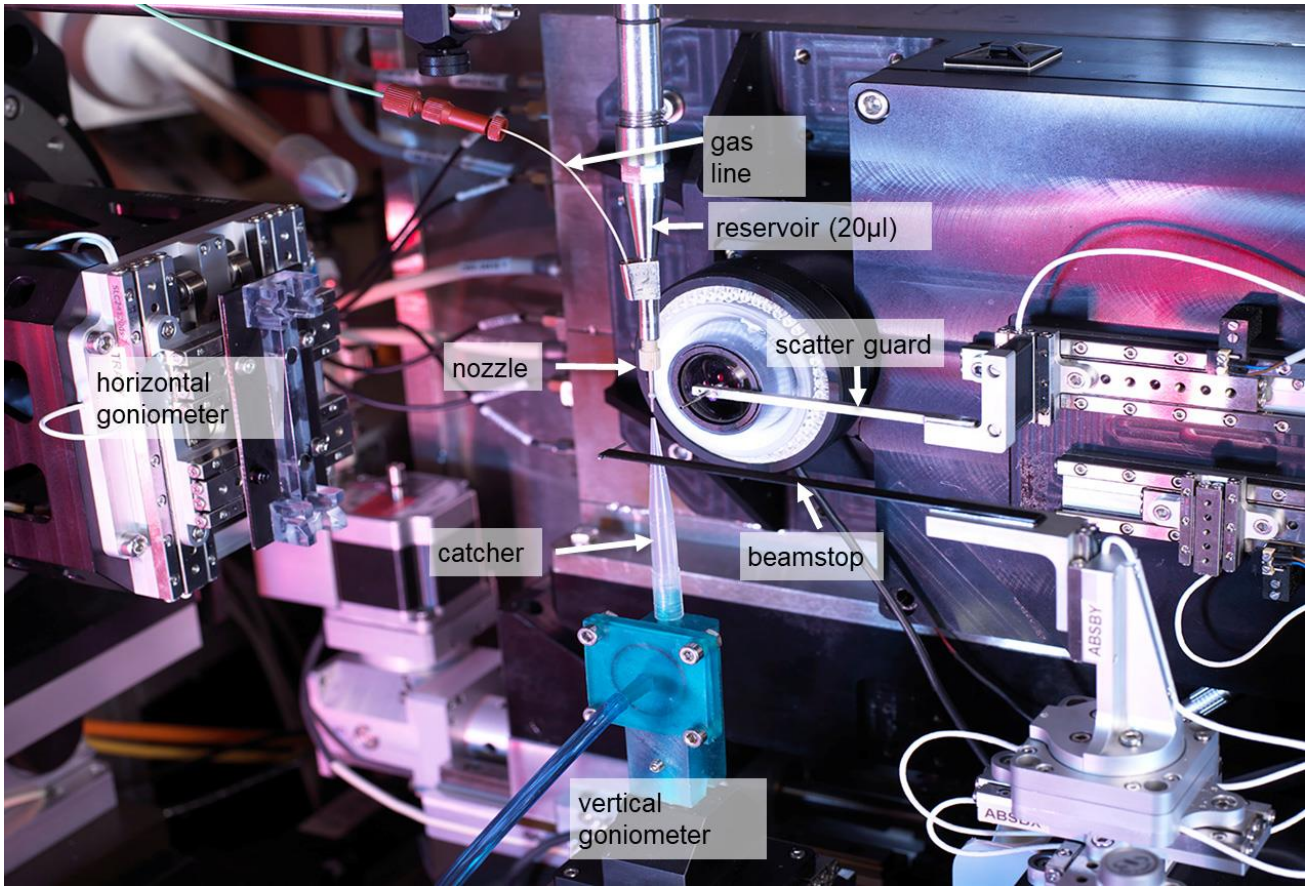

Figure 2.11: Viscous extruder mounted on beamline I24. The steel piston assembly extends off the top of the image and is fixed to the horizontal goniometer (left). The sample reservoir and nozzle are attached below, extending to just above the X-ray interaction region. The catcher attached to the blue support collects the extrusion below the interaction region.

At SACLA, the hit rate was monitored in close to real time using *Cheetah*, which performs spot-finding on a subset of images and determines hits based on a threshold for number of spots in an image (Nakane *et al.*, 2016; White *et al.*, 2016; Barty *et al.*, 2014). *Cheetah* runs with *CrystFEL* at SACLA as part of an automated image-analysis platform, with the *Cheetah Dispatcher* GUI for displaying results and initiating *Cheetah* runs. The analysed images were plotted on a graph of spot count versus image number, which corresponds approximately with experiment run time. At I24 at Diamond, a similar pipeline is available, which uses *DIALS* (Winter *et al.*, 2018; Waterman *et al.*, 2013; Winter *et al.*, 2022) to count spots and determine hits. It also indexes images to provide unit cell information in close to real time. Further details are available in §3.11, which describes the very similar real-time analysis used for fixed target data collection, and examples of plots are displayed in Figure 3.11.

2.9 Data processing

Diffraction data was processed via the command-line on the Diamond Light Source compute cluster using one of several software packages. With the exception of the SFX structure of McCP-β, all data was processed using algorithms from the *DIALS* (Diffraction Integration for Advanced Light Sources) toolkit (Winter *et al.*, 2022; Winter *et al.*, 2018). The time-resolved McCP-β data (§4.2.2–4.2.3) and the AxCP SFX structure (§4.2.4) were processed using the latest version of the dials.stills_process pipeline, which implements *DIALS* programs for importing data, spotfinding, indexing, integration, and geometry refinement (Winter *et al.*, 2018). This data was then merged using cctbx.prime (where prime stands for post-refinement and merging), part of the *cctbx* (Computational Crystallography Toolbox) package (Brewster *et al.*, 2016; Sauter *et al.*, 2013; Uervirojnangkoorn *et al.*, 2015), without any cycles of post-refinement. All DtpB data (§5.2), the resting state (§4.2.1) and dark control (§4.2.2) McCP-β SSX structures, and the AxCP SSX structure (§4.2.4), were processed using xia2.ssx (Winter, 2010; Winter *et al.*, 2013; Winter *et al.*, 2022), a pipeline under development for serial crystallography, which also implements *DIALS* programs. It includes provision for refining detector geometry using dials.refine, which was performed and optimised once per beamtime to produce a reference geometry which specifies the detector position relative to the sample. This data was merged with xia2.ssx_reduce, which uses the scaling and merging programs from *DIALS*, and can also be used to combine previously processed datasets from crystals under the same conditions. Resolution cut-offs were determined using $CC_{1/2}$=0.3 and multiplicity >10 for the highest resolution shell, while ensuring high completeness across the whole resolution range (>95 %). The McCP-β SFX structure (§4.2.1) was processed using *CrystFEL* (White *et al.*, 2012; White *et al.*, 2016), which was written specifically for processing SFX data from XFELs. Better results were obtained by omitting detector saturation correction when processing data from the SACLA MPCCD detector.

Due to the complexity of processing serial crystallography data with still images, the process of obtaining an MTZ file is more involved than for rotation data. For each dataset collected, processing parameters were optimised by testing on a subset of the data. Parameters that were changed included the minimum spot size, the spotfinding and indexing algorithms used, and the maximum number of crystal lattices to search for during indexing. Indexing was tested with and without providing a target space group and unit cell dimensions to

search for, and since the unit cell lengths vary by up to one ångstrom, the cell lengths provided and the length tolerance for filtering indexed crystals were also varied to obtain optimal results.

Molecular replacement was performed with *MOLREP* (Vagin and Teplyakov, 1997; Vagin and Teplyakov, 2010) within *CCP4i2* (Agirre *et al.*, 2023; Winn *et al.*, 2011). *REFMAC5* (Murshudov *et al.*, 2011; Vagin *et al.*, 2004) and *Coot* (Emsley *et al.*, 2010) were used for rounds of model building and refinement. Validation was performed with MolProbity (Williams *et al.*, 2018) and the JCSG Quality Control Check v3.1 (Elsliger *et al.*, 2010). All structural figures were generated with *CCP4 Molecular Graphics* (McNicholas *et al.*, 2011).

3. Development of light-initiated time-resolved SSX at I24

Time-resolved crystallography has existed as a technique for many years, but has been of limited use due to the difficulty of collecting a diffraction dataset from a crystal in a short enough period to clearly capture a single intermediate state in a reaction within a protein crystal (Brändén and Neutze, 2021). This has mostly restricted time-resolved crystallography to the slowest of enzymatic reactions with long-lived intermediates, besides some experiments with myoglobin which took advantage of the reversible photosensitive nature of CO dissociation from the haem, and which also used Laue crystallography (Wulff *et al.*, 1997; Moffat, 1998; Moffatt *et al.*, 1992). The advent of serial crystallography, however, has allowed much shorter total data collection time, and so lower absorbed dose, from any one crystal, as only a single diffraction pattern is collected from each crystal, as described in §1.3. As well as the development of data processing methods for still diffraction data (that is, collection of single images from crystals with no rotation), and sample delivery methods for rapidly bringing thousands of crystals into the beam (described in §1.4), time-resolved crystallography has benefitted from the advancements in X-ray sources. These advancements include much greater flux at modern insertion device beamlines and especially XFELs, and microfocus beams that can be focused onto crystals as small as 1 µm (Reich, 2013; Robinson, 2015; Owen *et al.*, 2014; Owen *et al.*, 2016). The increased photon flux of modern insertion device beamlines reduces exposure times and enables faster reactions to be captured with time-resolved serial crystallography. Exposure times of single digit milliseconds per image are used routinely, although total data collection time depends on the frame rate of the detector (Loeliger *et al.*, 2012). To get the most effective use from these high fluxes, modern hybrid photon-counting (HPC) detectors, such as the PILATUS or EIGER from DECTRIS (Donath *et al.*, 2023; Donath *et al.*, 2013; Kraft *et al.*, 2009), or integrating detectors as found at XFELs (Nango *et al.*, 2016; Leonarski *et al.*, 2018), are used.

The development of serial crystallography, including methods in sample delivery, and increased beamline flux and detector capabilities, provides the ideal environment for development of time-resolved serial crystallography from a niche method applicable to very few samples, into a highly capable technique that can be applied in a straightforward manner to many biological systems. Before TR-SX can become routine,

methods need developing with reliable but challenging samples as test cases, using iterative experiments to optimise beamline hardware and methodology. A major aspect of TR-SX that requires careful optimisation is reaction initiation, especially regarding using light for starting reactions within light-sensitive proteins (Standfuss, 2017; Suga *et al.*, 2019) or crystals containing photocages that protect substrates or catalytic groups (Monteiro *et al.*, 2021; Basa *et al.*, 2019; Wu *et al.*, 2004). Part of this optimisation includes quantifying the light exposure required to initiate a reaction without causing unwanted effects in the crystals due to multiphoton absorption or heating. This chapter presents the process and results of development of TR-SSX methods on beamline I24 at Diamond Light Source, applied to the fixed target sample delivery method for serial crystallography which was developed previously for use at the beamline (Horrell *et al.*, 2021; Sherrell *et al.*, 2015; Owen *et al.*, 2017). The chapter also details early experiments at I24 using UV laser illumination to cleave a nitric oxide photocage (Namiki *et al.*, 1997) that was soaked into crystals, and initiate binding of NO within two haem proteins. Although the structures collected here do not show true intermediate states which could not be obtained without time-resolved crystallography, the data was still collected in a time-resolved manner to develop the method. For simplicity, the term time-resolved is used throughout the thesis to describe the laser-initiated data collection and resulting data. Having the stable endpoint different from the starting structure makes it clear from the data that the reaction has been triggered in the crystals, without having to have the correct timepoint.

3.1 Sample delivery

The earliest time-resolved serial crystallography experiments have used liquid jets (GVDNs) or viscous extruders to deliver crystals into the beam (Nango *et al.*, 2016; Kern *et al.*, 2013; Kupitz *et al.*, 2014a; Aquila *et al.*, 2012; Tosha *et al.*, 2017), as described in §1.4.1–1.4.2, with reaction initiation by light activation (Togashi *et al.*, 2020; Kupitz *et al.*, 2016) or upstream mixing (Calvey *et al.*, 2016; Schmidt, 2013). Although effective for the study of short timepoints, no longer than tens of milliseconds, they are unable to access longer timepoints, as the crystals are passing transiently through the interaction region. This is where fixed targets (described in general in §1.4.3) are advantageous, as crystals are held stably on a support which can be held in the interaction region indefinitely (Horrell *et al.*, 2021; Owen *et al.*, 2017). Therefore there is no upper limit on the

timepoint that can be captured after reaction initiation with light or application of a reaction substrate by mixing, but short timepoints can still be captured equally well. For that reason, the TR-SSX methodology developed here has built on the silicon fixed-target chip system that is already in routine use for SSX at I24 (Horrell *et al.*, 2021).

3.2 Light activation

An alternative to substrate mixing for initiating a biological reaction is the use of light, which travels fast enough to be almost instantaneous, taking only about 40 fs for a beam to pass through a 10 µm protein microcrystal (Giacomo, 1984). Many time-resolved experiments make use of photocaging, where light induces cleavage of a protective group from either the pre-soaked substrate (Monteiro *et al.*, 2021), or from the protein itself (Deiters *et al.*, 2006; Wu *et al.*, 2004; Lee *et al.*, 2009), as described in §1.6.4. For studies of proteins that interact with nitric oxide, a form of caged NO is available, *N*,*N*′-bis-(carboxymethyl)-*N*,*N*′-dinitroso-*p*-phenylenediamine, with a molecular mass of 326.17 g·mol^{-1}, described in §1.6.5. This photocage compound has a strong UV absorbance peak at 300 nm, and upon illumination releases two molecules of NO in a reaction with a k_d of 2.96 × 10^4 s^{-1} at 23 °C (Namiki *et al.*, 1997). The photocage has been successfully used in a TR-SFX experiment with the haem NO reductase from the fungus *Fusarium oxysporum* (Tosha *et al.*, 2017).

3.3 Photocage synthesis

The previously published experiments with the NO photocage (Namiki *et al.*, 1997) that were performed at SACLA (Tosha *et al.*, 2017), mentioned above and in §1.6.5, used cage that was synthesised commercially (Dojindo Laboratories). From commercial suppliers, the chemical is available only in very small quantities (5 mg), costing several hundred pounds per milligram. This is unsustainably expensive for time-resolved experiments where approximately 2 mg is used in preparing each fixed target chip (see §2.6 for details on sample preparation). Therefore, the assistance of a collaborator was used in synthesising the cage in larger (400 mg) quantities which are sufficient for many time-resolved experiments. This required only common reagents that are available at reasonable cost, following the previously published protocol (Cabail *et al.*, 2002; Namiki *et al.*, 1997), shown in the scheme in Figure 2.7. As the synthesised photocage was untested, this was

an additional unknown in the first time-resolved experiments at I24. Therefore, as well as a test of the laser configuration and data collection methodology, the first I24 experiments sought to test the first batch of synthesised photocage. To this end, NO-bound structures with 308 nm laser activation have been collected (§3.6, §4.2.2), as well as a dark structure to show that the cage remains intact in the crystals and does not release NO without light exposure.

3.4 Laser configuration

As important as having a light-sensitive sample, either naturally or via photocaging, is being able to illuminate the sample on the beamline to initiate the reaction. For fixed-target SSX on beamline I24, described in detail in §1.4.3 and §2.6, that means directing laser light into an individual aperture of a silicon chip (shown in) to reach the crystal held within. As on the Oxford chip design used at I24, protein crystals are held in wells with a spacing of 125 µm between centres in X and Y dimensions (see Figure 1.5), care must be taken to illuminate only one well at a time, otherwise reactions could accidentally be triggered in neighbouring crystals at incorrect timepoints. This means the laser must be directed onto the sample in a spot of <125 µm and be aligned to centre accurately on the sample position.

For light activation of reactions in crystals in time-resolved experiments at beamline I24, Diamond Light Source, a tuneable femtosecond laser system (PHAROS, Light Conversion), called PORTO (PORTable pump-prObe), is available, and is shared between several beamlines. The laser has a maximum repetition rate of 50 kHz, with a peak power of 20 W at 1030 nm. The wavelength is tuneable in the range 210–2600 nm, covering a significant proportion of the UV region, and the entire visible range and near IR regions. These capabilities are ample to trigger a variety of reactions in protein crystals, including activation of light-sensitive proteins, lysis of photocages, and IR temperature jump experiments.

To get the laser light to the sample, a series of mirrors and a lens are used to direct the light from the output of the laser on its portable optical table. Movement of three mirrors is performed by kinematic mirror mounts (Thorlabs). The mirrors allow the light to approach the X-ray beampath from the side, as shown in cyan in Figure 3.1. The planoconvex lens is mounted on a motorised two axis stage (SmarAct) and focuses the light to

converge toward a 60 µm (FWHM) spot. A final smaller mirror is mounted on the arm of the upstream X-ray scatter guard, which sits between the lens for the on-axis viewer (OAV) and the sample position. The X-ray beam reaches the sample position via a drilled hole in the mirror and lenses within the OAV.

The mirror on the scatter guard arm directs the laser light to the sample at an approximately 15° angle off-axis from the X-rays, due to the mirror sitting to the side of the X-ray beam. The laser power is measured at the

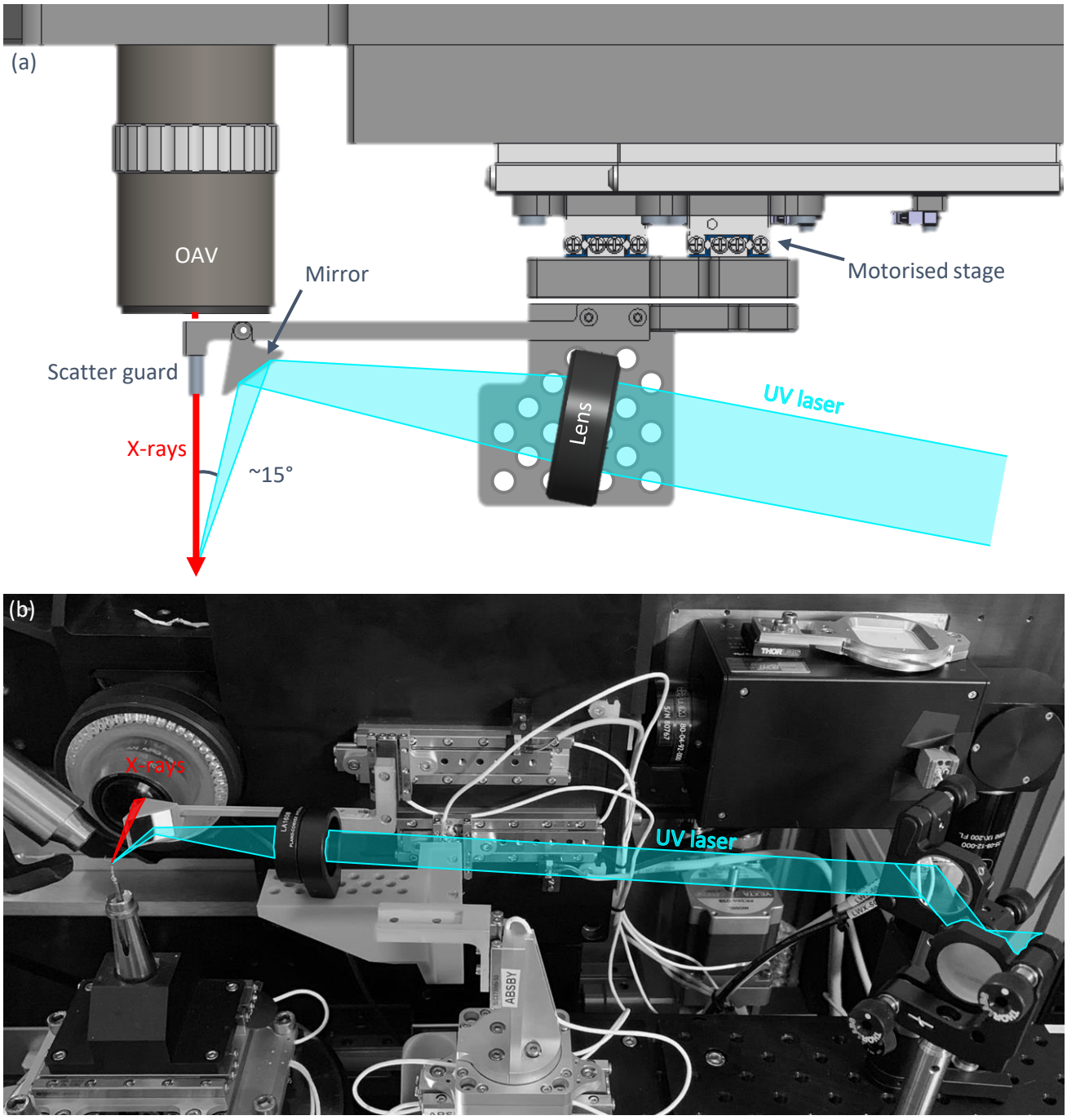

Figure 3.1: The I24 endstation, showing laser (cyan) alignment to the sample position via two circular mirrors, a focusing lens, and a final square mirror mounted on the scatter guard support arm. A top-down CAD drawing (a) and an annotated photograph (b) are shown, not to scale. The X-rays (red) reach the sample position via a drilled hole in the OAV and scatter guard. Note the approximately 15° offset between the X-ray and laser

beginning of a shift with a laser power meter (Thorlabs) mounted at the sample position, with a knife edge scan in two dimensions to determine focal spot size. The PORTO laser is integrated into the beamline personnel safety system to ensure that in normal user operation the output is disabled while people have access to the experimental hutch. A separate alignment mode allows people in the hutch with the laser on if the door is closed. The 15° offset of the laser from the X-rays as it hits the sample, rather than the two beams being coaxial, results in a parallax effect where laser alignment is sensitive to the position of the sample along the beam axis (Z axis). The position of the X-ray beam at the sample position is determined using a scintillator crystal (yttrium aluminium garnet) at the sample position. The laser alignment is subsequently adjusted to bring the focal spot onto this position using the crosshair in the OAV.

3.5 Laser energy calculations

The amount of energy used to initiate photolysis in the NO photocage, as determined by the power meter (see §2.7), was measured in microjoules (10^{-6} J). These laser energies refer to the total energy reaching the sample position during the exposure, and are not to be confused with the energy of a photon at 308 nm wavelength, which is 6.45×10^{-19} J, or 4.03 eV. The amount of laser energy measured at the sample position is not the same as the energy that reaches the crystals, nor that which is absorbed by the photocage. To have meaningful information on whether the laser energy used is providing single photon absorption by the photocage (Grünbein *et al.*, 2020), as described in §1.6.5, the energy delivered to the sample position requires converting into numbers of photons absorbed per photocage molecule.

For these calculations, a typical McCP-β crystal of 30 × 30 × 30 µm will be assumed, sitting face on in a 50 µm (FWHM) laser focal spot. With the Gaussian beam profile, approximately 58 % of the total beam energy is contained within the FHWM area in two dimensions (equal to 0.76^2). The 30 µm crystal occupies 46 % of the area of the laser focal spot within the FWHM. Having the laser focal spot slightly larger than the crystal is required to allow for tolerances in alignment and crystal positioning. Because the position of the crystal relative to the laser beam centre cannot be precisely controlled and can vary across the up to 25 600 crystals on a chip, the intensity of the beam within the FHWM area will be considered flat. Therefore, 26 % of the total

energy of the laser beam hits the crystal, although a significant proportion will pass through the sample without being absorbed.

To calculate the absorbance of the crystal (A), the Beer-Lambert law, $A = \varepsilon c l$, can be used. The concentration of the photocage in mM (c), the length of the light path in cm (l), and the extinction coefficient for the photocage in $mM^{-1} \cdot cm^{-1}$ (ε) are required. The extinction coefficient for the cage has been determined at 300 nm, with a value of 13.5 $mM^{-1} \cdot cm^{-1}$ (Namiki *et al.*, 1997), and should be similar with the 308 nm laser excitation used here. The photocage concentration for all experiments was 10 $mg \cdot mL^{-1}$ (31 mM). As the solvent content of the McCP-β unit cell is 63.6 %, the final photocage concentration within the crystal was 19 mM. With this photocage concentration, the extinction coefficient, and the crystal depth of 30 µm, an absorbance of 0.79 is obtained. As $A = -\log_{10} T$, this means that 16 % of the incident laser photons are transmitted through the crystal, and 84 % of the photons are absorbed.

From the proportion of photons hitting the crystal, and the proportion absorbed by it, the number of photons absorbed can be calculated. This also requires the photon energy of 6.45×10^{-19} J. For every microjoule of laser energy measured at the sample position, 1.55×10^{12} photons are provided by the laser. If only 26 % of these photons hit the crystal, and then only 84 % of those are absorbed, that equates to 3.4×10^{11} absorbed photons. From the number of absorbed photons for the whole crystal, a number per unit cell can be calculated. A 30 µm crystal with unit cell dimensions of 107.3 Å contains 2.2×10^{10} unit cells, therefore approximately 16 photons are absorbed per unit cell.

Each unit cell of an McCP-β crystal, in space group $P2_13$, contains 24 haem groups. This is a result of there being one haem group per monomer, with each of the 12 asymmetric units in the unit cell containing one protein dimer. At the 19 mM photocage concentration, approximately 15 photocage molecules occupy each unit cell. The quantum yield of the NO photocage at 308 nm has previously been determined to be 1.4 under optimum conditions (Tosha *et al.*, 2017). The yield is greater than one since two NO molecules are released from each cage molecule when a photon is absorbed. Therefore, with absorption of a single photon per cage molecule, 20 NO molecules are released per unit cell, which is close to the number of haem groups per unit

cell. Comparing the number of laser photons absorbed per unit cell with the approximately equal number of photocage molecules shows that, for a 1 µJ laser pulse, these experiments are comfortably within a single photon excitation regime. The relatively small number of photocage molecules per unit cell (15 at 19 mM concentration) suggests that this may be a limiting factor, and increasing photocage concentration may be beneficial.

These calculations do not consider any absorption of the laser by the thin film sealing the chip (see §3.9), the protein itself, or any mother liquor that remains around the crystal after chip loading (§1.4.3), which would decrease the number of photons absorbed per cage molecule to less than one. The thin film was chosen specifically for its low absorbance at 308 nm, so will not have a large effect on transmission. Protein absorption is also low at this wavelength. Additionally, the quantum yield of the photocage may be different under the conditions in this experiment. These factors would limit the NO available within the crystal to less than required for full occupancy within the structures. Inaccuracies in any of the starting values or assumptions could lead to deviation of the calculated values, although obtaining precise measurements for a complex experiment such as this is challenging.

3.6 Initial time-resolved experiments

In the first time-resolved experiments at I24 with the NO photocage, a starting energy of 16 µJ over 10 ms was used, attenuating the laser by reducing the pulse rate from 50 kHz down to 500 Hz, 1 % of full power. This starting point was based on the previously published results from crystals of P450 nitric oxide reductase (Tosha *et al.*, 2017), collected at SACLA using hydroxyethyl cellulose medium (Sugahara *et al.*, 2017) in a viscous extruder (Weierstall *et al.*, 2014). The first I24 experiment with the laser was performed, and immediate data processing showed a protein structure without NO bound in the haem site. Subsequent experiments during the same beamtime used increased laser exposures, of 10 and 100 times the starting exposure, which did show presence of NO in the haem site. Figure 3.2 shows omit maps for NO in the haem site of McCP-β structures collected with these exposures, which are described in detail in §4.2.2. None of these structures showed any density corresponding to the caging group or unlysed photocage, suggesting that the cage remains

unordered in the bulk solvent within the crystal, rather than binding to sites within the unit cell. From these first experiments, several alterations to the experimental setup and processes have been made, which have enabled much lower laser exposures to produce the same results with NO bound in the haem sites of McCP-β, with the results described in §4.2.3.

3.7 Improvements to laser alignment

The first of the improvements made to the hardware for time-resolved SSX at I24 was to the way the laser is aligned. During the very first experiments, the laser was aligned to the sample position using its position relative to the crosshairs on the OAV, while it was shone at the scintillator crystal in the sample position. The laser spot position appeared to move by a large distance between checks of the beam alignment. By coincidence, while performing laser tests with a sample in place, the laser was discovered to be misaligned on the Mylar® film on the front of the sample holder, despite appearing aligned on the scintillator. Since this observation, laser alignment has been performed using a layer of Mylar® mounted in a sample holder, rather

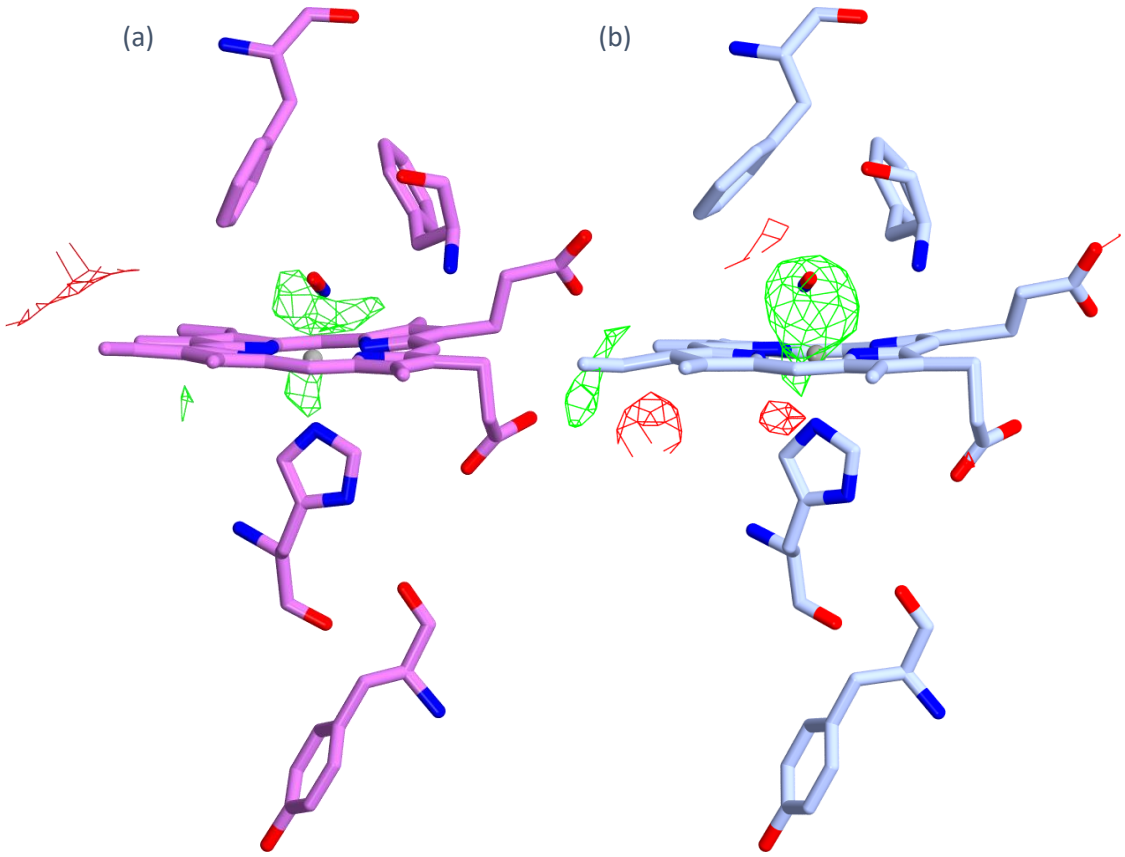

Figure 3.2: Haem sites of McCP-β with NO bound in the distal pocket, from TR-SSX structures to 2.00 Å resolution, collected with (a) 160 μJ and (b) 1600 μJ of laser energy. $F_{obs} - F_{calc}$ omit maps contoured at 3σ show electron density corresponding to the NO.

than the scintillator crystal. The laser spot is directly visible on the Mylar® film, and at high enough powers can melt holes in it. Once correctly aligned, the alignment has been stable, requiring only very minor adjustments over the course of a beamtime shift (24 h), provided the laser wavelength is not changed. Further improvements to laser alignment have enabled the laser to be focused to a smaller spot size (50 µm FWHM, down from 60 µm), increasing the power density, allowing a reduction in the total energy delivered to the sample position, while still providing the same energy to the crystal to initiate binding. In a crystal with a 30 × 30 µm cross-section facing the beam, reducing the beamsize from 60 µm to 50 µm (FHWM) increases the proportion of the area within the FHWM occupied by the crystal from 31.8 % to 45.8 %.

An improvement was also made by reorienting the chips for laser-activated experiments compared to static data collection. The silicon fixed target chips, as described in §1.4.3 and §2.6, have been used routinely for SSX at I24 before the first time-resolved experiments (Horrell *et al.*, 2021). In static experiments, the chips are mounted in their holders such that the tapered wells face the detector, as per Figure 3.3a, which is not drawn to scale. This allows the X-ray beam (shown in red, usually 20 × 20 µm) to hit the crystals through the apertures, and the diffracted X-rays have a clear path to the detector through the rear layer of thin film. The first laser-induced NO binding experiments used the same chip orientation, with the UV laser (cyan) illuminating the crystals through the aperture (b). As the laser focal spot is much larger than the chip apertures, only a small proportion of the light reaches the crystals. For a chip with a 10 × 10 µm aperture, and a 50 µm

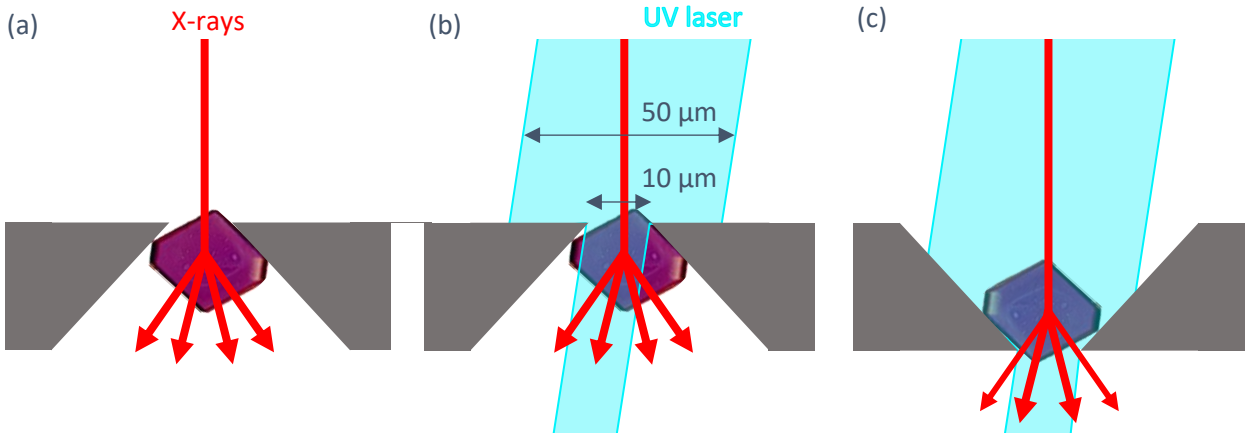

Figure 3.3: Laser illumination of crystals in wells of a fixed target chip (grey). Static SSX uses the layout shown in (a), where the X-ray beam (red) illuminates the sample through the chip, and the diffracted photons have a free path to the detector. Early laser-initiated experiments (b) used the same chip orientation, but the chip itself blocked most of the pump light (cyan). Further laser-initiated experiments (c) used the chip in the opposite orientation, where the laser can illuminate the whole crystal. Not to scale.

diameter laser spot, only 5 % of the laser light reaches the crystal, as the aperture is 5 % of the area of the laser spot. To counter this, in future experiments the chips were reoriented with the apertures facing the laser on the upstream side (c), rather than the detector. This allows the whole crystal to be illuminated with the laser, but potentially sacrificing some of the intensity from the higher resolution reflections.

3.8 Methods of laser attenuation

In the initial experiments with the laser, changes in laser power between experiments were performed by adjustment of the laser pulse rate in the laser control software, reducing it from its maximum of 50 kHz. Measurements of the power reaching the sample position showed discrepencies in that the changes in meaured power did not correspond to the changes in requested pulse rate. This is likely due to the laser power meter (described in §2.7) being designed to measure continuous wave lasers, whereas the PORTO laser was operating in a pulsed mode. This means that the changes in measured power value did not precisely correlate with changes in repetition rate. There was also some dependence of the power meter on laser alignment, but the temporary method of mounting the power meter at the sample position, required due to the multiple uses of the I24 endstation (Owen *et al.*, 2017), prevented reproduceable alignment of the meter in the sample position. These discrepencies introduced unreliability in the laser energy imparted into the crystals, as it is impractical to measure the power at the sample after every change in requested laser power, and especially in between measuring every chip, which would consume precious beamtime. Discussions with scientists responsible for the PORTO laser led to changing the way in which the laser was attenuated, to the widely accepted method of using ND filters.

Subsequent experiments used UV-reflective ND filters (Thorlabs) inserted into the beam to reduce the intensity, while keeping the laser pulse rate the same across all experiments. This meant that the laser power needed measuring only once per beamtime. The ND filters can be combined to cover more than five orders of magnitude in laser power, with precision of ±5 %. Attenuating the beam in this manner instead of changing the pulse rate also means that experiments with different laser energies still receive the same number of pulses, with the only difference being the intensity per pulse. For example, reducing the pulse rate by five

times is equivalent to an ND filter combination of OD 0.7. Although both methods should provide the same power, reduction of the pulse rate proved unreliable in early experiments. When performing similar experiments in chips at SACLA using the portable fixed target endstation (Sherrell *et al.*, 2015), the optical pump laser available operates differently. Rather than operating at kilohertz repetition rates, the laser operates at 30 Hz, providing a single intense pulse for every XFEL pulse, with the delay time between the pump and probe pulses available to be adjusted (Togashi *et al.*, 2020). The PORTO laser at Diamond is also capable of delivering laser excitation in a single pulse mode, so collecting additional datasets with this mode would provide a more direct comparison between laser pulse regimes, and rule out this as an explanation for any differences seen between SSX and SFX datasets.

3.9 Comparing thin films

Another modification made to the time-resolved experiments after the initial results was to change the thin films used to seal the samples. In normal SSX experiments without light activation (Horrell *et al.*, 2021), the usual choice of sealing film is Mylar®, available in 6 µm thick sheets (Spex SamplePrep). Mylar® is a brand of biaxially oriented polyethylene terephthalate (BoPET), a type of polyester. The properties of Mylar® that make it suitable for sealing chips in crystallography include its low moisture permeability, ensuring crystals are not damaged by dehydration during sample alignment and data collection, and its low X-ray scattering, which avoids background counts on the detector, enabling accurate measurements of low intensity diffraction spots. The very high laser energies used in the first experiments were enough to melt the film using the beam (melting point 254 °C). Unfortunately, Mylar® has a strong UV absorbance peak at 300 nm (shown in blue in Figure 3.4), making it ill-suited for experiments with the NO photocage (described in §3.2), as this is triggered by UV light in a narrow range also centred at 300 nm (Namiki *et al.*, 1997). To avoid the absorbance at 300 nm, two alternative thin films were sought and compared, including testing their UV-Vis absorbance and X-ray scattering properties. These films are EVAL™ EF-F (Kuraray), a brand of ethylene vinyl alcohol copolymer (EVOH), commonly used for food packaging (Mokwena and Tang, 2012), and polyvinylidene chloride (PVDC), both 12.5 µm thick. The grey and orange lines in Figure 3.4 show these films lack the strong absorbance peak

of Mylar® at 300 nm, making them suitable for initiation of reactions with UV light. The spectra were recorded from a single layer of film, with accumulation of 50 exposures of 10 ms.

The X-ray scattering properties of three thin films were compared by collecting detector images of scatter from a 20 × 20 µm X-ray beam passing through two layers of film in air before hitting a beamstop. This data was collected with a flux of 1.6×10^{12} ph·s^{-1} at 20 keV, using a DECTRIS EIGER2 X CdTe 9M positioned 320 mm from the films, with a beamstop immediately downstream of the films. Each frame was collected on a fresh area of the film each time, as would be the case for most crystallography experiments, with 10 ms exposures. Images summed from 10 detector frames are shown in Figure 3.5, compared with the X-ray scattering from air alone, with the edges of the detector corresponding to 1.8 Å resolution. Average counts from 100 images have been plotted against resolution for each film and air in Figure 3.6, with minor smoothing applied to the data. The gaps in the plot centred 5.4 Å and 2.6 Å correspond to gaps between modules in the detector, which are 12 pixels (900 µm) and 38 pixels (2850 µm) between panels in the horizontal and vertical directions respectively. Air scatter alone shows an average of 90 counts at lower scattering angles, with counts dropping off further from the beam centre. This drop-off means background scatter is less where the weaker higher resolution spots occur on the detector, which is beneficial for accurate data processing. Mylar® has several

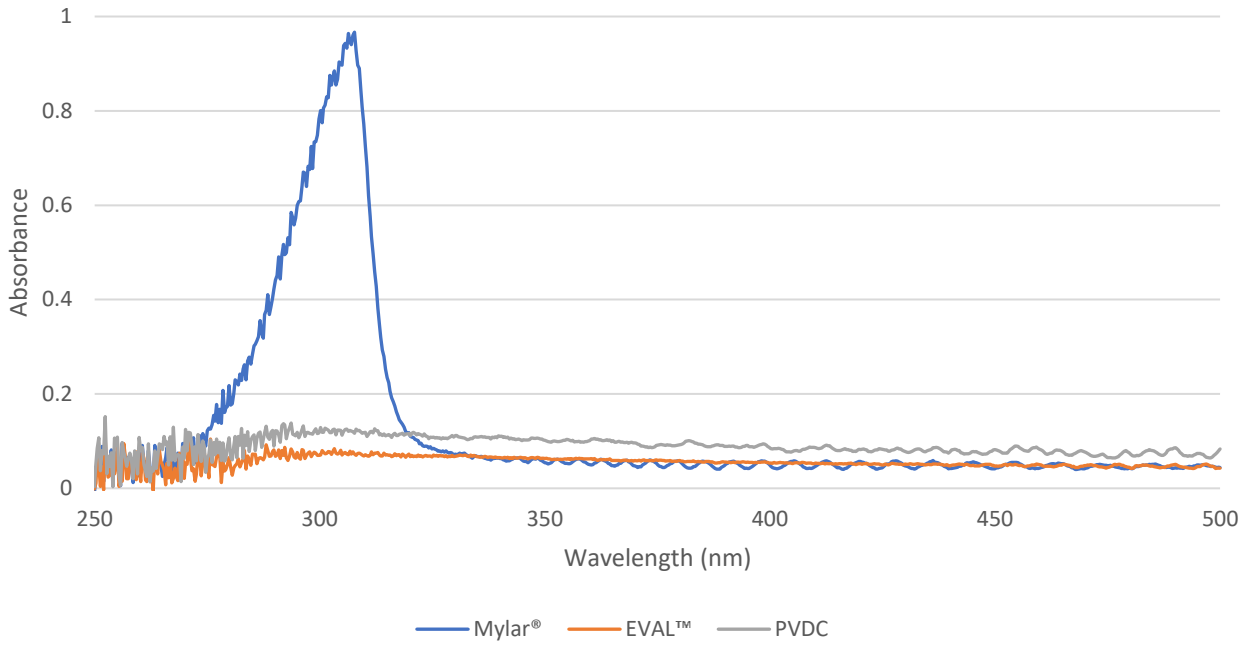

Figure 3.4: UV-Vis absorbance spectra for three thin films. Mylar® (Spex SamplePrep); 6 µm thick biaxially oriented polyethylene terephthalate (BoPET). EVAL™ EF-F (Kuraray); 12.5 µm thick ethylene-vinyl alcohol copolymer (EVOH). PVDC; 12.5 µm thick polyvinylidene chloride.

visible scatter rings, with the most prominent being at 5.5 Å resolution. EVAL™ has only three visible rings, although one of them at 4.6 Å is more intense than those of Mylar®. PVDC clearly scatters the most, with many more rings at higher intensities than Mylar®, making it less suited to sealing chips during data collection. Thinner films of Mylar® are available, but testing with 3 µm thickness concluded the lower X-ray scattering is not worth the trade-off of reduced mechanical strength while handling the film and loading samples.

EVAL™ was chosen for further fixed target experiments using the NO photocage, due to its balance of transparency at 300 nm and relatively low X-ray scattering. Because light activation only occurs from the front

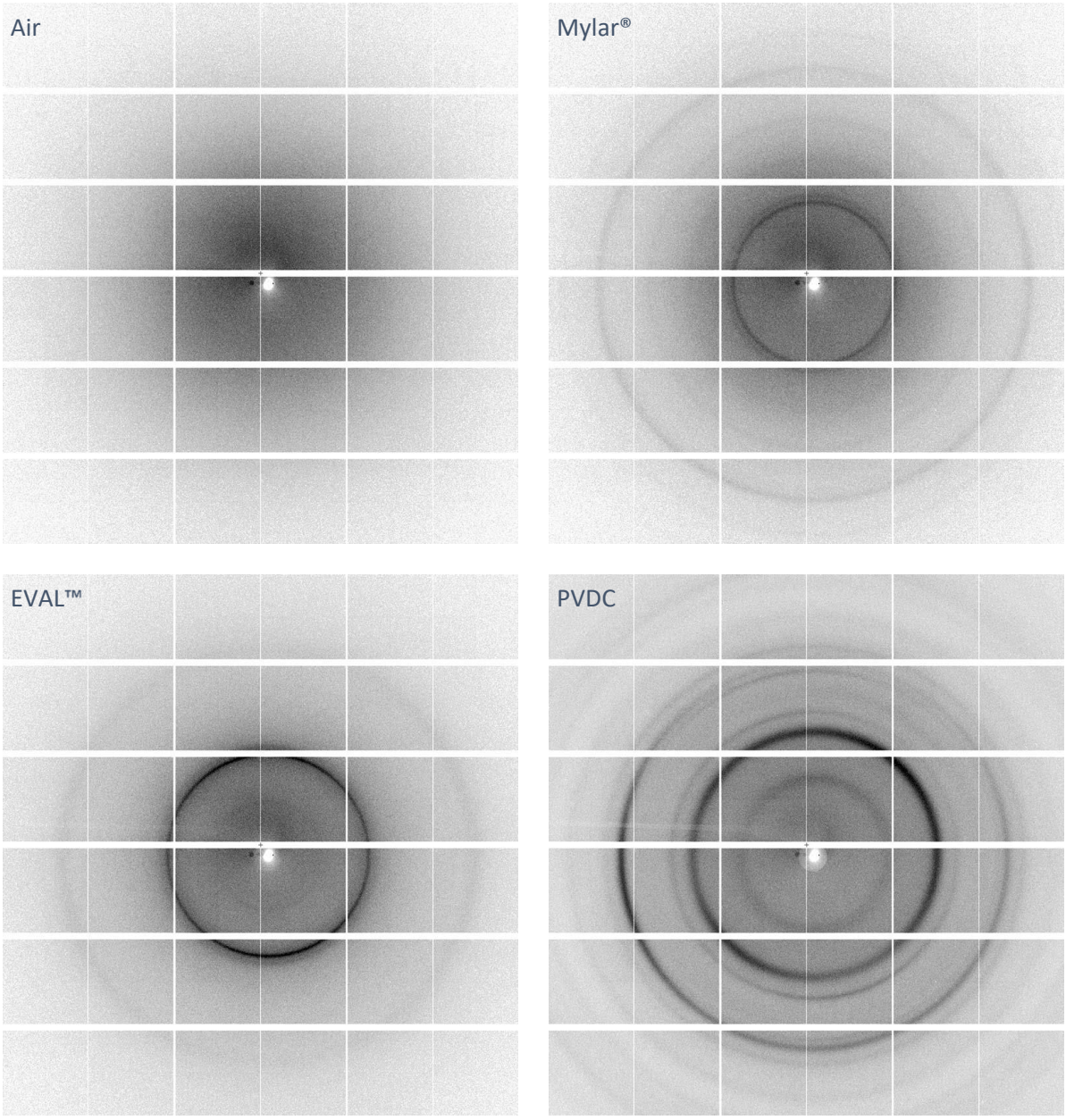

Figure 3.5: X-ray background scatter for three thin films and air alone, collected at 20 keV. Scatter from air alone is shown as a comparison in the first image. The edge of the detector corresponds to 1.8 Å resolution.

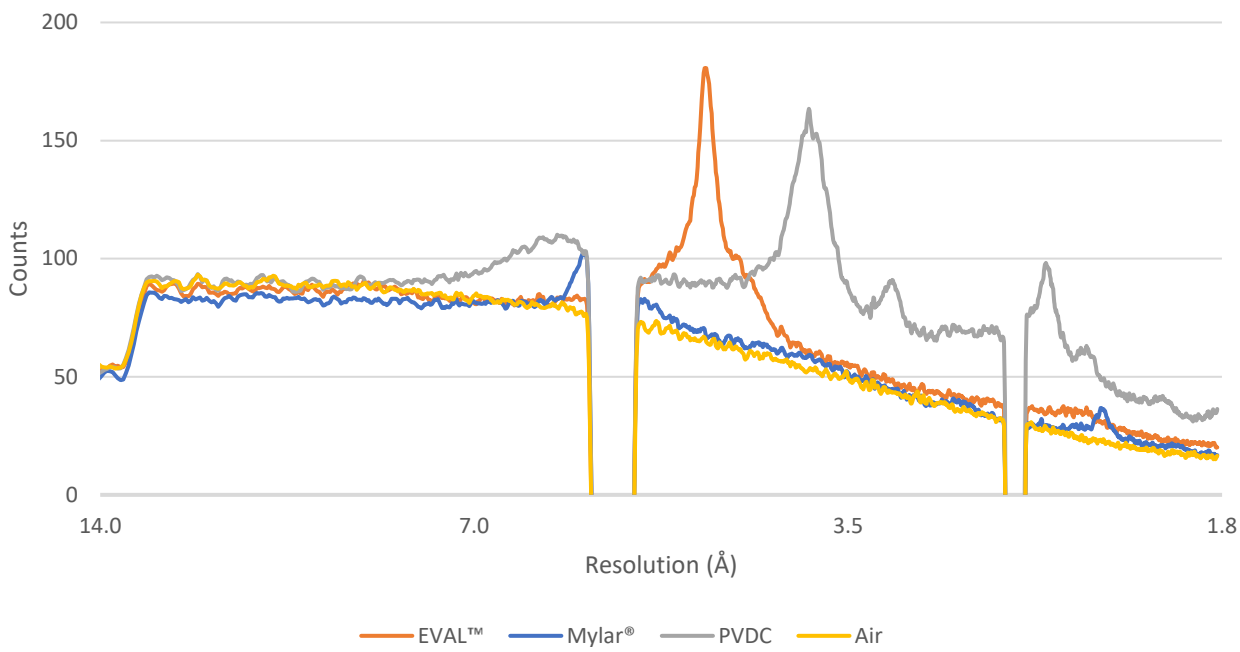

Figure 3.6: Counts of X-ray background scatter plotted against resolution for air and three thin films. The gaps at 5.4 Å and 2.6 Å correspond to gaps between modules in the detector. Data is averaged from 100 images and smoothened.

side of the chip (facing the X-ray beam), Mylar® was still used for the rear side of the chips (facing the detector) because of the lower X-ray scattering. EVOH films are commonly used in packaging of shelf-stable food in inert environments due to their low oxygen permeability (Mokwena and Tang, 2012), which is important for preventing lipid oxidation in the food. This makes them suitable for anaerobic crystallography experiments, where samples can be prepared in an anaerobic glovebox, and need to remain anaerobic for the approximate 20 min duration of the experiment, including aligning the chip on the beamline and collecting the data. An oxygen-releasing photocage exists (Ludovici *et al.*, 2002), which will enable time-resolved crystallography with proteins that bind oxygen, such as myoglobin (Barends *et al.*, 2015; Copeland *et al.*, 2006), or that use oxygen as a reaction substrate, such as formylglycine-generating enzyme (FGE), a copper-dependent oxidase (Miarzlou *et al.*, 2019). Controlling reaction initiation with oxygen is a challenge due to its abundance in air, but an oxygen-releasing photocage is convenient as light easily passes through a sealed oxygen-impermeable thin film barrier around the sample, to release oxygen within only one crystal at a time.

3.10 Time-efficient collection of long and short timepoints

Although fixed targets are capable of collecting short timepoints of reactions occurring in crystals, their main advantage is in collecting long timepoints that are unreachable by liquid jet and viscous extruder sample

delivery methods (Owen *et al.*, 2017; Schulz *et al.*, 2018; Sierra *et al.*, 2018; Pearson and Mehrabi, 2020), as described in §1.4. Unlike liquid jets and extruders, fixed targets have the advantage of being able to stop the motion of the sample in the beam. With the fixed target chips at I24, data is collected for short timepoints, up to a few hundred milliseconds, by bringing a well of the chip into the interaction region and performing both pumping and probing before moving onto the next position and a new crystal. That means a cycle consisting of moving position; laser flashing (pumping); waiting a delay period; collecting diffraction data (probing); and moving position again to repeat. This process is directed by the Geobrick LV-IMS-II motion controller (DeltaTau) responsible for the stage that repositions the sample (Sherrell *et al.*, 2015). When the motion controller indicates the stage has reached the correct position, a Zebra timing controller (Quantum Detectors, developed at Diamond) initiates the laser flash by sending a TTL signal to the laser. After the requested delay period, the Zebra sends a signal to the X-ray detector to initiate data collection for a given exposure time. Delay times and laser exposure duration are set by entering values in a GUI described in §2.6. With this mode of time-resolved data collection, the time to collect data from all 25 600 positions on a chip is only extended compared to static (non-time resolved) data collection by the delay time multiplied by the number of positions, for example a 20 ms timepoint would take 512 s (8.5 min) longer than the usual approximately eight-minute collection time.

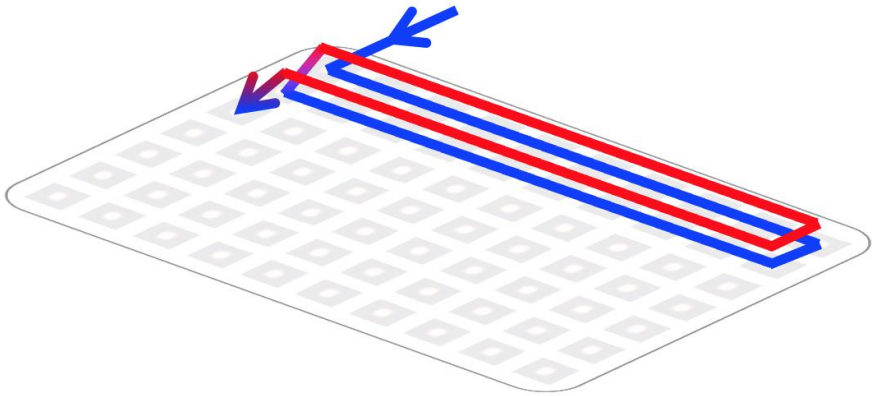

Figure 3.7: The excite and visit again method for collecting longer timepoints from reactions within crystals in fixed target chips. The arrow represents the position of the chip in the interaction region over time. The first visit, where the crystal is pumped via a visible or UV laser, is show in blue, followed by the probe visit to the same wells for diffraction data collection, shown in red.

For longer timepoints of >200 ms, waiting at each position on the chip for the whole duration of the delay time would be prohibitively time-consuming, taking several hours per chip. Instead, the ability of fixed targets to repeatably locate the same crystal and bring it to the interaction region multiple times is used. This approach is known as “hit and return” (Mehrabi *et al.*, 2020; Schulz *et al.*, 2018) or “excite and visit again” (Horrell *et al.*, 2021). With this method, several successive crystals on an even number of rows on a chip can be brought to the interaction region for light initiation, shown in blue in Figure 3.7, then the same positions returned for the probe X-ray pulse and diffraction data collection, shown in red. By altering the number of rows visited for pumping before returning to probe the crystals, the delay time can be adjusted. Because the delay time for one crystal is spent visiting other positions to either pump or probe them, rather than waiting with the fixed target stationary in one position, the total data collection time is not impacted by changes in the timepoint being studied.

Another method of time-resolved data collection at I24 takes advantage of the almost continuous X-ray beam at the synchrotron by collecting a series of several timepoints from one chip. This works similarly to the data collection method for short timepoints, but the detector is triggered multiple times in succession on the same position, to capture several closely spaced timepoints. It is based on dose series data collection using MSS, described in §1.6.6, just with the addition of the laser illumination. Data collection occurs immediately before reaction initiation (green in Figure 3.8), to get a resting state structure, then every 10 ms during and after laser

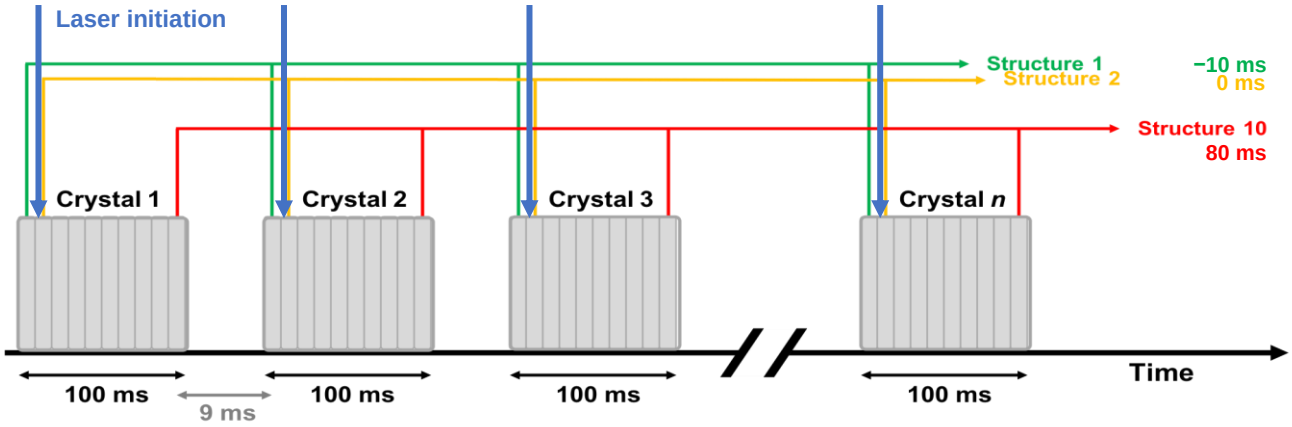

Figure 3.8: Schematic of the timeseries data collection mode using fixed target chips. For each crystal, one image is collected 10 ms before laser initiation (green), then another (yellow) is collected during the laser exposure (blue), followed by eight more images up to 80 ms (red). The images from all crystals are binned into timepoints to solve 10 structures. Data collection takes 100 ms per crystal, plus approximately 9 ms to move between positions. Adapted from Ebrahim *et al.* (2019a) with permission of the IUCr.

initiation (blue), for 100 ms total (yellow and red), with nominal timepoints of 10 to +80 ms. The resting state structure, although it should not contain any laser-induced effects, is collected from crystals that contain- any photocage used. This means that in successful experiments, it can be used as a negative control, showing that any changes seen in the structure compared to the resting state are due to the laser activation, and not the presence of the photocage. This timeseries method collects several timepoints in a very time and sample efficient manner, although the data is linked by both reaction timepoint and increasing dose, which could affect assignment of the mechanism involved for any changes in structure over time. This method of data collection has been tested unsuccessfully with NO binding. Unfortunately, this data shows presence of NO in the first timepoint, which is before the laser exposure that is meant to initiate binding, and the NO occupancy remains steady throughout the experiment, indicated by the flat line in Figure 3.9. Further explanation of possible reasons for the unsuccessful experiment is included in §5.2.2 and §5.3.2.

3.11 Monitoring of data collection and number of crystals

Due to the nature of serial data collection with still images, rather than rotation series, data collection does not just require collecting diffraction from a sufficient space group-dependent angular range from a crystal (or

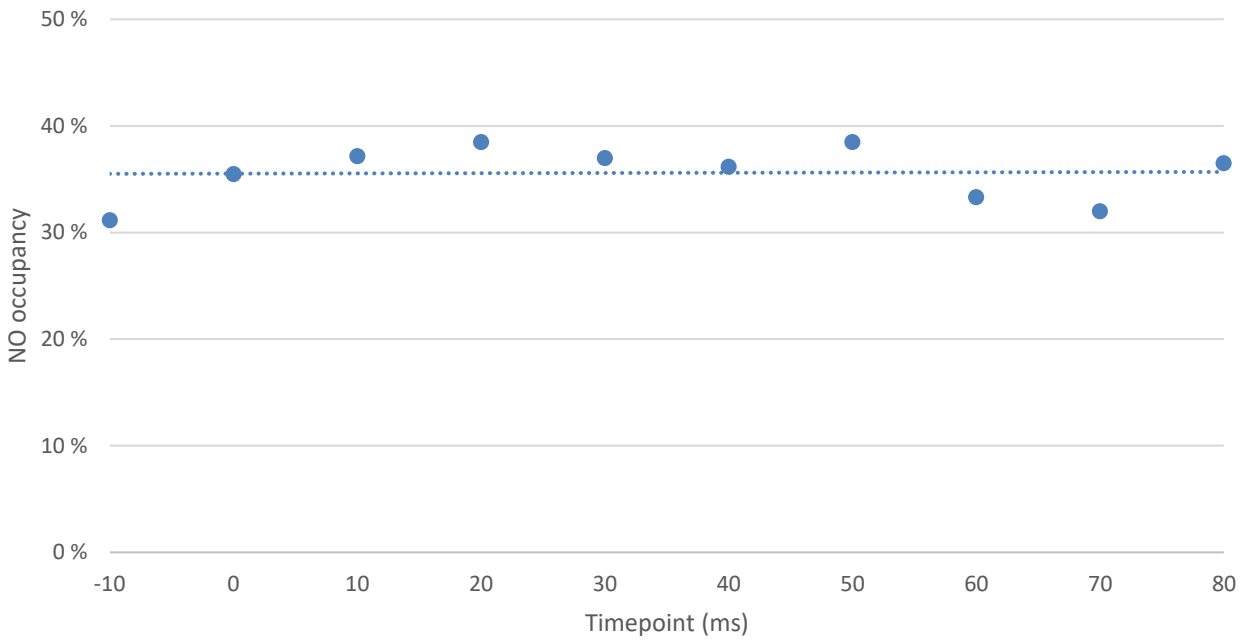

Figure 3.9: A graph of NO occupancy in dye-type peroxidase B over time before and after NO release from a photocage via UV laser exposure. The NO occupancy is flat over all timepoints, indicating NO release from the photocage occurred prior to laser exposure.

crystals). As described in §1.1, because only partial reflections are captured, integration of stills involves measuring each reflection with sufficient multiplicity (>10) to get an accurate value for the intensity. This requires having enough diffraction patterns to sample each reflection sufficiently (Winter *et al.*, 2018). Although ligands can be identified in datasets with lower numbers of images (Moreno-Chicano *et al.*, 2019), a minimum of 5000 images per dataset was aimed for in these experiments. This is demonstrated with imidazole binding in the distal haem pocket of DtpAa (Ebrahim, 2020) in Figure 3.10. The number of images required to accurately determine the structure is dependent on space group of the crystals, with higher symmetry groups requiring fewer crystals, as there are fewer unique reflections to sample. For the lower symmetry orthorhombic space group $P2_12_12_1$, in which DtpB crystallises, datasets of 5000 or more indexed crystals provide confidence when comparing NO binding geometry between structures and estimating occupancies for the bound ligand in the laser-initiated structures (Gorel *et al.*, 2021).

To monitor the number of diffracting crystals present, called hit finding, use was made of the real-time processing scripts available for SSX at I24. These scripts use *DIALS* algorithms (Winter *et al.*, 2022) to perform basic spot finding and indexing on the compute cluster at Diamond. The number of spots on each image is plotted as a graph against image number (0–25 599), which updates dynamically as data collection and real-time analysis progress, shown with data from the resting state McCP-β SSX structure in Figure 3.11a. A threshold of 15 spots is used to classify an image as a hit (blue horizontal line), from which a hit rate as a percentage is calculated (Horrell *et al.*, 2021). Clicking on a datapoint opens the relevant image in dials.image_viewer for manual inspection. In addition to the number of spots, the indexing results are used to plot graphs of unit cell parameters (c), including scatterplots comparing cell axis lengths, and histograms of all six parameters. Indexing results allow validation that the spots being identified are Bragg peaks, and that they come from the expected crystal form as determined by previous experiments. In some cases, the automatic spot finding used to generate the hit rate can identify noise as spots. This can lead to a false assumption about the true hit rate, which is why obtaining a count of indexed crystals is also important. In a small number of crystal systems, unit cell changes occur in response to reactions within the protein molecules. This means that with sufficient previous characterisation of the crystals, it becomes possible to identify reaction initiation from

indexing, without inspection of electron density maps. A heatmap of spot counts (b) can also be generated and used to judge crystal loading across the chip (see §1.4.3 and §2.6).

For basic resting state structures, a hit rate and indexing count are sufficient to monitor the data being collected, including to determine when enough diffraction patterns have been collected to sufficiently solve the structure. Further processing, including optimising parameters, can then be performed after the beamtime. However, for the more complex time-resolved experiments, it is important to have close to real-time feedback on the structure of the protein with electron density maps. This allows monitoring of whether the reaction has been initiated in the crystal. Therefore, excluding the few cases where unit cell changes indicate reaction initiation, automated processing needs to include integration, and once enough diffraction patterns have been collected to produce an interpretable electron density map, scaling and merging should

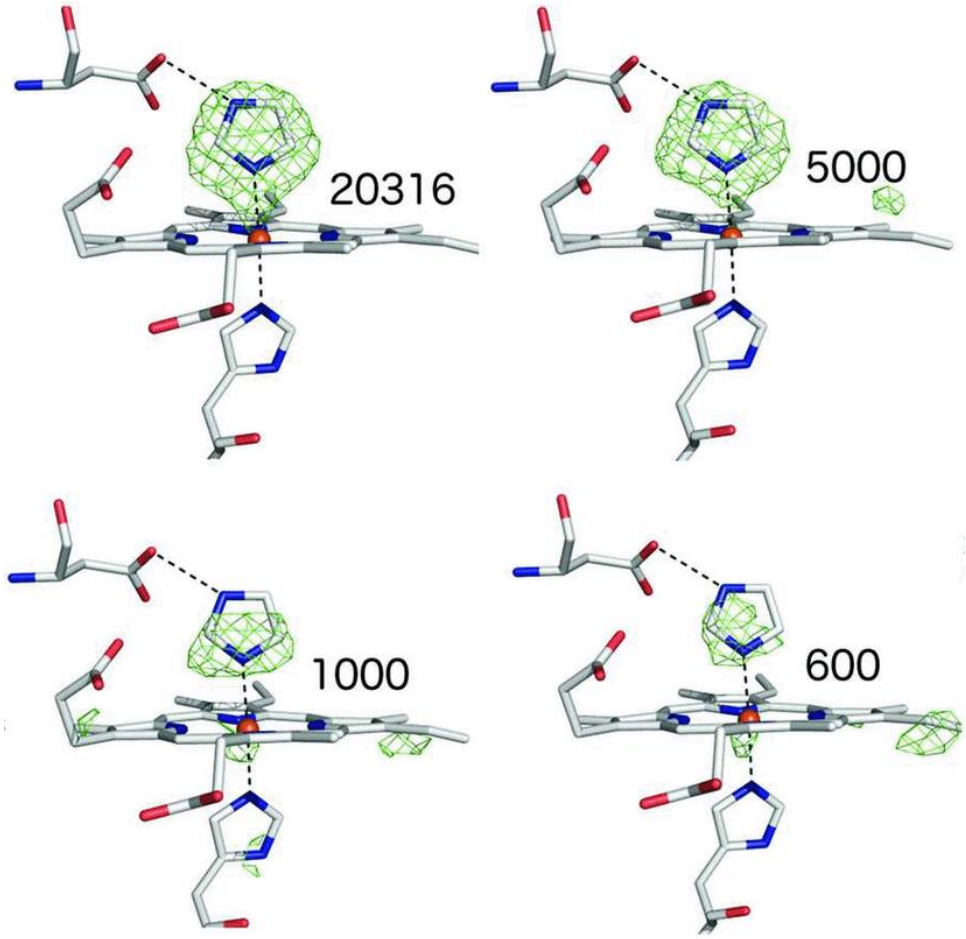

Figure 3.10: Comparison of $F_{obs} - F_{calc}$ difference maps with imidazole omitted, contoured at 3σ, for imidazole binding to DtpAa. The maps are from the same serial dataset, but with different numbers of merged crystals. Although presence of the ligand in the haem distal pocket can be identified in the maps with low numbers of images, the positioning of the imidazole is not clear in the maps generated from fewer than 5000 images. Reproduced from Moreno-Chicano *et al.* (2019) with permission from the IUCr.

be performed. Interpretation of the structure, especially $F_{obs} - F_{calc}$ difference map features, provides feedback for the next experiments during the same beamtime shift, where parameters such as timepoint, laser power, and soaking times can be altered.

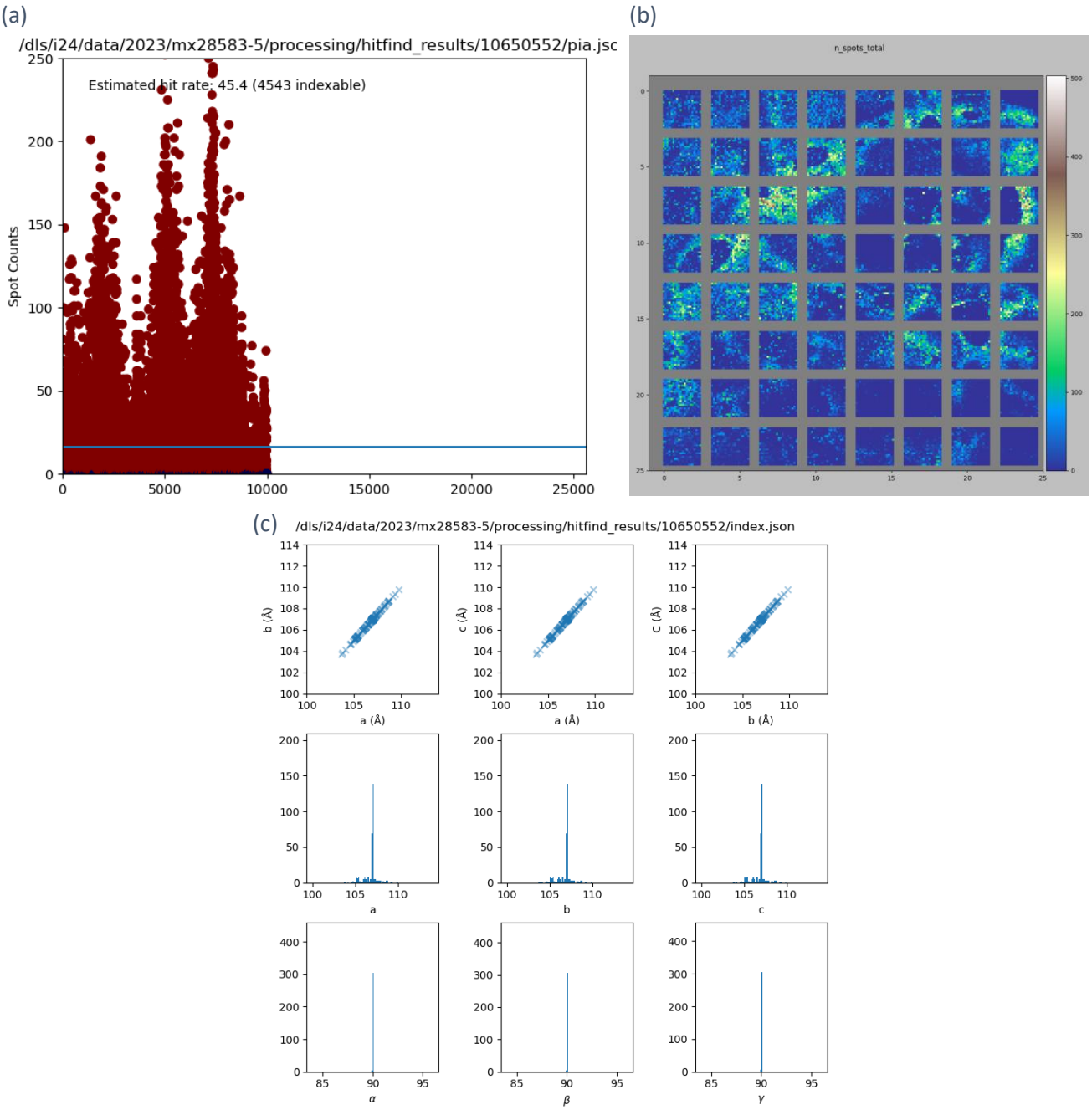

Figure 3.11: Real-time monitoring plots available for SSX at I24, showing data from the resting state McCP-β SSX structure (§4.2.1). (a) A graph of spot counts against image number in a data collection from a chip (25 600 images), as seen during the data collection when only 10 000 images have been collected and processed. Clicking on a datapoint opens the image for manual inspection. Hits are classed as having 15 or more spots (blue line). An updating hit rate as a percentage is given, along with the number of indexable images. (b) A heatmap of spot counts, showing distribution of crystals on the chip. (c) Unit cell parameters from indexed patterns. The top row shows scatterplots comparing pairs of cell axis lengths. The middle and lower rows show histograms of cell lengths and angles (which are constrained to 90° by the specified space group).

3.12 Discussion

By combining a photocage that releases NO upon UV excitation, a fixed-target sample holder with considered choice of thin film, a system for coordinated motion control and event triggering, and a laser system with mirrors to deliver light to the sample position, a protocol has been developed for performing time-resolved serial crystallography on beamline I24 at Diamond Light Source. The system can be used with both DECTRIS detectors available at the beamline, a PILATUS3 6M and an EIGER2 X CdTe 9M. As described in §4–5, this method has been demonstrated with two proteins that bind NO in the active site, which acts as a basis for studying other protein systems and reactions. Simple changes in laser wavelength make it possible to initiate other reactions using photocages, including using oxygen as a substrate (Ludovici *et al.*, 2002). The system has also been used to initiate changes in naturally light sensitive proteins, including human γ-D-crystallin, which is destabilised in the presence of UV light, leading to cataract formation in the eyes (Hill *et al.*, 2024). Investigation into reversibly switchable fluorescent proteins has also been performed at I24, using two smaller laser sources combined with a dichroic mirror (Baxter *et al.*, 2022).

As well as being installed for use at I24 at Diamond, the same fixed target chip system has been transported to the SACLA XFEL in Japan, and successfully used for time-resolved SFX experiments (Sherrell *et al.*, 2015). The portable endstation allows the same sample environment to be used to collect both long and short timepoints in a reaction, using the similar pump laser available at SACLA (Togashi *et al.*, 2020). At I24, the motion controller reports correct stage positioning, and the UV-Vis pump laser and X-ray detector are then triggered. At SACLA, the X-ray source is pulsed at 30 Hz and provides the master control for the timing. The pump laser is triggered in advance of the X-ray probe pulse, with the time gap set according to the reaction timepoint being studied. The movement of the sample stage between apertures occurs directly after each X-ray pulse, so it is ready for the next pump pulse. The detector readout is synchronised to the X-ray pulse, although the 10 fs pulse length controls the exposure time.

This chapter shows development of TR-SSX capabilities at I24 from the first unsuccessful experiments, through iterative improvements that have improved the capabilities and increased reliability. Despite this, there are

still improvements that can be made as part of future development of the beamline. For these experiments, the beamline did not feature a fast X-ray shutter which could open and close in sync with serial data collection, at the command of the motion control and timing system. The slow shutter that was available is set to open at the beginning of the data collection and not close until data collection is finished. Therefore, the shutter remains open and exposing crystals for the whole experiment. For shorter timepoints, this means that a crystal is being exposed to the beam for the duration of the delay period from triggering the reaction, and therefore has already received some X-ray dose prior to detector triggering and diffraction data collection. For longer timepoints, via the excite and visit again mode of data collection, each crystal is exposed to the X-ray beam while it is visited for laser pumping, then again for diffraction data collection. However, the crystal is still only used for one image, so dose does not accumulate throughout the whole data collection for the chip. To improve this and reduce the dose that the crystals are exposed to, a fast solenoid shutter can be installed, set to remain open only when the detector is triggered. At the time of writing, such a shutter has been fabricated and installed, but is yet to be tested or used in any experiments.

To improve the laser alignment and remove the parallax between the X-ray and laser beams, a new on-axis viewer has been installed on I24, allowing the laser (and other pump light sources) to illuminate the sample through the viewer using a system of mirrors and an optical fibre. A CAD drawing is shown in Figure 3.12, with the important new parts being the fibre connection, and the two new mirrors in black 45° holders. One of the mirrors is semi-transparent and allows the photoexcitation light (cyan) to be merged with the viewing microscope (magenta), then go through the final drilled mirror (not shown) and lens, onto the X-ray beam axis (red). This eliminates the effect of sample Z axis positioning on the laser alignment and should simplify laser alignment to make it more stable, saving precious beamtime. Early testing with the laser being brought to the sample position via the new OAV has had problems with focusing of the photoexcitation light to a small enough spot size, as the light must pass through some of the same lenses as the sample viewing light on the X-ray beam axis. Additionally, the scatter guard in front of the lens interacts with the laser light.

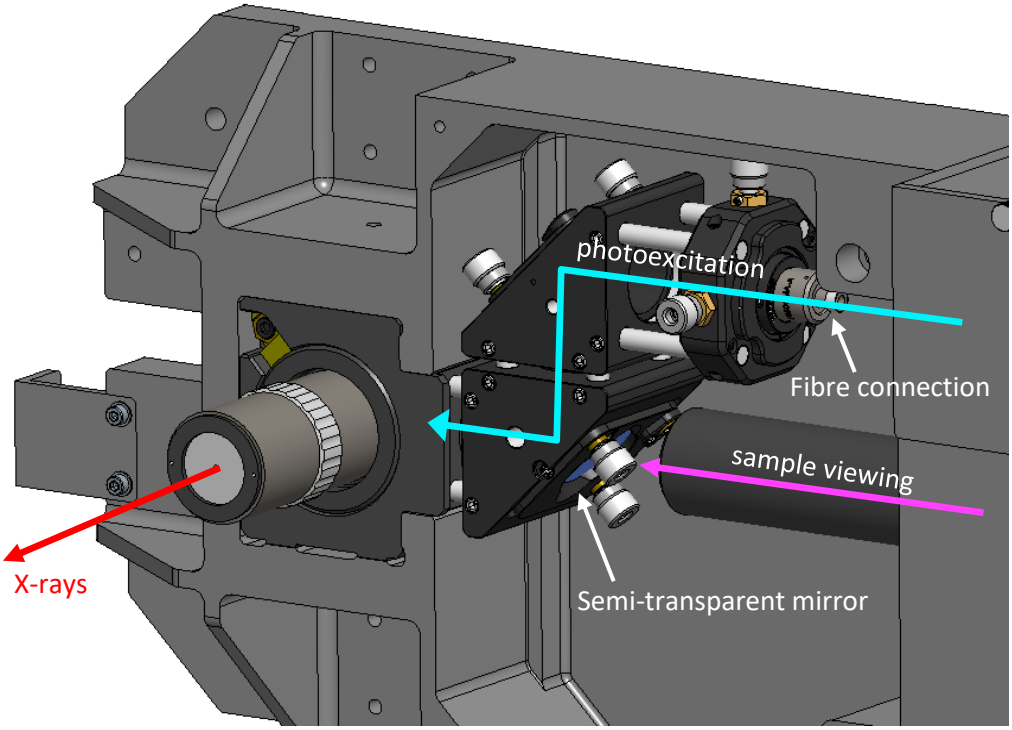

Figure 3.12: Upgraded on-axis viewer for beamline I24. Via a series of mirrors, UV-Vis photoexcitation (cyan) from an external source can be directed to the sample position on the same axis as the X-ray beam (red) along with the sample viewing microscope (magenta). A mirror (not visible) and lenses with a central drilled hole for the X-ray beam to pass through bring both illumination and viewing onto the X-ray beam axis. This enables simple alignment of the X-ray and laser beams with the sample.

In addition to using light, the same chip system at I24 has been used with substrate mixing to initiate reactions in crystals. The same chip stage and timing control is used, but the laser and photocage is replaced with a piezoelectric ejector (microdrop Technologies), which can reliably fire picolitre sized drops of ligand solution into the interaction region and hit a specific well of the chip. In this setup, the thin film on one side of the chip is replaced by a humidified local atmosphere to allow ligand drops to reach the crystal, while still preventing crystal dehydration (Shilova *et al.*, 2022; Glerup, 2022).

4. Nitric oxide binding to McCP-β

4.1 Introduction

Cytochromes *c* are haem proteins containing a *c*-type haem group, and featuring a conserved CXXCH motif that covalently binds the haem to the peptide component (Bowman and Bren, 2008). Cyts *c*′ specifically have pentacoordinate haem iron in their resting state and are found in the periplasm of gram-negative bacteria. The majority of cyts *c*′ characterised have been alpha-helical (Choi *et al.*, 2005; Fujii *et al.*, 2017; Kekilli, 2015), forming four helix bundles, although a smaller number of examples have been found with primarily beta-sheet structures, termed cytochromes *c*′β (Elmore *et al.*, 2007; Adams *et al.*, 2019). These have been identified in methanotrophic (Bergmann *et al.*, 2000) and ammonia oxidising (Abendroth *et al.*, 2022; Liew *et al.*, 2019) bacteria, and are homologous to cytochrome P460 proteins, often found in the same organisms, see Figure 1.11. However, cyts P460 have a haem-Lys crosslink which is absent in cyts *c*′β (Adams *et al.*, 2019). This crosslink provides the protein with the ability to catalyse conversion of hydroxylamine (NH_2OH) to nitrous oxide (N_2O) via nitric oxide (NO). As cyts *c*′β lack the crosslink, their function is different from P460s, resembling that of cyts *c*′α, and are known to bind NO. A more detailed overview of cytochromes is found in §1.7.

4.1.1 Cytochrome *c*′-β from *Methylococcus capsulatus* (Bath)

A cyt *c*′β has been identified- in the obligate methanotroph *Methylococcus capsulatus* (Bath) and has been characterised spectroscopically and crystallographically (Adams *et al.*, 2023; Adams *et al.*, 2019). It is abbreviated as McCP-β. This protein is of interest because of its potential role in the global nitrogen cycle, and because few cyts *c*′-β have yet been characterised. This chapter describes the acquisition of the first crystal structures of McCP-β at room temperature. This includes a damage-free structure from serial femtosecond crystallography (SFX) and investigation of the binding of NO to the haem site using time-resolved serial synchrotron crystallography (TR-SSX). A crystal structure collected at room temperature is more likely to reflect the state of the protein, including bound ligands, at physiologically relevant temperatures, than one collected under cryogenic conditions (Sharma *et al.*, 2023; Fischer *et al.*, 2015). X-ray photoreduction of metal centres is also important as these are where catalytic and ligand binding activities take place (Bowman *et al.*,

2016). Radiation damage and serial crystallography techniques are described further in §1.2 and §1.6.1–1.6.2. Room temperature data collection also allows investigation of protein dynamics with time-resolved crystallography, as described in §1.6.4 and §3.

Reaction initiation is an important part of time-resolved crystallography, as a reaction needs to be started in a whole crystal at the same time, as described in §1.6.5. To do this, use has been made of a photocage, an organic molecule which releases a NO by photolysis with UV radiation, described in §1.6.4–1.6.5. Cytochromes *c*' make suitable targets for developing time-resolved crystallography, as NO binds to the haem via a vacant distal site (Winter *et al.*, 2022; Barty *et al.*, 2014), in contrast to many other haem proteins, which bind small gas molecules via displacement of a water molecule. The NO binding in the dry pocket provides an obvious change in electron density, as two non-hydrogen atoms bind in place of none, whereas NO displacing water involves the more subtle change of one non-hydrogen atom to two. The obvious electron density change can be used to quickly identify successful NO binding, when real-time processing is required to provide fast feedback on the outcome of an experiment, so as to inform subsequent experiments following in the same beamtime shift and so aid in the development of time-resolved crystallography methods using photocages.

4.1.2 Cytochrome *c*' from *Alcaligenes xylosoxidans*

The cytochrome *c*' in *Alcaligenes xylosoxidans* (AxCP), like that of *Methylococcus capsulatus*, has the conserved CXXCH motif for binding *c*-type haem (Kekilli *et al.*, 2017a), and both dimeric proteins bind NO at the haem site. Whereas McCP-β has a primarily β-sheet structure, AxCP consists of an antiparallel four α-helix bundle, with the haem in the centre, shown in Figure 4.1. As well as structural differences, the two proteins have different NO binding mechanisms. NO binding to McCP-β is discussed above (§4.1.1). In AxCP, NO binding proceeds via a multi-step process involving two NO molecules in different binding sites, depicted in Figure 1.12. First, NO binds to the distal side of the haem, similarly to in McCP-β. However, in AxCP this causes a dissociation of the proximal histidine residue (His 120) to a distal five-coordinate iron species, then allowing a second NO molecule to bind in place of the haem. Finally, the first NO molecule dissociates to leave a five-coordinate product with NO bound proximally.

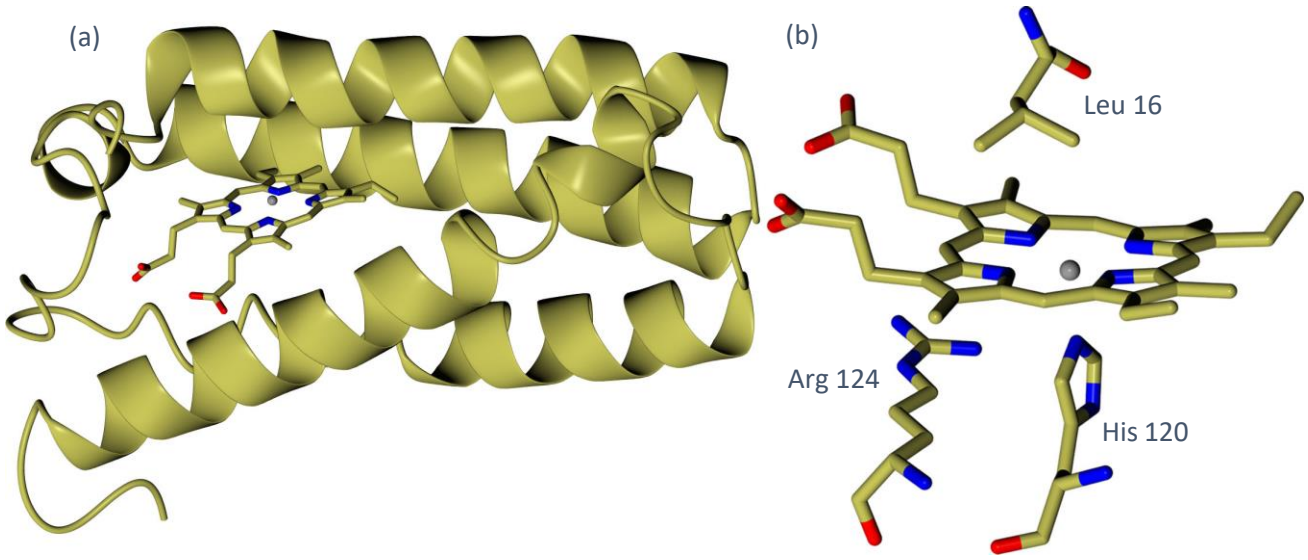

Figure 4.1: Cryogenic structure of AxCP in its ferrous (reduced, Fe^{2+}) form. (a) Ribbon view of a single monomer, showing the helical bundle and haem position. (b) The pentacoordinate haem site, including the three relevant neighbouring residues. The conserved His 120 is shown above the haem on the proximal side. PDB: 4cda from Kekilli *et al.* (2014).

Besides its NO binding mechanism, AxCP is also of interest because of the conformational change that occurs to the haem site arginine residue upon change in the redox state of the haem iron. In its ferric (Fe^{3+}) form, the guanidino group of the Arg 124 sidechain sits parallel to the plane of the haem (Einsle *et al.*, 2007). In the reduced ferrous form (Fe^{2+}), the arginine side chain mostly (70 % occupancy) reorients perpendicularly to the plane of the haem, as shown by the arrow in Figure 4.2. This change is seen in structures that were chemically reduced prior to cooling, or X-ray photoreduced at 160 K or 180 K, but not the structure photoreduced at 100 K, implying that at 100 K the arginine is cryotrapped (Kekilli *et al.*, 2014). This structural change can be

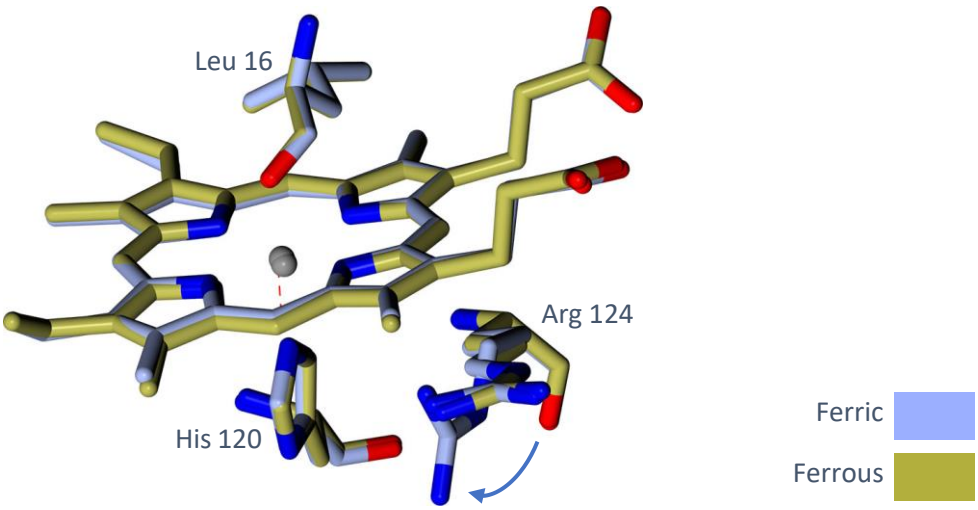

Figure 4.2: The haem site of AxCP, with the ferric (yellow, PDB: 4cda) and ferrous (blue, PDB: 4cjo) redox states superimposed. In the ferric form, the guanidino group of Arg 124 sits parallel to the plane of the haem, but in the reduced form, the residue reorients with 70 % occupancy perpendicular to the haem.

used to infer the redox state of the iron without more specialised experimental techniques such as X-ray spectroscopy, resonance Raman spectroscopy (Kekilli *et al.*, 2017a) or online UV-Vis spectroscopy (Hough *et al.*, 2008), as long as the data collection temperature is greater than 160 K. Redox state determination using only standard diffraction data allows serial and time-resolved crystallography techniques and sample delivery to be used more easily (see §1.4). This allows protein dynamics to be accurately investigated in the correct validated redox states, and also provides a convenient way of measuring photoreduction as a symptom of specific X-ray damage to the sample (see §1.6.6).

4.2 Results

4.2.1 Room temperature structures of McCP-β

McCP-β was expressed and purified as described in §2.1. Crystallisation conditions have already been optimised for growing large single crystals (~100 µm), published in Adams *et al.* (2019). The large crystals, grown with vapour diffusion, used a protein concentration of 15 mg·mL^{-1}, and polyethylene glycol 550 concentration of 35 % (v/v). To adapt this condition for growing microcrystals, on the order of 20 µm in each dimension, the concentrations of precipitant and protein were increased, and the crystallisation method was changed. Microcrystal conditions were trialled with batch crystallisation, mixing protein and precipitant directly in a tube. This allows the mixture to start in the nucleation zone of the crystallisation phase diagram (see §1.5 and Figure 1.7), so thousands of nucleation events occur soon after mixing, rather than nucleation occurring only upon reaching the nucleation zone by vapour diffusion. Optimisation of the conditions was performed to have sufficient nucleation for thousands of crystals to grow, without being overconcentrated and forming non-crystalline precipitates. Crystallisation trials with 20 µL volume were performed in PCR tubes, before scaling up to 400 µL batches in 1.5 mL microcentrifuge tubes. Test batches were examined under a microscope after 24 h, and their dimensions measured. Knowing crystal dimensions is important for preparing fixed target chips, so that a chip with the correct aperture size is selected to avoid crystals passing through the apertures and being lost. Crystal size is also important for reaction initiation because of the limited penetration depth of the laser light. A haemocytometer was used to measure the concentration of crystals within the

sample. To increase crystal concentration, crystals were sedimented over time or by gentle centrifugation, and a volume of the mother liquor supernatant removed. The ideal concentration of crystals for loading a fixed target chip is approximately 10^7 crystals·mL^{-1}. Final conditions were reached with 68–70 % (v/v) PEG 550, and 40 mg·mL^{-1} protein, as described in §2.2. The crystallisation was found to be variable between batches of purified protein, so separate trials were conducted for each. Figure 2.2a shows crystals of McCP-β in mother liquor, with approximate dimensions of 30 × 30 × 30 μm.

An SFX structure of McCP-β was collected at the Spring-8 Angstrom Compact Free Electron Laser (SACLA, Japan) by other members of the research group, using the fixed-target chip system described above (§1.4.3 and §2.5). Data from four chips was merged and the structure was solved to 1.80 Å by molecular replacement (Vagin and Teplyakov, 2010), using a conventional structure collected at 100 K (PDB: 6hih) as a search model (Adams *et al.*, 2019). This SFX model shows the structure of the protein with zero X-ray dose and at room temperature. An SSX structure was also collected at beamline I24 at Diamond Light Source, using the same method. This was solved to 2.15 Å resolution. Data processing and refinement statistics for the structures are listed in Table 4.1.

Data processing methods for serial crystallography are less mature than for conventional rotation crystallography, and the nature of collecting still images with partial reflections makes processing more complex (Kirian *et al.*, 2010; Brewster *et al.*). Obtaining the best quality data requires optimisation of spot finding and indexing parameters and ensuring that an accurate description of the detector geometry is provided. For these serial structures, interactive examination of images was required to obtain the best quality maps for solving accurate protein structures. For each beamtime, the data from the first chip collected was used to refine the detector geometry, specifying the precise detector distance and where on the detector the beam centre lies. This provided a geometry file which could be used for all datasets from the beamtime. Spotfinding and indexing results from a subset of the data were compared against each other during iterative optimisation, to find the best parameters to apply to the whole dataset. To evaluate processing parameters, the number of indexed images from the dataset was compared, as including more images improves data

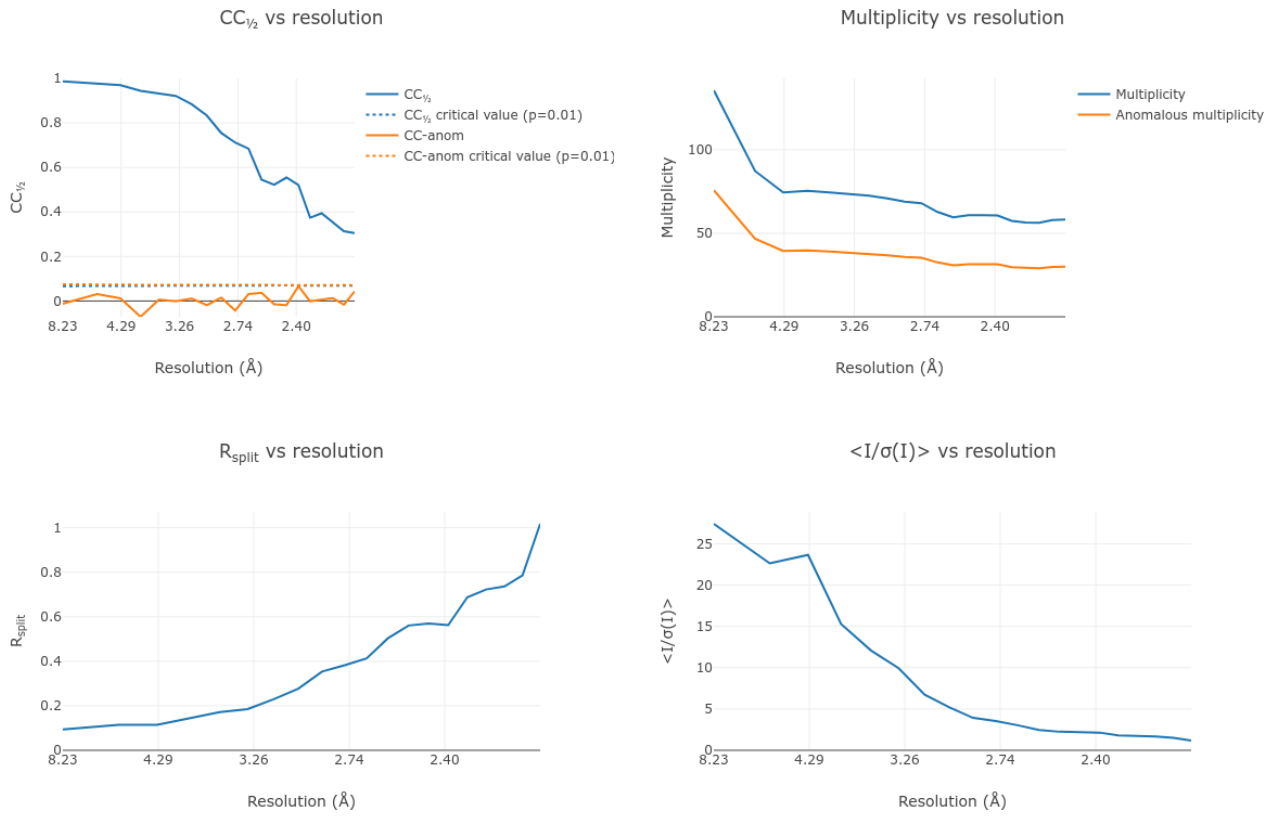

Figure 4.1: Data quality for the resting state SSX structure of McCP-β to 2.15 Å resolution, plotted against resolution, for data processing in xia2.ssx (Winter *et al.*, 2022).

quality, especially the completeness and multiplicity. Examples of data processing statistics plotted against resolution, taken from the log files of xia2.ssx, are shown in Figure 4.3. Most data processing was performed using *DIALS* (Waterman *et al.*, 2013; Winter *et al.*, 2018; Winter *et al.*, 2022), except for the SFX structure of McCP-β, which was processed using *CrystFEL* (White *et al.*, 2012; White *et al.*, 2016). See §2.9 for more details on data processing and structure refinement.

The β-sheet homodimeric structure can be seen in Figure 4.4, which is a superposition of the 100 K (yellow), SFX (blue), and SSX (coral) structures. A ribbon view of the protein dimer (a) shows differences for some β-strands between the zero dose SFX structure and the two synchrotron structures, especially in the residue ranges 15–21 in chain A and 73–81 in chain B. Viewing these regions more closely (b–c) reveals these differences extend only to side chain positioning. Met 20 in chain A has a shared conformation in the SFX and 100 K structures, and two conformations that are unique to the SFX and SSX structures. In chain B, Val 74 has a second conformation only seen in the SFX structure, and Lys 79 occupies a different conformer at room

temperature compared to 100 K. The B-factor for the Cε atom toward the end of the sidechain is much higher (42.85 Å²) in the SSX structure compared to the SFX structure (22.99 Å² and 25.75 Å²) and the 100 K structure (36.73 Å²), so the difference seen in the SSX structure may be an artefact of the disorder in this residue and resolution of the structure, and not a true different conformation. Looking at the B-factors for the whole structures, shown with the ribbons view in Figure 4.6, there is higher disorder in the terminal regions and loops, which are to be expected. This is seen especially in the region with the residue range 24–30, labelled in chain B in (a), potentially explained by the separation of the region from the main β-sheets allowing more

Dataset	Resting state SFX	Resting state SSX	1600 µJ laser excitation	160 µJ laser excitation	PROLI NONOate soak	Dark negative control
Resolution (Å)	33.89−1.90 (1.93−1.90)	75.82–2.15 (2.19–2.15)	37.90−2.00 (2.03−2.00)	43.79−2.00 (2.03−2.00)	37.94−1.95 (1.98−1.95)	75.88–2.20 (2.24–2.20)
Space group	$P2_13$	$P2_13$	$P2_13$	$P2_13$	$P2_13$	$P2_13$
Unit cell (Å), $a = b = c$	106.7	107.2	107.2	107.3	107.3	107.3
Merged crystals	14 394	6692	18 976	19 929	12 517	4368
Unique reflections	61 896 (3121)	22647 (1156)	27 994 (1378)	28 056 (1395)	30 288 (1487)	21205 (1049)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Multiplicity	233.7 (159.9)	69.9 (58.1)	159.8 (19.0)	223.4 (18.0)	140.9 (16.5)	40.0 (34.0)
$CC_{1/2}$	0.938 (0.639)	0.984 (0.279)	0.986 (0.489)	0.983 (0.488)	0.981 (0.459)	0.974 (0.373)
R_{split}	0.204 (0.492)	0.188 (1.017)	0.118 (0.561)	0.133 (0.572)	0.129 (0.660)	0.184 (0.753)
Mean I	284 (1.92)	28.0 (7.9)	21.9 (4.5)	15.5 (4.2)	38.7 (5.6)	44.1 (3.3)
I/σI	4.0 (2.1)	7.7 (1.2)	12.9 (8.6)	13.3 (8.3)	1.39 (0.46)	9.0 (1.7)
N_{obs}	14 465 627 (498 904)	1 583 238 (67 218)	4 473 881 (26 237)	6 267 430 (25 124)	4 267 720 (24 491)	848 796 (35 661)
R_{work}	0.185	0.184	0.175	0.193	0.183	0.185
R_{free}	0.213	0.227	0.218	0.223	0.222	0.241
RMSD bond length (Å)	0.0115	0.0131	0.0170	0.0151	0.0233	0.0131
RMSD bond angles (°)	2.178	2.623	2.678	2.830	2.790	2.793
Ramachandran favoured (%)	95.91	95.91	92.91	95.90	94.03	96.28

Table 4.1: Data processing and refinement statistics for McCP-β structures, including resting states collected at SACLA and I24, time-resolved NO-bound structures collected using a photocage, and an NO-bound structure soaked with PROLI NONOate. Values in parentheses refer to the outermost resolution shell.

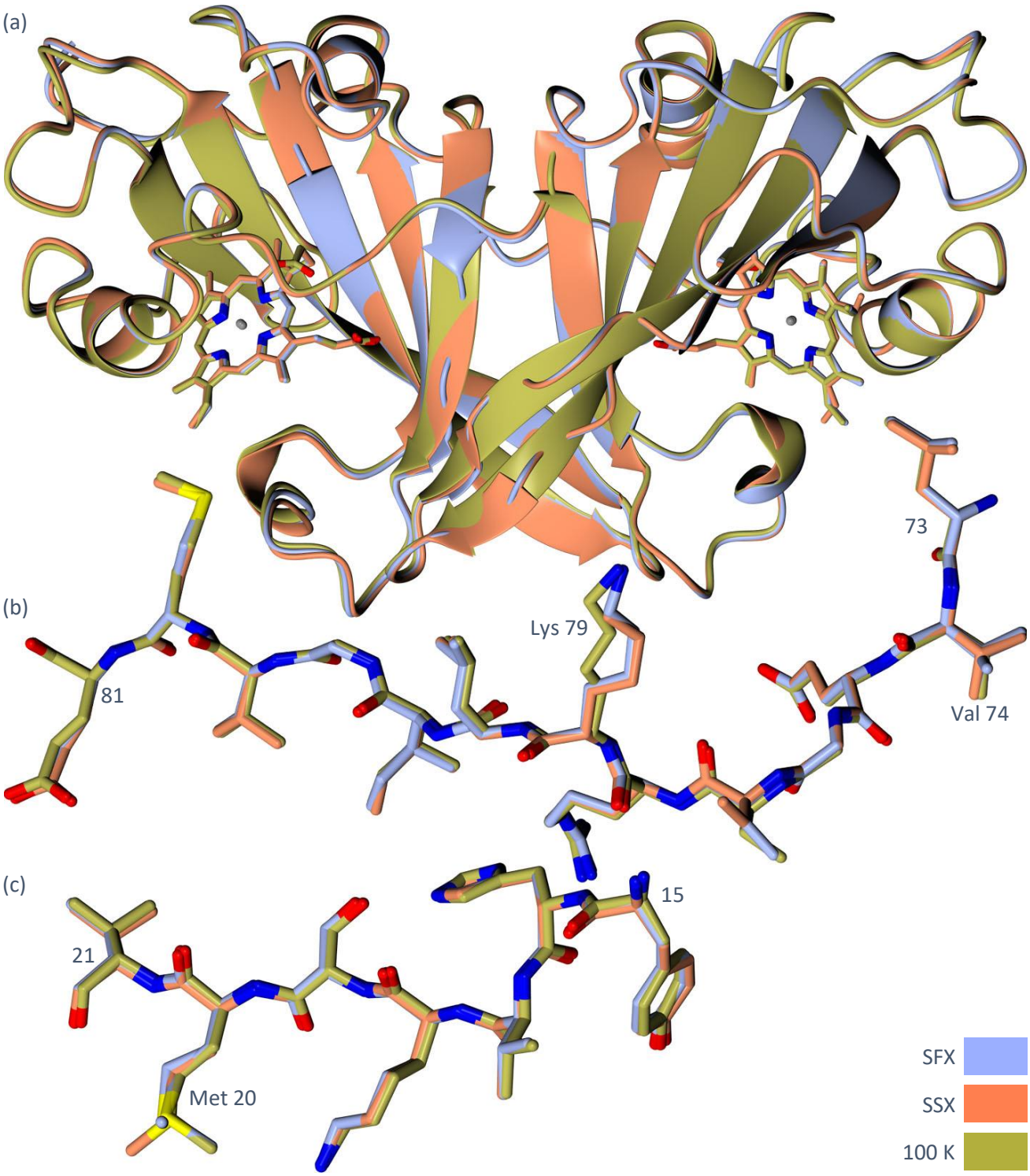

Figure 4.2: Superposition of three McCP-β structures. (a) Ribbon view of the dimer, showing some differences between the β-strands of the SFX structure (blue, 1.90 Å resolution) compared to the 100 K (yellow, 1.61 Å, PDB: 6hih) and SSX (coral, 2.15 Å) structures. Residues (b) 15–21 of chain A and (c) residues 73–81 of chain B show that differences between structures are confined to sidechain conformations, including Met 20 in chain A, and Val 74 and Lys 79 in chain B.

flexibility. The B-factors are also higher than average in the Phe 32 residue in all structures, indicating the flexibility to move further from the haem upon NO binding (see below).

The *c*-type pentacoordinate haem is shown in Figure 4.5, and lacks the lysine crosslink that is found in McP460 and other proteins with the P460 fold (see §1.7.1). The haem site includes His 123 on the proximal side, which is conserved in many cytochromes (Bowman and Bren, 2008). The histidine is hydrogen-bonded to the backbone of Tyr 131. The distal side of the haem contains residues Phe 32 and 61, which form a hydrophobic pocket for ligand binding. The Phe 32 sidechain occupies a different position, further from the haem in the

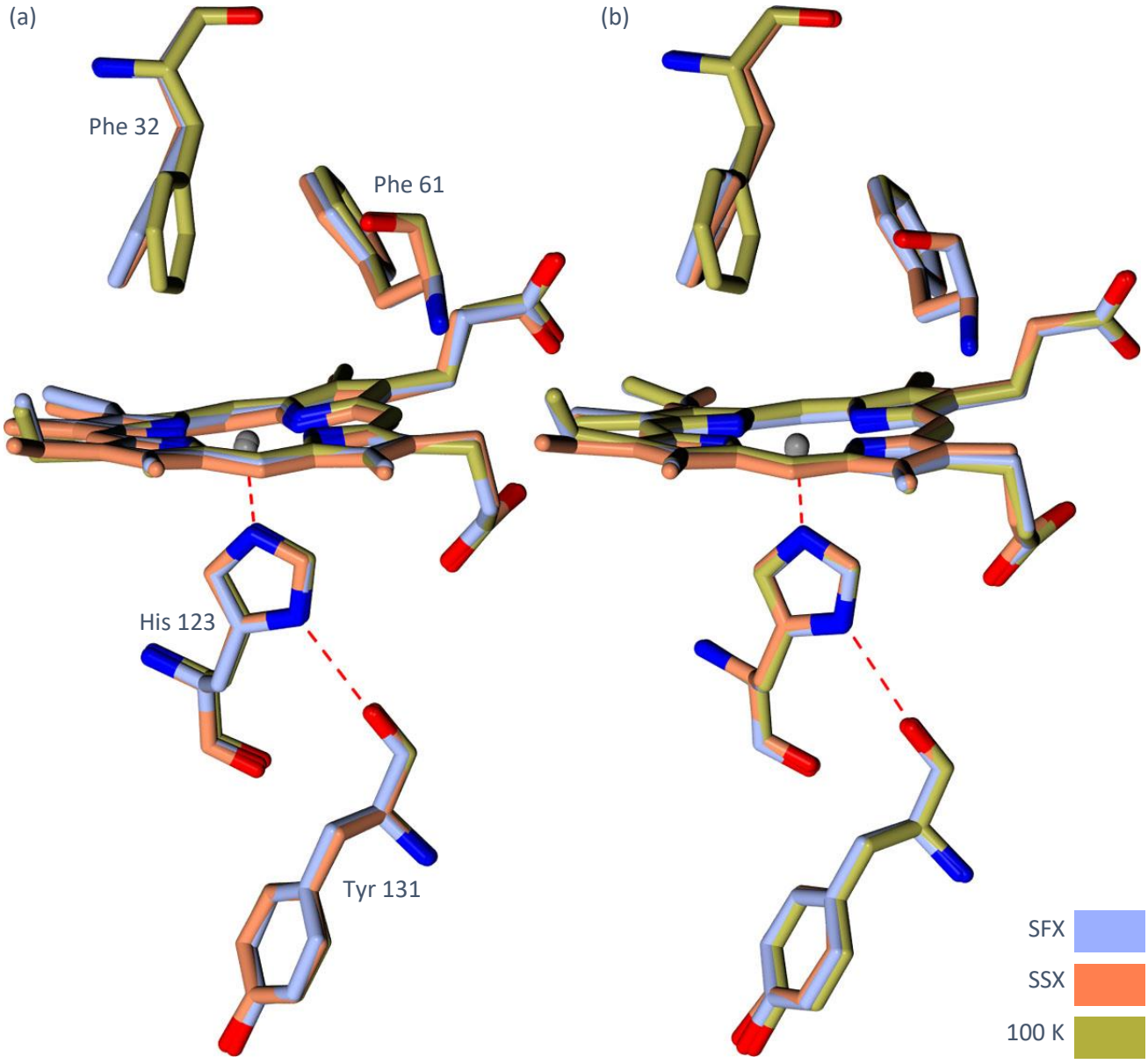

Figure 4.3: Superposition of the haem sites of three McCP-β structures, showing chains A and B. Displacement of Phe 32 away from the haem is visible in both room temperature structures. The colours are the same as Figure 4.4: SFX; blue, 1.90 Å resolution, 100 K; yellow, 1.61 Å, PDB: 6hih, and SSX; coral, 2.15 Å.

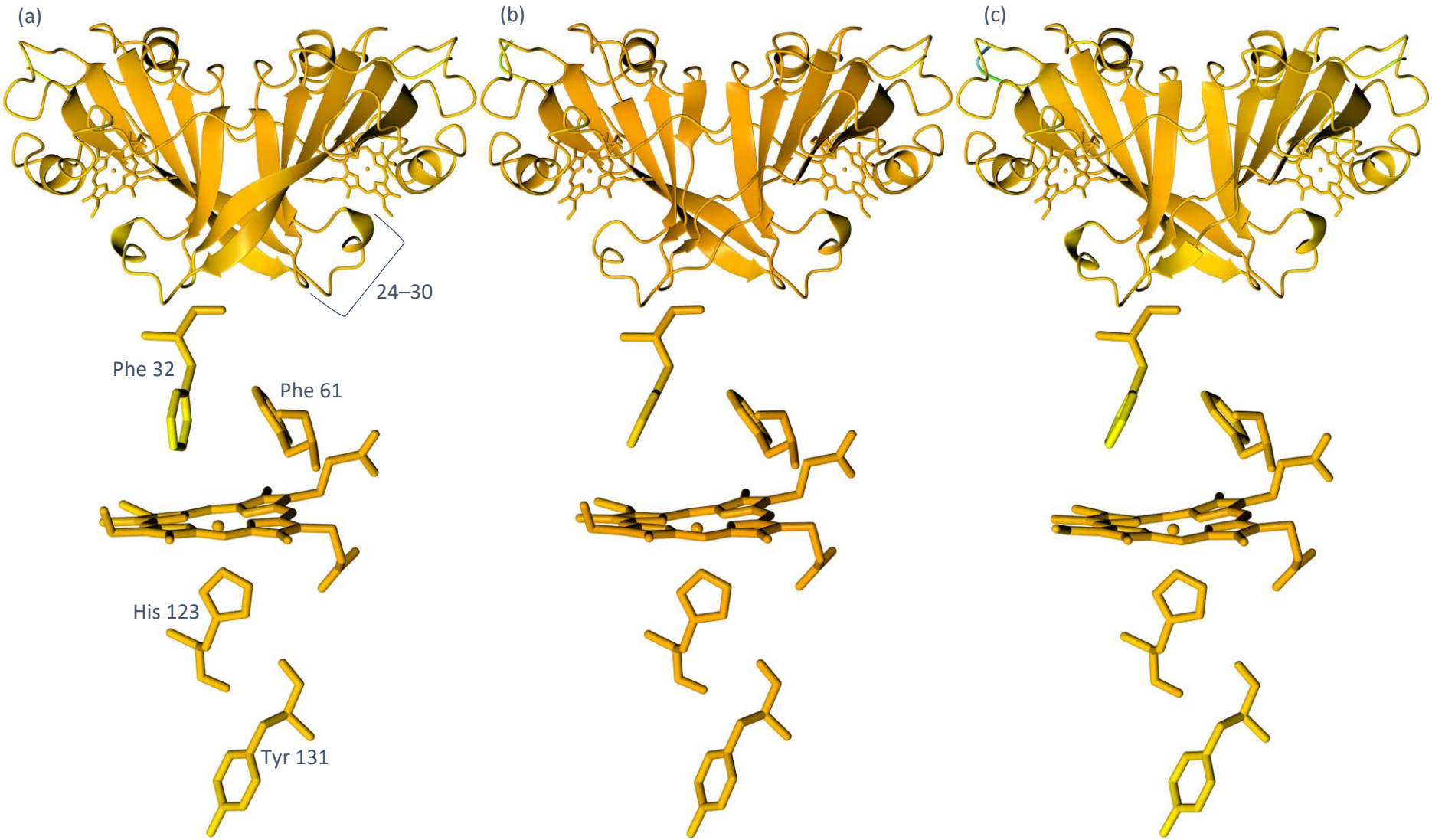

Figure 4.4: McCP-β structures coloured by B-factor. (a) 100 K (PDB: 6hih), (b) SFX, (c) SSX. Orange indicates low B-factor, increasing through green to blue.

room temperature structures compared to the 100 K structure. The Cζ atom of the phenylalanine benzyl ring is displaced by 1.13 Å and 0.99 Å in chains A and B respectively in the SFX structure. The water-free (dry) distal pocket of the haem in the 5-coordinate resting state is useful for developing time-resolved methods for initiating ligand binding, as the ligand binds in an empty space, rather than displacing a water molecule, so the change in electron density is more obvious, especially with potentially suboptimal parameters used in initial processing during the beamtime.

Prior to the laser-initiated NO binding experiments described below (§4.2.2), a positive control was collected in the form of a structure from crystals soaked with an excess of NO, collected to 1.95 Å resolution, with data processing and refinement statistics given in Table 4.1 above. The NO was supplied from PROLI NONOate, a pH-dependent donor compound (Saavedra *et al.*, 1996), which releases NO in response to the pH change that occurs when mixed with the crystal slurry, described in §2.6. This structure showed a difference density feature in the distal pocket into which NO was modelled, shown in Figure 4.10d, consistent with the NO-bound structure collected from a ferrous crystal at 100 K (PDB: 7zps), published in Adams *et al.* (2023). Also consistent with the cryogenic structure is the lack of a five-coordinate species formed by dissociation of the proximal histidine residue, which is seen in solution at low pH, revealed by UV-Vis, resonance Raman, and EPR spectroscopies (Adams *et al.*, 2023). The Fe-His bond length does not increase upon binding of NO either (decreases by 0.03 Å), also indicating-His bond length does not increase upon binding of NO either (decreases by 0.03 Å), also indicating no dissociation of the proximal His 120.

The two NO-bound crystal structures were compared to reveal any changes in binding associated with the temperature or redox state differences (see §1.6.1). Figure 4.7 is a superimposition of the structures, showing the whole dimer (a) and the haem sites in both chains (b–c). In the high-resolution cryogenic structure, solved to 1.56 Å resolution, two binding conformations can be seen for the NO in chain B (c), with equal occupancies. The first conformer is common across both chains, although with a 0.12 Å and 13° variation in Fe–N bond lengths and Fe–NO bond angles between chains, and the second conformation is unique to chain B. The bond length of the first conformer is similar to the Fe–N bond length in AxCP at 100 K (Kekilli *et al.*, 2014), although

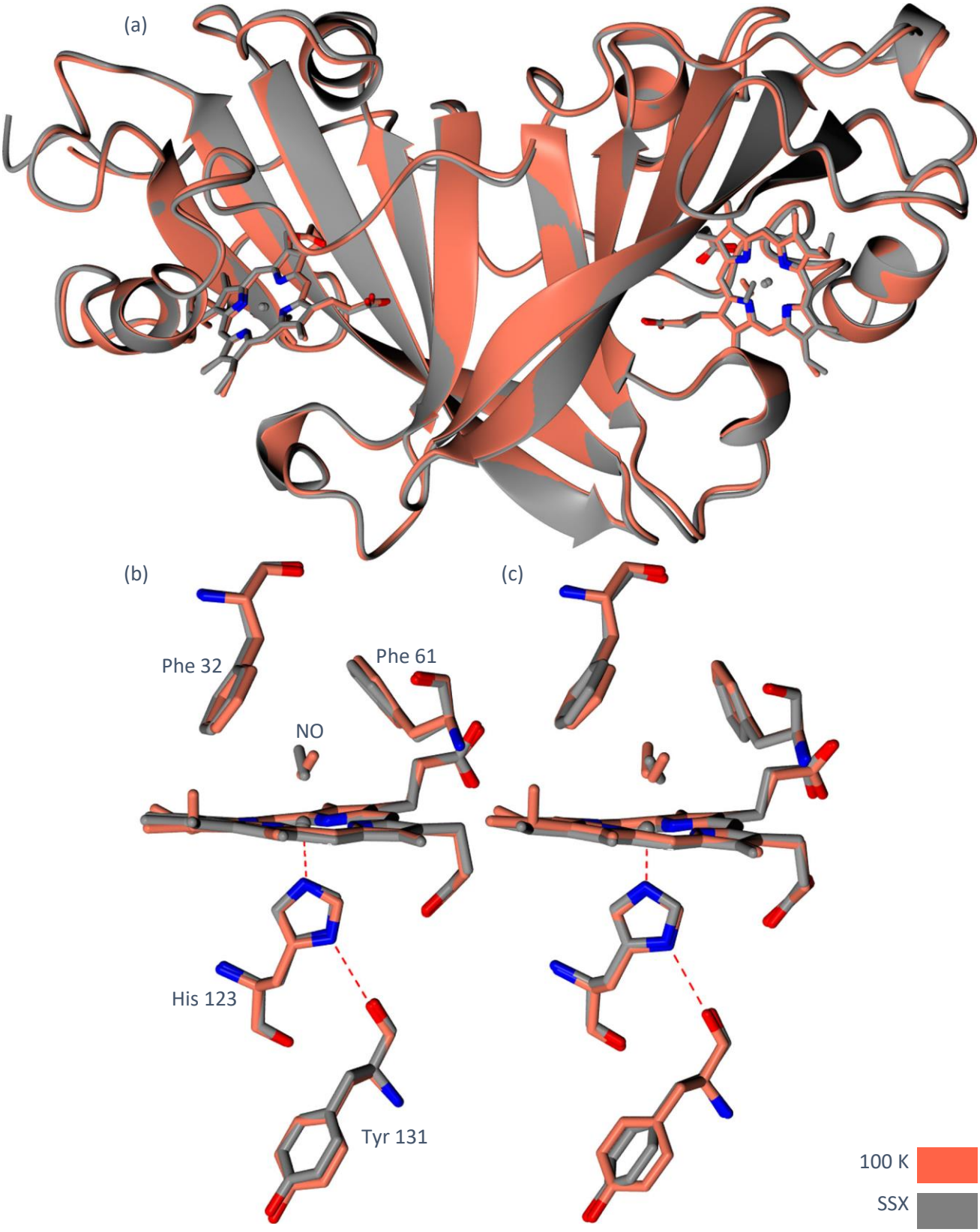

Figure 4.5: Superimposed structures of McCP-β bound to NO, collected at 100 K (red, PDB: 7zps), from Adams *et al.* (2019) and from TR-SSX at room temperature (grey). A ribbon view of the homodimer (a) is shown, and details of the haem sites of chain A (b), and chain B (c). Positioning of the NO within the distal pocket varies between structures and chains.

no room temperature structure of NO-bound AxCP exists to compare with the serial structures of McCP-β collected here. In the room temperature structure of McCP-β, no alternate conformers are seen for the NO. An Fe–N bond length of 1.99 Å is seen in both chains, matching the bond length of the conformer seen in only chain B of the 100 K structure. The Fe–NO bond angles for the structures vary -considerably between- chains and laser energies (listed in Table 4.2). Despite the variation between room temperature structures, the angle is consistently greater in chain A, averaging 144° across all room temperature structures, compared to 114° in chain B. Comparing the distal pockets of the two chains in Figure 4.8 shows significant differences in the positioning of both phenylalanine cap residues. The positioning of Phe 32 especially changes the available space above the NO, which is likely responsible for the differing NO angles between chains. As NO is coordinated to the haem iron by the nitrogen atom, its position within the binding site is more constrained than the oxygen, which lacks this coordination. This is demonstrated by the B-factor of the oxygen atom being higher in most (9 of 12) NO molecules in the six RT NO-bound structures collected, on average 6.81 Å² higher. This implies flexibility in the oxygen location in the unit cell, and hence the Fe–NO bond angle, so interpretation of the variation of the angles between the models must be done with caution. The low NO occupancy in some of the structures (as low as 30 %) means the electron density for the NO is weaker, also making the modelled bond angles less reliable. Additionally, the room temperature structure was collected from ferric crystals,

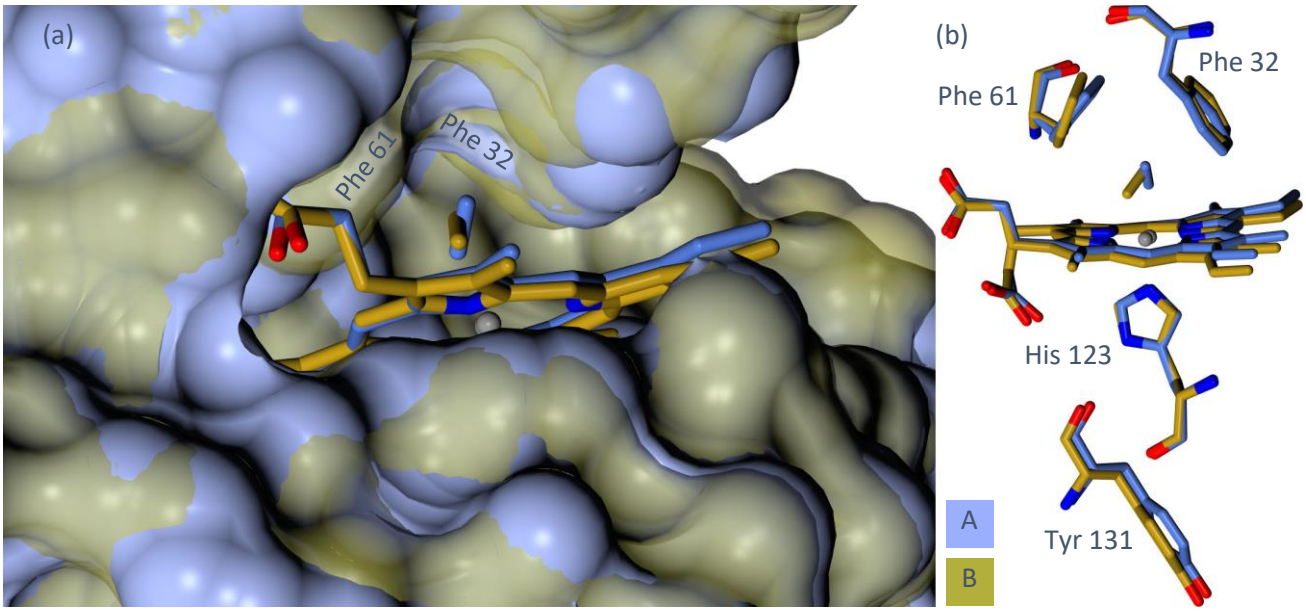

Figure 4.6: Comparison between chains A (blue) and B (gold) for an NO-bound McCP-β structure at room temperature. The surface view (a) shows changes in space occupied by the Phe residues, potentially causing the flatter positioning of the NO (smaller Fe–NO bond angle) in chain B.

rather than the ferrous redox state of the cryogenic structure, therefore any changes seen between the two structures could be due to either differing temperature or redox state. Furthermore, no validation of redox state was performed for the room temperature structure, meaning that the crystals which were ferric before data collection may have been partially reduced by the beam during the experiment. Besides the haem sites, no significant differences can be seen in the haem site or elsewhere in the structures (RMSD 0.29 Å).

	Fe–N bond length (Å)				Fe–NO bond angle (°)				B-factor (Å²)	
Laser energy (μJ)	A	B	Mean	SD	A	B	Mean	SD	N	O
1600	1.94	1.97	1.96	0.02	138	96	117	21	56.54	58.08
160	2.03	2.01	2.02	0.01	134	91	110	19	25.15	28.40
9.5	1.91	1.93	1.92	0.01	150	109	129	20	46.31	67.46
0.95	1.98	2.01	2.00	0.01	173	140	157	16	42.69	52.33
0.19	1.88	1.99	1.94	0.06	152	146	149	2.9	44.12	50.22
PROLI NO (RT)	1.99	1.99	1.99	0.00	120	104	112	8.1	25.83	25.01
Mean (all RT)	1.96	1.98	1.97	0.04	144	114	129	24	40.10	46.91
PROLI NO (100 K)	1.77	1.89 1.99	1.86	0.09	123	135 139	130	7.4	27.47	33.35

Table 4.2: Bond lengths and angles for NO bound in the haem sites of McCP-β, in structures with different laser energies, as well as structures soaked with PROLI NONOate collected at different temperatures. Mean values are shown for each structure, and for all room temperature structures. SD; standard deviation. The two values for chain B at 100 K represent multiple conformations. The mean B-factors for each atom type in the NO is also shown.

4.2.2 Laser-initiated NO binding

Time-resolved SSX experiments were performed at beamline I24, to investigate NO binding in McCP-β and to develop methodology for time-resolved SSX. These experiments were repeated on several occasions to iteratively develop and improve approaches for initiating the NO binding using the photocage and UV laser. The earliest time-resolved experiments with McCP-β involved collecting structures with two different laser intensities. The microcrystals were prepared under red light by soaking them in a solution of photocage dissolved in mother liquor, as described in §2.7. The photocage (Namiki *et al.*, 1997) was prepared at a concentration of 10 mg·mL^{-1}, which corresponds to a final concentration of 19 mM within the unit cell, using the crystal solvent content of 63.6 % (Kantardjieff and Rupp, 2003). This concentration was chosen to provide enough NO within the unit cell, while being economical with the supply of photocage, as explained in §3.5.

The time-resolved structures were collected at a 1.44 s timepoint, using the excite and visit again mode of data collection (§3.10). The first structure was collected with the 308 nm laser at full power, resulting in 1600 µJ of laser light at 308 nm reaching the sample position over 10 ms, operating pulsed at 50 kHz (3.2 µJ per pulse). The second structure was collected with the laser attenuated by 10 times, via reduction of the laser pulse rate down to 5 kHz. A dark structure was also collected as a negative control, where photocage was soaked into the crystals, but no laser illumination occurred. The electron density maps for the control structure, shown in Figure 4.9, feature no density for the NO in the distal pocket, like the resting state SFX structure. This indicates that release of the NO from the photocage is caused by the laser illumination, and not a result of accidental ambient light exposure or another form of chemical degradation. Figure 4.10 shows the results of the first laser-initiated binding experiments on I24. An NO-bound structure is superimposed with the resting state SFX structure (a), showing the outwards movement of the Phe 32 residue upon NO binding, and slight distortion of the haem, both consistent with the changes seen for NO binding at 100 K (Adams *et al.*, 2019; Adams *et al.*, 2023). The Cζ atom of the benzyl ring is displaced by 0.77 Å in chain A, and 0.81 Å in

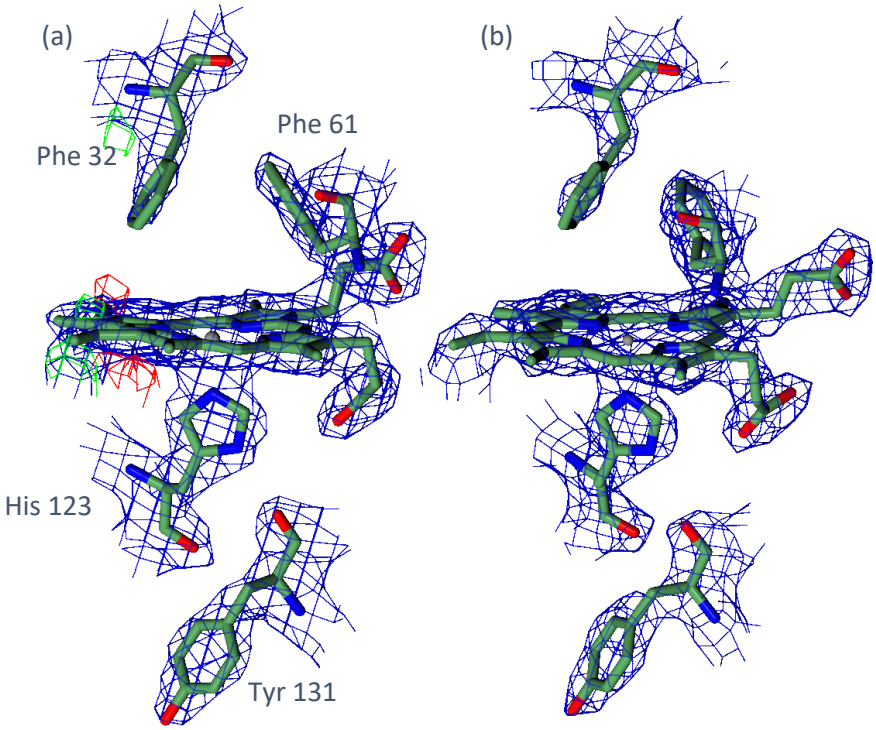

Figure 4.7: Electron density maps for the haem sites of a dark control structure. For this structure, NO photocage was soaked into the crystals, but no laser illumination occurred. Chains A and B are shown in (a) and (b). Density for the NO is not present in the distal haem pockets in this structure, which implies that the laser is responsible for the appearance of NO in the laser-initiated structures. $2F_{obs} - F_{calc}$ map contoured at 1σ, $F_{obs} - F_{calc}$ map contoured at 3σ.

chain B, compared with the room temperature resting state. The variation in NO position between structures is discussed above in §4.2.1. No density corresponding to the caging group or unlysed photocage is evident in any of the structures.

As well as providing room temperature structures of NO bound to McCP-β, the laser-initiation experiments were also an investigation into required laser illumination for photocage activation in the fixed-target chips, as part of the development of TR-SSX capabilities at I24. The laser illumination of the sample is measured in microjoules, and referred to here as laser energy, but is not to be confused with the photon energy of the UV laser, which at 308 nm is equal to 6.45×10^{-19} J. The time-resolved structures are shown in Figure 4.10, with highest laser energy (1600 μJ, blue) having electron density corresponding to 80 % ligand occupancy, and at the lower laser energy (160 μJ, purple) the occupancy is reduced to 40 %. As a comparison, a difference density peak corresponding to 100 % occupancy is shown in Figure 4.10d, taken from the PROLI NONOate soak structure, which had an excess of NO available (§4.2.1). These occupancies were estimated by manually editing the occupancy in the structures and comparing the B-factors of the ligand atoms with those of the adjacent haem atoms after a further round of refinement. The relative peaks in the $F_{obs} - F_{calc}$ omit maps were used to check the estimated values were accurate. These ways of estimating occupancy were used because of the unreliability of automated occupancy refinement available in Phenix. Further discussion of current methods of occupancy estimation is included in §5.3.1, and more robust occupancy estimation methods were investigated and used for later experiments detailed in §5.2.3–5.2.4.

The aim of the first time-resolved experiments with the laser at I24 was simply to achieve NO binding initiated using the NO photocage. Although this aim was achieved, the high laser energies (1600 μJ and 160 μJ) required to initiate NO binding were more than expected, and more than the approximately 40 μJ used for the previously published experiments with the photocage (Tosha *et al.*, 2017). Using the same calculations detailed in §3.5, but with the larger beam size of 60 μm, the laser energy means 3.8×10^{14} photons are absorbed by the crystal. This can be converted to values of approximately 1200 and 120 photons absorbed per photocage molecule. Although, as detailed in §3.7 and §3.9, not all expected photons were likely reaching

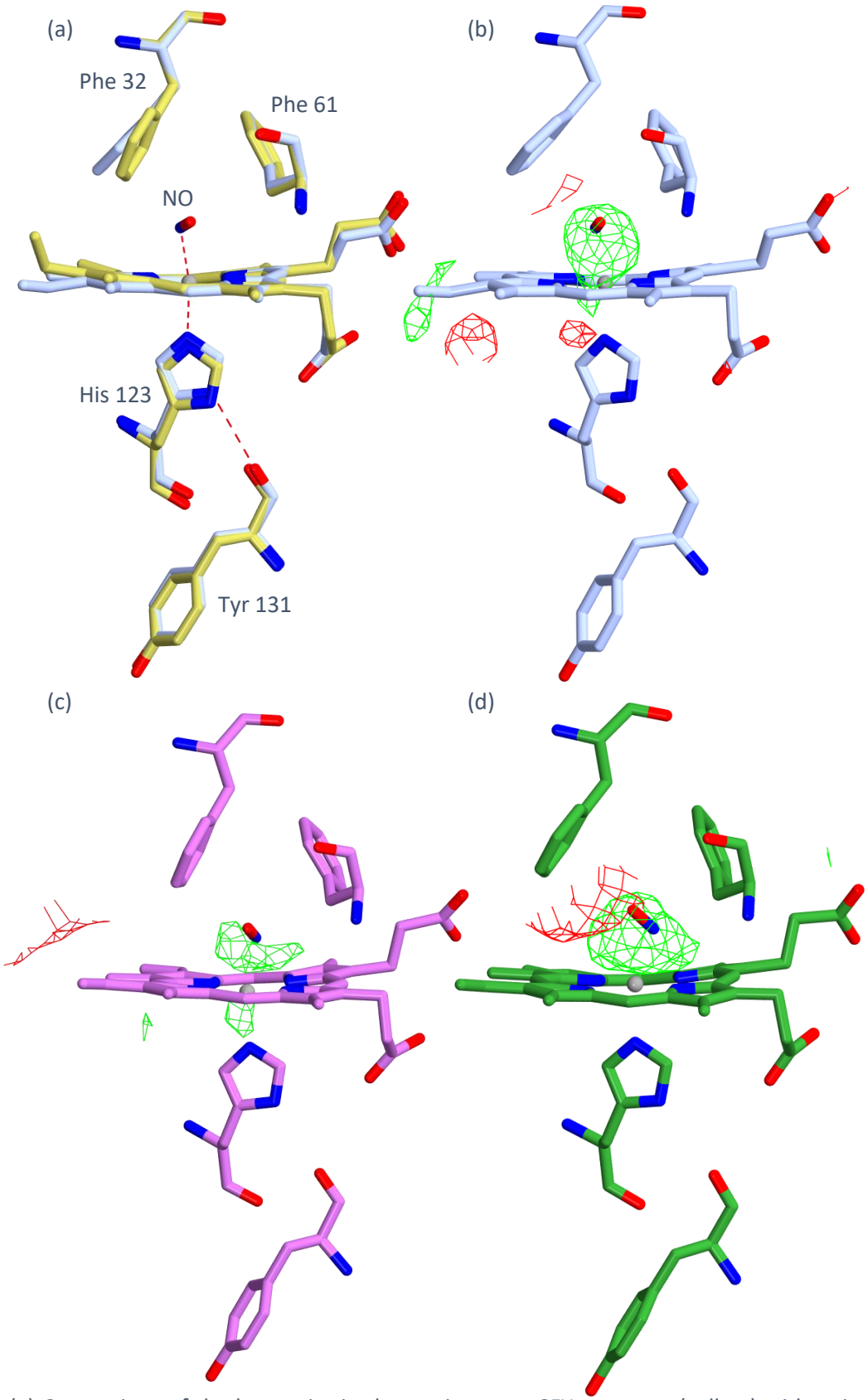

Figure 4.8: (a) Comparison of the haem site in the resting state SFX structure (yellow) with a time-resolved structure (2.00 Å resolution, blue) at 1.44 s after NO release. (b–d) Comparison of electron density for NO in omit maps (contoured at 3σ) for structures collected with (b) 1600 µJ (c) 160 µJ laser energy (2.00 Å), and (d) a PROLI NONOate soaked structure (1.95 Å). The NO occupancies in (b–d) are approximately 80 %, 40 %, and 100 %.

the crystal, this is far from the single photon excitation regime (Grünbein *et al.*, 2020). However, it is not outside the range of pump laser energies used for some other TR-SX experiments (Brändén and Neutze, 2021; Suga *et al.*, 2019). From these first experiments, areas of weakness in the methodology were identified, and developments were made to reduce the laser energy required for reaction initiation. The improvements are detailed in §3.7–3.9, with the results from further McCP-β experiments shown in §4.2.3 below.

4.2.3 Improvements in laser excitation

Measurements of the laser power reaching the sample position showed some discrepancy when reducing the pulse rate from its maximum of 50 kHz, in that the change in power was not proportional to the requested change in repetition rate. As it is impractical to measure the laser power after every change in repetition rate between samples, changes in laser power using repetition rate are unreliable. To combat this, attenuation by insertion of ND filters into the beam was used for in future experiments, described in §3.8.

As well as changing the method of attenuating the laser beam, changes were made to the sample holders to aid in illumination. In the first experiments with the NO photocage (the structures with 1600 µJ and 160 µJ laser energy), the chips were sealed using thin films of Mylar® on both sides of the holder, as is standard practice for chip preparation to prevent crystal dehydration. Unfortunately, Mylar® has a strong UV absorbance peak centred at 300 nm, so absorbs a significant amount of the laser energy delivered to the sample position, preventing it from reaching the photocage in the crystals. To reduce the laser intensity required, in future experiments, the Mylar® film on the upstream side of the chip -was replaced with EVAL™, an alternative film which lacks the strong 300 nm absorbance peak. Further detail on thin film choice is given in §3.9.

In addition to changing the film to seal the samples, the way the chips were mounted in the holders was also changed to aid laser excitation. The tapered aperture design for trapping crystals means that the crystals sit in wells on one face of the chip, easily accessible from one side, only accessible from the opposite side through the 7–30 µm square apertures (Mueller *et al.*, 2015). In the orientation used for resting state structures, the 50 µm laser beam on the upstream side of the chip could only illuminate the crystals through the much smaller

apertures, so only a fraction of the laser light reaches the crystal. To reduce the laser excitation from the 1600 µJ used in the first experiment, the chips were flipped over, so the laser illuminates the crystals more directly, described further in §3.7. In addition, as the laser light delivery hardware improved, the laser could be focused onto the sample position with a reduced spot size by fine positioning of the final focusing lens on a motorised stage, reducing beamsize from 60 µm to 50 µm. This means 44 % more energy from the laser reaches the crystal, rather than being wasted hitting the surrounding area on the edges of the well of the chip.

Using experience gained from collecting the first datasets, further experiments were conducted, involving the collection of three further time-resolved NO-bound structures with binding initiated by different laser energies. When measured immediately before the experiments, the laser power at the sample position was 1.2 mW, so a 16 ms laser exposure without additional attenuation, consisting of 80 pulses at 5 kHz, would impart 19 µJ into the sample. Three datasets were collected using ND filter combinations that attenuated the beam to provide 9.5 µJ, 0.95 µJ, and 0.19 µJ of laser energy at the sample position. These correspond to approximately 10, 1, and 0.2 photons absorbed per cage molecule, respectively. These structures are shown superimposed in Figure 4.11, coloured by laser energy, and are consistent (RMSD 0.28). The only significant difference besides NO occupancy is displacement of the Phe 32 capping residue further from the haem in the structure with the higher NO occupancy (9.5 µJ). The Cζ atom of the benzyl group is displaced by 0.88 Å and 0.65 Å in chains A and B respectively. This is consistent with previous NO-bound structures, including the laser-initiated structure superimposed with the resting state structure in Figure 4.10a above, which shows Phe 32 displacement upon binding. The processing and refinement statistics for the structures with improved laser initiation are shown in Table 4.3.

The three laser-initiated structures were analysed for NO binding, including estimation of the ligand occupancies and comparing any differences in active site geometry between room temperature and 100 K. As with the previous experiments, there was near 100 % occupancy at the higher laser energy used, although this was much less than the energies used for the first experiments, for the reasons described above. Despite the

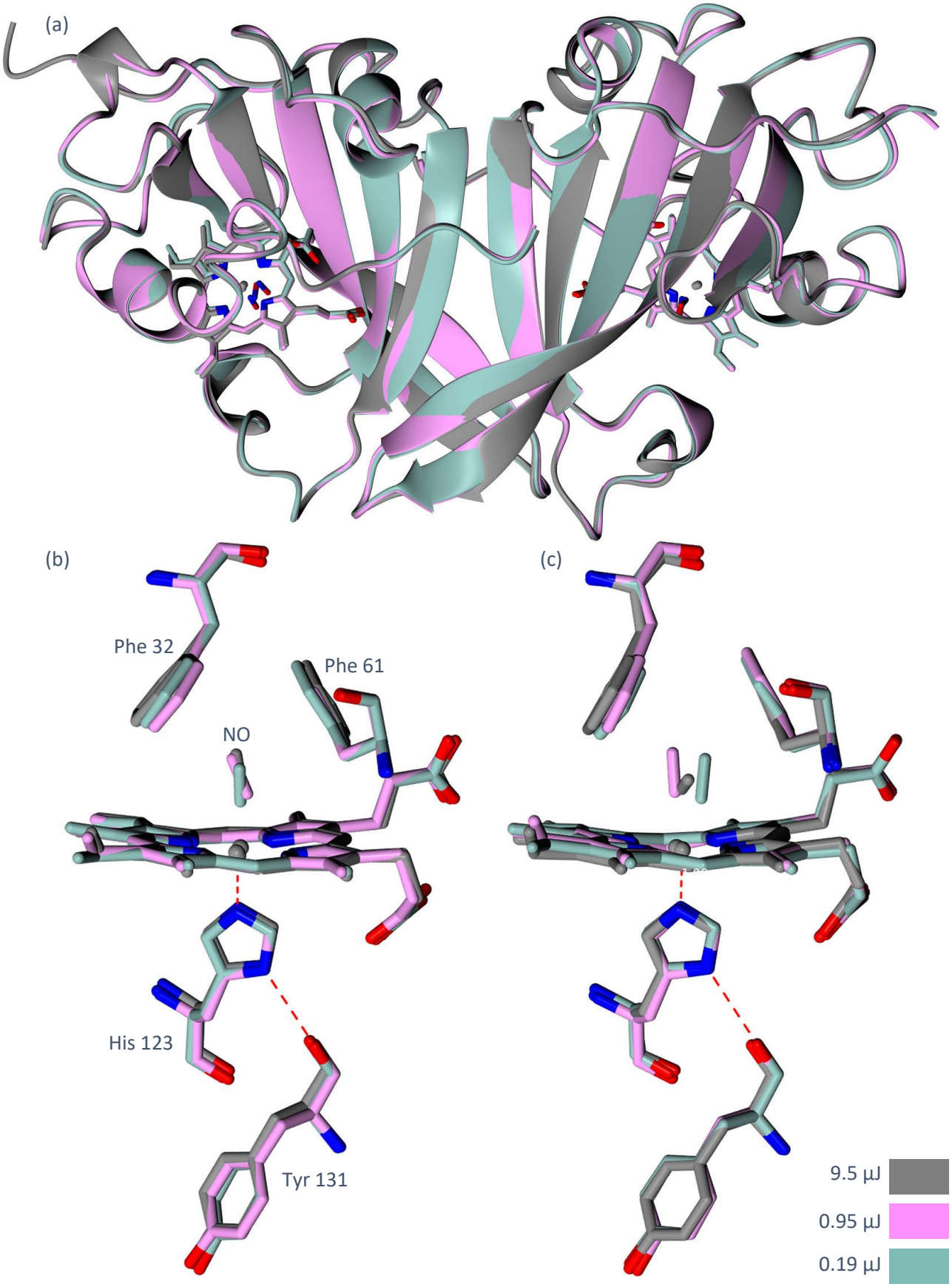

Figure 4.9: Superposition of three time-resolved NO-bound structures of McCP-β, collected with laser powers of 9.5 µJ (grey), 0.95 µJ (pink), and 0.19 µJ (green). In chain B (c), the structure with highest laser power shows the NO in the alternative conformation as seen in the 100 K structure (Adams *et al.*, 2023).

improvements in technique, at the two lower laser energies, separated by a factor of five, the estimated NO occupancies do not follow the expected trend of higher laser energy equating to higher occupancy. Occupancy values are listed in Table 4.4.

Dataset	9.5 µJ	0.95 µJ	0.19 µJ
Resolution (Å)	43.81–2.35 (2.39–2.35)	43.84–2.15 (2.19–2.15)	43.83–2.05 (2.09–2.05)
Space group	$P2_13$	$P2_13$	$P2_13$
Unit cell (Å), $a = b = c =$	107.3	107.4	107.4
Merged crystals	9778	7713	17 766
Unique reflections	17 432 (874)	22 724 (1123)	26 144 (1296)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Multiplicity	96.1 (24.3)	77.99 (22.26)	144.54 (36.06)
$CC_{1/2}$	0.962 (54.92)	0.952 (0.644)	0.973 (0.597)
R_{split}	0.171 (0.515)	0.193 (0.487)	0.125 (0.453)
Mean I	25.1 (5.1)	32.3 (6.4)	45.5 (5.4)
I/σI	1.56 (0.53)	1.65 (0.61)	2.27 (0.49)
N_{obs}	1 327 272 (21 195)	1 772 245 (24998)	3 778 854 (46 734)
R_{work}	0.178	0.190	0.173
R_{free}	0.248	0.245	0.221
RMSD bond length (Å)	0.0136	0.0155	0.0158
RMSD bond angles (°)	2.479	2.582	2.407
Ramachandran favoured (%)	91.85	95.17	94.42

Table 4.3: Data processing and refinement statistics for time-resolved McCP-β structures with an NO photocage. Values in parentheses refer to the outermost resolution shell. NO occupancies shown as by manual estimation.

For these structures, the occupancies were estimated using two different methods based on real space B-factors and electron density peak heights which are explained in detail in sections §5.2.3 and §5.2.4 respectively. Using the difference map peak heights to estimate occupancy, values 39 %, and 41 % were obtained for the two lower laser energies. By comparing B-factors for the ligand nitrogen atom with those of the adjacent haem iron, while changing the occupancy of the NO in the range 0–100 %, occupancy values of 39 % and 63 % were obtained. The B-factor differences from the output from the occupancy refinement script are displayed in Figure 4.12, with the interpretation of this type of plot explained in §5.2.3. The arrows indicate the point where the B-factor difference is zero, which is determined to be the true occupancy value. In Figure 4.13, results from both occupancy estimation methods are plotted against the laser energy. A general agreement is seen between the methods, with the script that compares B-factors providing

consistently slightly higher results. Tentative lines of best fit are drawn, although having only three datapoints severely limits any conclusions that can be drawn, especially with the poor correlation of the points. Subsequent to the initial submission of this thesis, further analysis has been performed, based on an approach described by Barends et al. (2024). This analysis is explained in §5.3.1, and uses the $F_{obs} - F_{calc}$ difference map after refinement to determine occupancy. The results of this analysis on the McCP-β structures are shown as red points labelled 'refined density' in Figure 4.13, and show good agreement with the two methods used originally.

	NO occupancy (%)											
	Peak height comparison				Occupancy refinement script				Refined difference densities			
Laser energy (µJ)	A	B	Mean	SD	A	B	Mean	SD	A	B	Mean	SD
9.5	68	100	84	16	100	100	100	0	100	100	100	0
0.95	27	50	39	11	48	28	39	11	38	50	44	6
0.19	42	40	41	1	68	58	63	5	51	35	43	8

Table 4.4: NO occupancy as a percentage, for time-resolved McCP-β structures with binding initiated by with different laser energies. Values are shown for occupancy estimation by three methods, as described in §5.2.3–5.2.4, and §5.3.1.

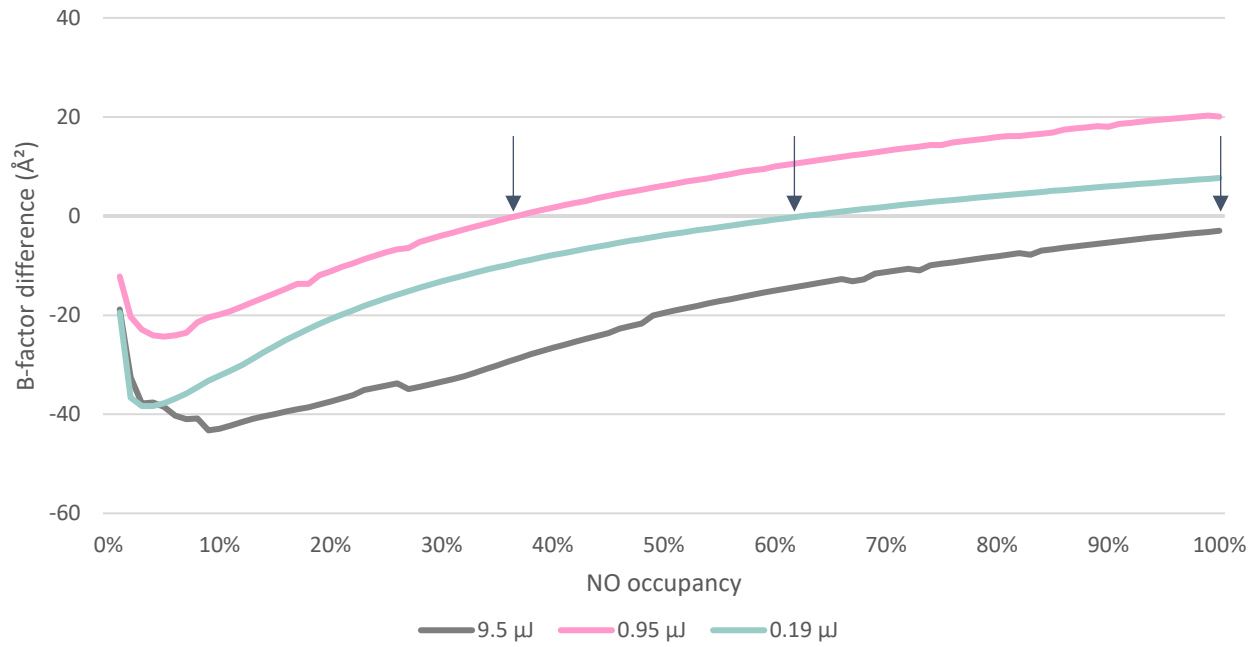

Figure 4.10: A graph of the difference in B-factors between the NO nitrogen and the adjacent haem iron, plotted against NO occupancy. Positive values represent the nitrogen B-factor being higher than the iron. The points at which the curve crosses the X axis, marked with arrows, represent the occupancy values, as described in §5.2.3. The data represents three different laser excitations and shares a colour scheme with Figure 4.11.

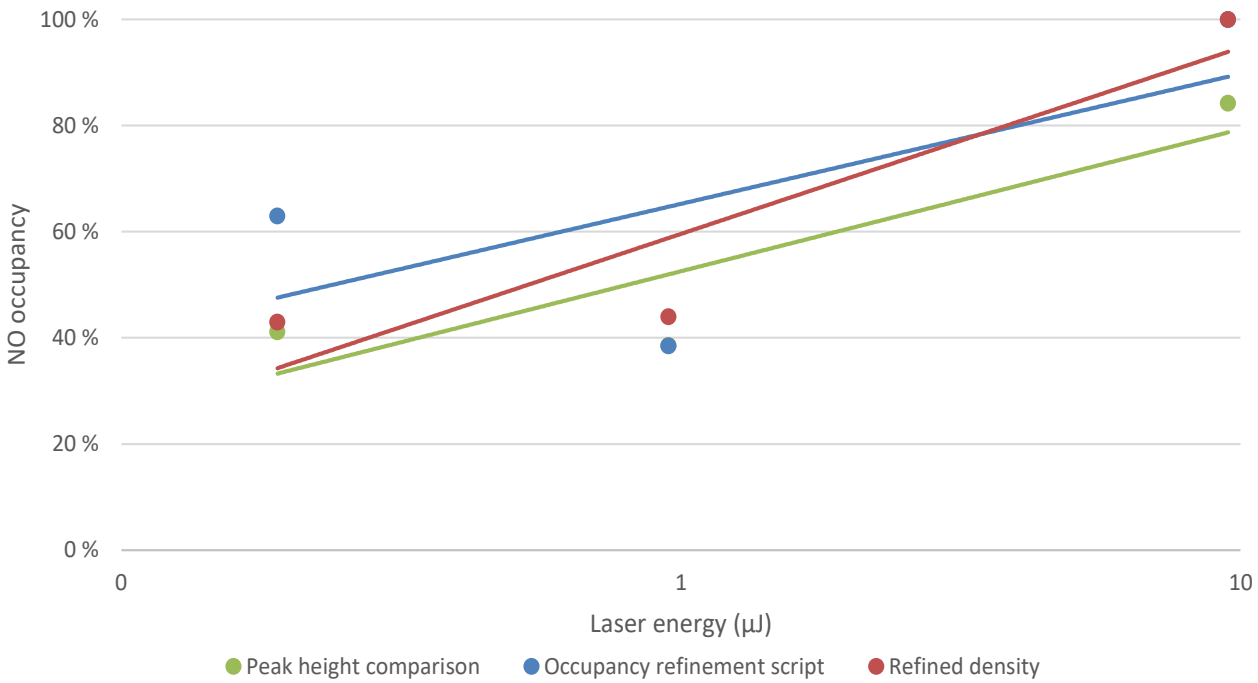

Figure 4.11: A graph of estimated NO occupancies plotted against laser power for three structures, using manual examination of omit maps and B-factors, an automated script that compares B-factors of refined structures, and magnitude of $F_{obs} - F_{calc}$ difference maps after refinement.

4.2.4 Room temperature structures of AxCP

AxCP was expressed and purified as described previously (Barbieri *et al.*, 2008), detailed in §2.3. Previous crystallisation of the protein was in space group $P6_522$, producing only large crystals 200–1000 µm in length (Kekilli, 2015; Antonyuk *et al.*, 2011), which are unsuitable for the serial crystallography techniques used here (see §1.5). AxCP was crystallised in a novel morphology, in space group I422 (§2.4). Microcrystals measuring approximately 15 µm in each dimension grew within 24 h, shown in Figure 2.2b. Two serial room temperature structures of AxCP have been collected, using a fixed-target chip at I24, and a zero-dose SFX structure using a viscous extruder at SACLA (Japan), with methods described in §2.6 and §2.8. These have been collected to 1.75 Å and 2.00 Å resolution respectively and can be used to evaluate any differences between data collection methods, although the differences between the sample delivery methods (chip versus extruder) are convoluted with the differences due to the X-ray source (synchrotron versus XFEL). Data processing and refinement statistics are shown in Table 4.5. The fact that these are the first serial and room temperature structures of AxCP is important, as temperature has been shown to affect ligand binding conformations in AxCP (Kekilli *et al.*, 2014), and X-ray radiation can photoreduce the iron in the haem centre.

Figure 4.14 is a superposition of three AxCP structures, including the SSX (coral) and SFX (blue) structures, as well as a 100 K ferric structure, validated with online spectroscopy (yellow, PDB: 4cda), published in (Kekilli *et al.*, 2014). Minor differences are seen between the helices in the structures, and in the loops between helices. The serial structures show close similarity to the 100 K structure, with RMSD values of 0.73 Å and 0.77 Å for the SSX and SFX structures compared to the 100 K structure. The two serial structures have an RMSD of 0.18 Å, so more closely match each other than the 100 K structure. The arginine 124 residue close to the haem, which indicates the redox state of the iron, is in the same conformation in all structures, sitting parallel to the plane of the haem to indicate ferric iron.

Dataset	SSX	SFX
Resolution (Å)	50.95–1.75 (1.63–1.75)	25.44–2.00 (2.03–2.00)
Space group	I422	I422
Unit cell (Å), *a* = *b* , *c*	72.0, 94.7	71.9, 94.6
Merged crystals	12 513	4163
Unique reflections	16 637 (778)	8711 (690)
Completeness (%)	99.8 (95.9)	100.0 (100.0)
Multiplicity	133.2 (68.2)	129.73 (73.70)
$CC_{1/2}$	0.999 (0.819)	0.960 (0.772)
R_{split}	0.041 (0.355)	0.246 (0.475)
Mean I	408.7 (18.6)	435.9 (75.2)
I/σI	48.4 (5.1)	2.22 (0.59)
N_{obs}	2 216 835 (53 096)	788 694 (30 498)
R_{work}	0.247	0.253
R_{free}	0.270	0.311
RMSD bond length (Å)	0.0120	0.0111
RMSD bond angles (°)	2.537	2.353
Ramachandran favoured (%)	97.58	98.39

Table 4.5: Data processing and refinement statistics for AxCP structures. Values in parentheses refer to the outermost resolution shell.

4.3 Discussion

4.3.1 NO binding to McCP-β

These experiments have provided the first room temperature structures of McCP-β, including a zero-dose serial structure collected at the SACLA XFEL, and low dose synchrotron structures from beamline I24. Differences between the structure of the protein at room temperature and 100 K are limited to movement of a few sidechains, including Phe 32 that forms part of the hydrophobic distal pocket. Room temperature NO-

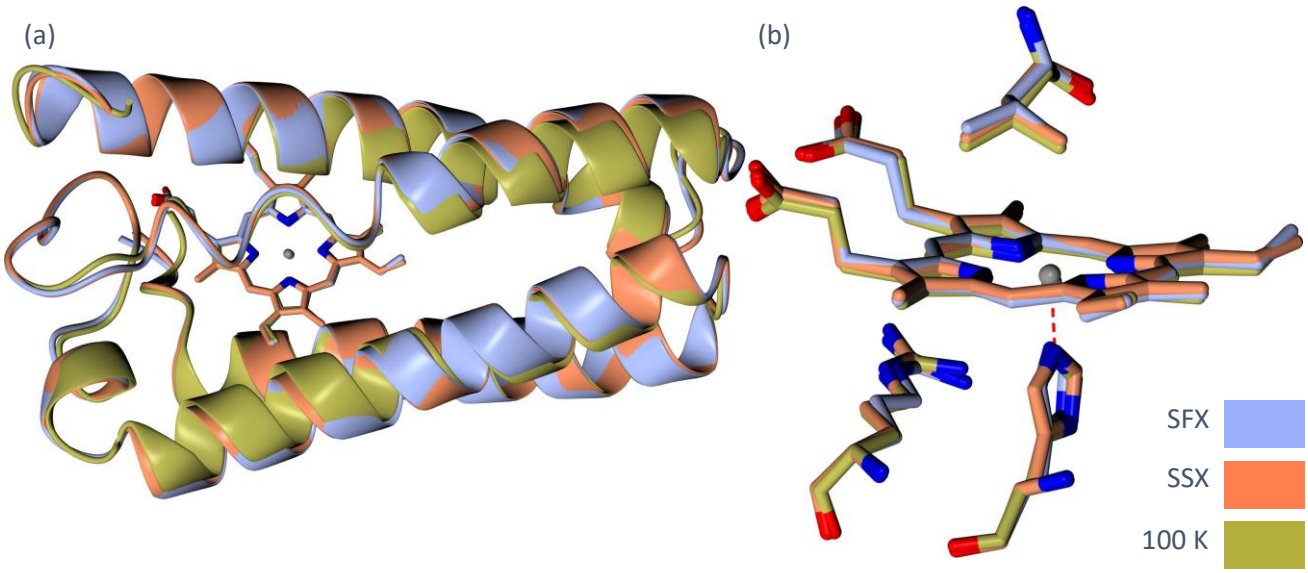

Figure 4.12: Superposition of three AxCP structures. (a) Ribbon view of chain A, showing that any differences between the structures are confined to the loops and terminal regions. (b) A close view of the haem site, with. Yellow: Published 100 K ferric structure (PDB: 4cda) from Kekilli *et al.* (2014). Blue: SFX; coral: SSX.

bound structures reveal a 6c distal species, with NO bound in the distal pocket. Further careful optimisation of crystallisation conditions, soaking parameters and data collection could obtain higher resolution structures with increased confidence in ligand conformations. To deconvolute the cause for the different NO conformations seen in the room temperature structures, and identify temperature or redox state as the cause, a repeat of the PROLI NONOate SSX data collection could be performed, but by first reducing the crystals with ascorbic acid in the same manner as the 100 K structure (Adams *et al.*, 2023). As part of collecting a comprehensive set of McCP-β structures from various conditions, a cryogenic structure of NO-bound ferric McCP-β could be collected with online UV-Vis spectroscopy to confirm the redox state in the crystal(s), ruling out photoreduction during data collection (Kekilli *et al.*, 2017a; Owen *et al.*, 2011). This has been performed previously for AxCP at cryogenic temperatures (Kekilli *et al.*, 2017b).

The time-resolved structures of McCP-β binding NO show that time-resolved SSX in the fixed-target chip system on beamline I24, including using light initiation with the NO photocage. The first successful experiment used 1600 µJ of UV laser energy to initiate binding, the most energy that could be delivered in the 10 ms exposure time. Through iterative development of the methodology, a several hundred-fold reduction in required laser energy has been demonstrated, down to 9.5 µJ for a full or near-full occupancy structure, and approximately half occupancy at 0.19 µJ. This is a large improvement of four orders of magnitude in required

laser energy, greatly reducing the chances of adverse effects caused by the laser (Moffat, 1998). The data collected at different laser energies provides information about required excitation energy to lyse the photocage in the fixed-target setup. The aim of laser energy optimisation was to provide single photon excitation to the NO photocage within the crystals. Using the calculations detailed in §3.5, the structure with 0.95 µJ laser energy provides approximately one photon absorbed per chromophore (photocage molecule), so single photon absorption has been reached. The concentration of the photocage in the crystals provides approximately the number of photons required for full NO binding in the crystal. NO is not available in excess, especially considering the quantum yield of 1.4, rather than the theoretical two NO molecules per cage.

Further experiments with increased NO concentration could provide NO in excess within the crystals, and lead to structures with confident 100 % occupancy. With increased photocage concentration, care must be taken with uneven light intensity. As the laser beam passes through a crystal, its intensity is attenuated, therefore providing uneven illumination from front to back (Grünbein *et al.*, 2020). With the 16.2 % transmission through the crystals with 10 mg·mL^{-1} photocage, the beam is already six times brighter at the front of the crystal compared to the rear, and increased absorbance by a more concentrated chromophore would exacerbate the gradient in light intensity. With this gradient, although the average number of photons per photocage may be in the single photon regime, the molecules seeing the most intense beam could exhibit multiphoton absorption. Another issue to consider is the solubility of the photocage. Anecdotal evidence suggests the NO photocage dissolves easily to 10 mg·mL^{-1} in the PEG concentrations used for crystallisation of McCP-β (§2.2) and DtpB (Lučić *et al.*, 2020b), although is not very soluble in ammonium sulfate-based mother liquors. There is potential for exchanging grown crystals into solutions more amenable to high photocage concentrations, but this is another condition to optimise, and an additional barrier to time-resolved crystallography.

Another consideration required with these experiments is the possibility of photolysis of the Fe–NO bond by the laser. Intense light flashes can be used to trigger photolysis of bonds between the haem iron and its axial ligands, as extensively demonstrated with the model system of CO bound to myoglobin (Barends *et al.*, 2015; Barends *et al.*, 2024; Schotte *et al.*, 2004). Flash photolysis of NO has been demonstrated with other haem

proteins in solution, including mitochondrial cytochrome *c* (Silkstone *et al.*, 2012), but no crystallographic studies have yet been performed. Photolysis is possible due to the long laser exposure time (multiple pulses over 5 ms), allowing NO to be produced, bind to the haem, and then be photodissociated. This is unlikely to affect the measured data, since both the photolysis and relaxation back to the liganded state are on the order of nanoseconds, so NO will be fully rebound within the 1.44 s delay time before diffraction data collection. Additionally, the optimal wavelength for flash photolysis is likely to be greater than 400 nm, far from the 308 nm used here. Further work with McCP-β could involve intentional flash photolysis of the bound NO and capturing the structure of the dissociated state. This would require a very short timepoint, and very short exposure times, currently out of reach of I24 due to limited flux (Owen *et al.*, 2012) and count rate limits of the detectors (Donath *et al.*, 2023; Loeliger *et al.*, 2012). The experiments could be performed at an XFEL with TR-SFX, employing the fixed-target sample delivery system used for the data in this thesis. NO could be added into crystals (in the form of PROLI NONOate) prior to loading to form the Fe–NO state. A laser pulse could be used to trigger photolysis picoseconds before a femtosecond X-ray pulse used to collect diffraction data.

Further work on reaction initiation could investigate the effects of laser exposure time on the samples, while changing flux to keep the total energy the same. This could be of interest in how it affects laser heating of the samples (Grünbein *et al.*, 2020; Miller *et al.*, 2020; Moffat, 1998). Repeated high-quality datasets could potentially show effects of high temperatures in the crystals on the conformations of sidechains in the protein. McCP-β has provided a suitable test case for developing the method, as it provides robust and relatively plentiful microcrystals.

4.3.2 AxCP as a future TR-SSX target

For the AxCP structures collected, the positioning of the Arg 124 residue on the proximal side of the haem can be used to indicate the redox state of the iron. In the reduced, ferrous state, the arginine reorients perpendicular to the haem, which is not seen here. The ferric state is expected in the SFX structure (2.0 Å resolution), as data collection from the 10 fs X-ray pulses occurs on a shorter timescale than the photoreduction and subsequent reorientation of the arginine. In the SSX structure (1.75 Å), the arginine

positioning also indicates the ferric redox state. AxCP can also be used as a model system for further development of time-resolved methods at I24 as well-diffracting microcrystals can be made, and the binding has been well-characterised in solution (Kekilli *et al.*, 2017b; Kekilli, 2015; Barbieri *et al.*, 2008). When moving beyond capturing a single endpoint, as has been demonstrated with NO binding in McCP-β and DtpB (see §5), AxCP is a logical next step, as the information gained on laser energy for triggering photolysis of the NO photocage is directly applicable, and would involve minimal changes to data collection. Short timepoints of AxCP could hopefully capture intermediate steps in the distal to proximal conversion at the haem site, as described in §1.7.1. Very short timepoints, less than one millisecond, would need to be collected with SFX, as the longer X-ray exposure times required when using synchrotron radiation would be unable to capture the shortest-lived transient states in a reaction, which may only be present for tens of microseconds. Longer exposure times required at synchrotrons would result in a mixture of states being captured, which would be very difficult to interpret (Mehrabi and Schulz, 2023). Having structural information for the intermediate steps involved in NO binding of AxCP would be valuable in elucidating the function of the protein and its role within *Alcaligenes xylosoxidans* and the wider nitrogen cycle (Kim and Shoda, 1999; Bordeleau and Bartha, 1972; Hargreaves *et al.*, 1992). The proximal to distal conversion upon NO binding is also seen in mammalian mitochondrial cyt *c*, where NO is involved in apoptosis (Silkstone *et al.*, 2010), so structural studies of this mechanism can lead to better understanding of human biology.

5. Nitric oxide binding to DtpB

5.1 Introduction

The dye-decolourising peroxidase (DyP) family is named due to their ability to degrade synthetic dyes (Lučić *et al.*, 2020b; Rožman, 2022), and is described in §1.7.2. The peroxidase activity of the type B DyP from the bacterium *Streptomyces lividans* (DtpB) has been studied, including solving the resting state ferric structure via SFX (see §1.6.2), and a structure of Compound I, generated following addition of hydrogen peroxide to the haem state to form an Fe^{4+}=O haem species carrying a porphyrin cation radical (Lučić *et al.*, 2020b; Rožman, 2022). Importantly for developing time-resolved serial crystallography methods using a nitric oxide photocage, as well as the peroxidase activity, the ferric protein also binds NO. This is illustrated in the UV-Vis absorbance spectra in Figure 5.1, provided by Dr Marina Rožman, which were collected from protein in solution at 20 µM with 20 mM phosphate pH 7, 150 mM NaCl, at 20 °C, using a Cary 60 spectrophotometer (Agilent). The resting state ferric spectrum is shown in blue. Upon addition of 10 µL PROLI NONOate, a pH-dependent NO donor compound (Saavedra *et al.*, 1996), the spectrum (red) features a redshift of the Soret peak from 401 nm to 419 nm, and the Q band splits into multiple peaks. These changes, which are similar to those in McCP-β (Adams

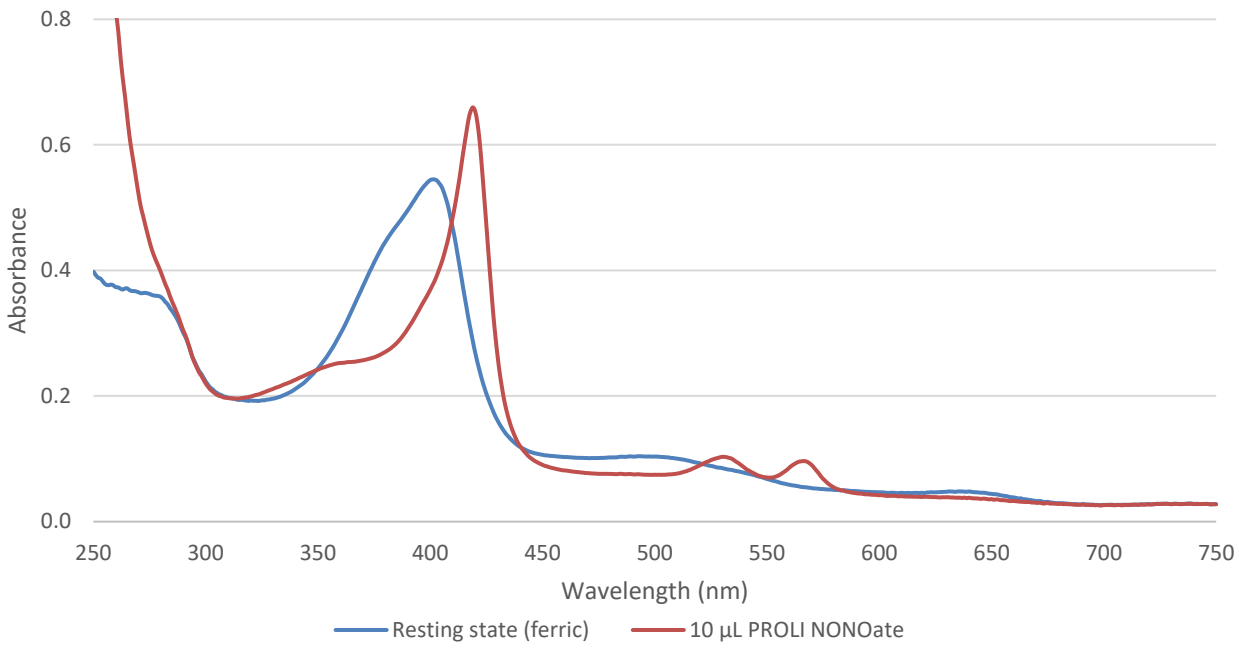

Figure 5.1: Solution UV-Vis absorption spectra, provided by Dr Marina Rožman, for DtpB in its ferric resting state (blue) and after the addition of 10 µL PROLI NONOate (red). Note the redshift of the Soret peak from 401 nm to 419 nm, and the splitting of the Q band into multiple peaks.

et al., 2023), indicate NO binding to the haem. Unusually for a DyP, the distal pocket in the resting state is dry, that is free from water (Lučić *et al.*, 2020b), so the presence of an NO ligand bound to the haem would be easy to identify in an electron density map. The protein is also hexameric, so provides six haem sites per structure in which to analyse the modes of NO binding.

5.2 Results

Using the methodology described in §3, time-resolved SSX structures of DtpB with NO bound have been collected to 2.40 Å resolution at beamline I24, Diamond Light Source, using crystals provided by Dr Marina Rožman. DtpB structures were solved by molecular replacement, using for a search model a previously published (Lučić *et al.*, 2020b) SFX structure (PDB: 6yrj), which was collected to 1.85 Å resolution at SACLA using the same fixed target chip system (Sherrell *et al.*, 2015). A resting state SSX structure was also collected, limited to 2.75 Å resolution due to the quality of the batch of crystals used. Data processing and refinement

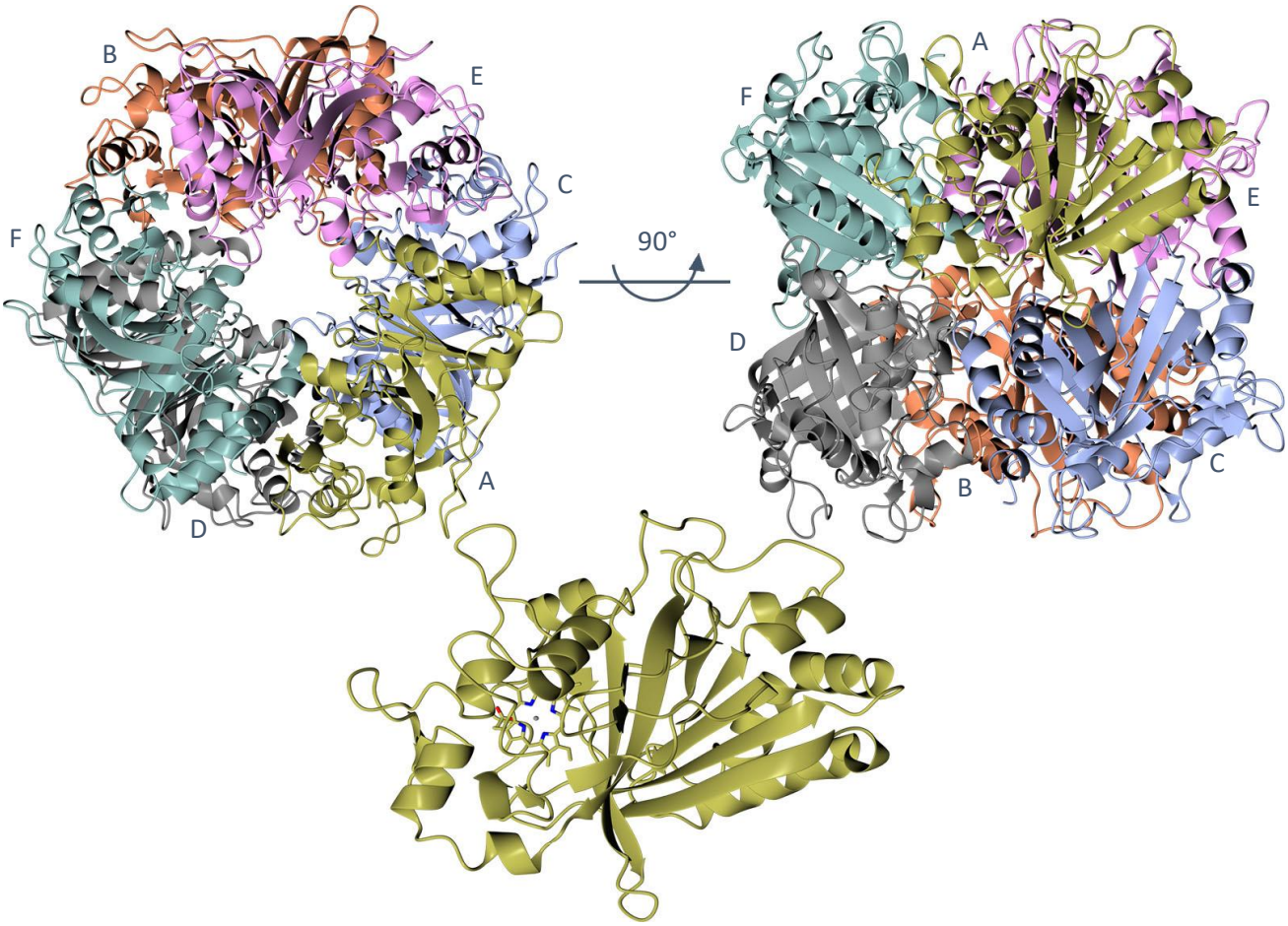

Figure 5.2: Ribbon view of a time-resolved SSX structure of DtpB illustrating its hexameric structure. Chains A–F are shown in yellow, coral, blue, grey, pink, and green respectively. The structure was collected with 32 µJ of laser energy. Chain A is also shown separately.

statistics for the structures are shown in Table 5.1. A ribbon view of a typical TR-SSX structure with NO bound is shown in Figure 5.2, demonstrating the homohexamer in the asymmetric unit. Figure 5.3 is a superposition of chain A from the published resting state SFX structure (blue, 1.85 Å resolution), the resting state SSX structure (coral, 2.75 Å), and a TR-SSX structure with 32 µJ laser energy (yellow, 2.4 Å). NO is confirmed to bind in the distal haem pocket, forming a six-coordinate iron species coordinated via the nitrogen atom, shown in Figure 5.3b and for all chains in Figure 5.5. The largest difference in the haem sites between structures is Arg 243 being displaced away from the distal pocket upon NO binding. The arginine residue has increased B-factor, indicating its flexibility to allow the NO to bind. Figure 5.3a shows that differences outside the haem sites are limited to the flexible loops and terminal regions.

Examining B-factors in Figure 5.4 shows that the loop regions, where differences between structures can be seen, are more disordered (blue-green) than the rest of the protein, indicating the differences between structures in the loops are a result of disorder, not of significant conformational changes. Quantitative comparison of the structures was performed using secondary-structure matching (Krissinel and Henrick, 2004), comparing each chain individually. The average RMSD values for all combinations of the structures were

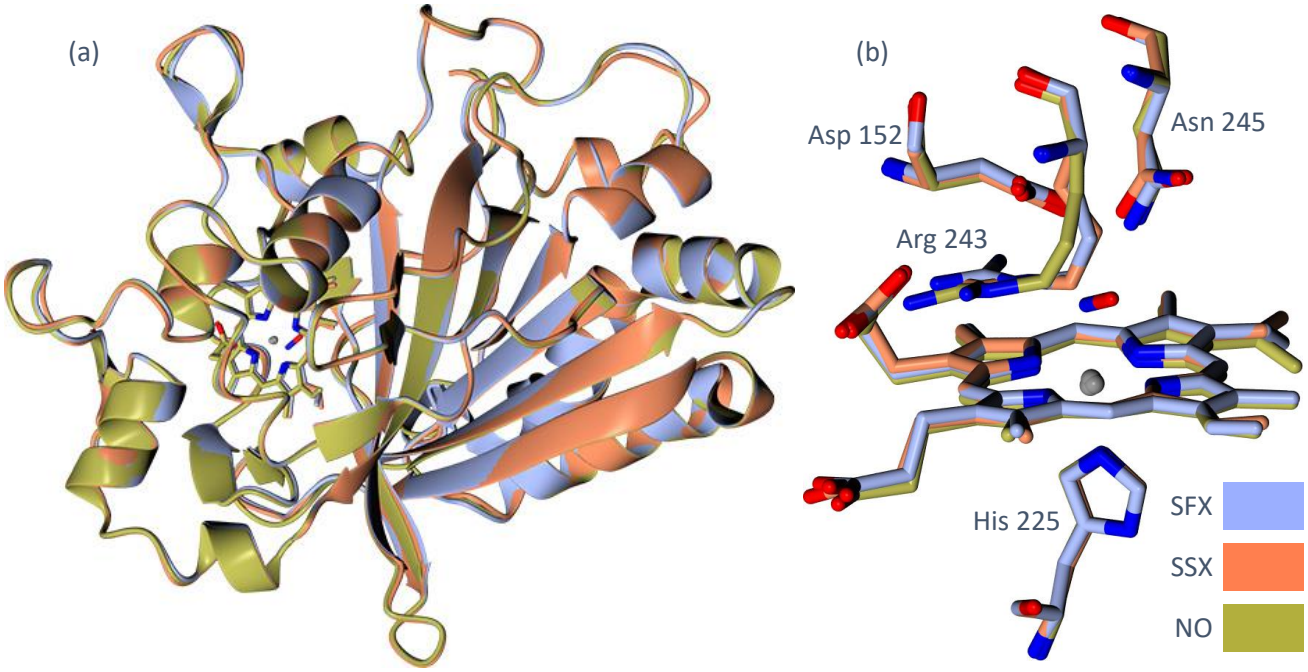

Figure 5.3: Superposition of three DtpB structures. (a) Ribbon view of chain A, showing that any differences between the structures are confined to the loops and terminal regions. (b) A close view of the haem site from that chain, with displacement of Arg 243 visible upon NO binding. Blue: SFX (PDB: 6yrj); coral: resting state SSX; yellow: NO-bound TR-SSX with 32 µJ laser energy.

similar (0.33 Å comparing a single chain between structures), showing there are no obvious differences between the X-ray source used (XFEL or synchrotron) or the presence of NO. This implies that the low dose SSX structures do not feature X-ray damage that can be identified in the structure at this resolution, although photoreduction of the haem iron would not be visible in the electron density (see §1.6.2).

Dataset	64.4 µJ	32.2 µJ	16.1 µJ*	8.05 µJ*	0.81 µJ	Resting state
Resolution (Å)	79.75–2.40 (2.44–2.40)	76.87–2.40 (2.44–2.40)	79.71–2.40 (2.44–2.40)	79.66–2.40 (2.44–2.40)	79.71–2.40 (2.44–2.40)	103.66-2.75 (2.80–2.75)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell (Å), *a*, *b*, *c*	87.2, 123.2, 195.5	86.7, 121.4, 198.2	87.1, 123.0, 195.4	87.1, 122.9, 195.0	8731, 122.9, 195.6	86.4, 121.4, 199.3
Merged crystals	6870	5536	3612	3109	5003	5895
Unique reflections	82 992 (4087)	82 434 (4038)	82 823 (4084)	82 535 (4074)	82 827 (4089)	55 276 (2716)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Multiplicity	39.3 (30.2)	34.1 (25.3)	23.3 (17.4)	20.6 (15.2)	28.7 (22.1)	22.2 (19.2)
$CC_{1/2}$	0.976 (0.346)	0.968 (0.585)	0.958 (0.311)	0.958 (0.379)	0.971 (0.343)	0.916 (0.297)
R_{split}	0.214 (1.332)	0.170 (0.589)	0.234 (1.157)	0.197 (0.819)	0.192 (0.878)	0.303 (1.018)
Mean I	27.6 (3.3)	60.0 (9.5)	37.7 (5.7)	73.2 (10.9)	40.7 (6.2)	36.4 (7.9)
I/σI	6.4 (0.9)	12.0 (2.4)	7.7 (1.4)	10.1 (1.9)	8.2 (1.5)	5.1 (1.3)
N_{obs}	3 257 619 (123 533)	2 814 455 (102 057)	1 926 745 (70 886)	1 701 745 (62 075)	2 374 546 (90 252)	1 225 800 (52 176)
R_{work}	0.159	0.147	0.155	0.146	0.147	0.179
R_{free}	0.236	0.225	0.233	0.222	0.224	0.260
RMSD bond lengths (Å)	0.0149	0.0152	0.0150	0.0158	0.0147	0.0167
RMSD bond angles (°)	2.413	2.309	2.383	2.443	2.282	2.591
Ramachandran favoured (%)	95.12	94.85	95.56	96.16	95.83	92.38
NO occupancy from B-factors (%)	84	58	65	50	39	0
NO occupancy from peak heights (%)	77	56	56	41	34	0

Table 5.1: Data processing and refinement statistics for time-resolved DtpB structures with different laser energies. Values in parentheses refer to the outermost resolution shell. *Structures at 16.1 µJ and 8.05 µJ are less reliable due to the low numbers of crystals merged in the datasets (see §3.11).

5.2.1 Data collection with different laser energies

As well as determining the NO-bound structure of DtpB, systematic experiments were performed to quantify the energy needed to release the NO by lysis of the photocage within the crystals. For these experiments, five NO-bound datasets were collected, with deposited laser energy as measured at the sample position in the range 0.81–64 µJ. This data was collected during beamtime on two consecutive days, with no intervening

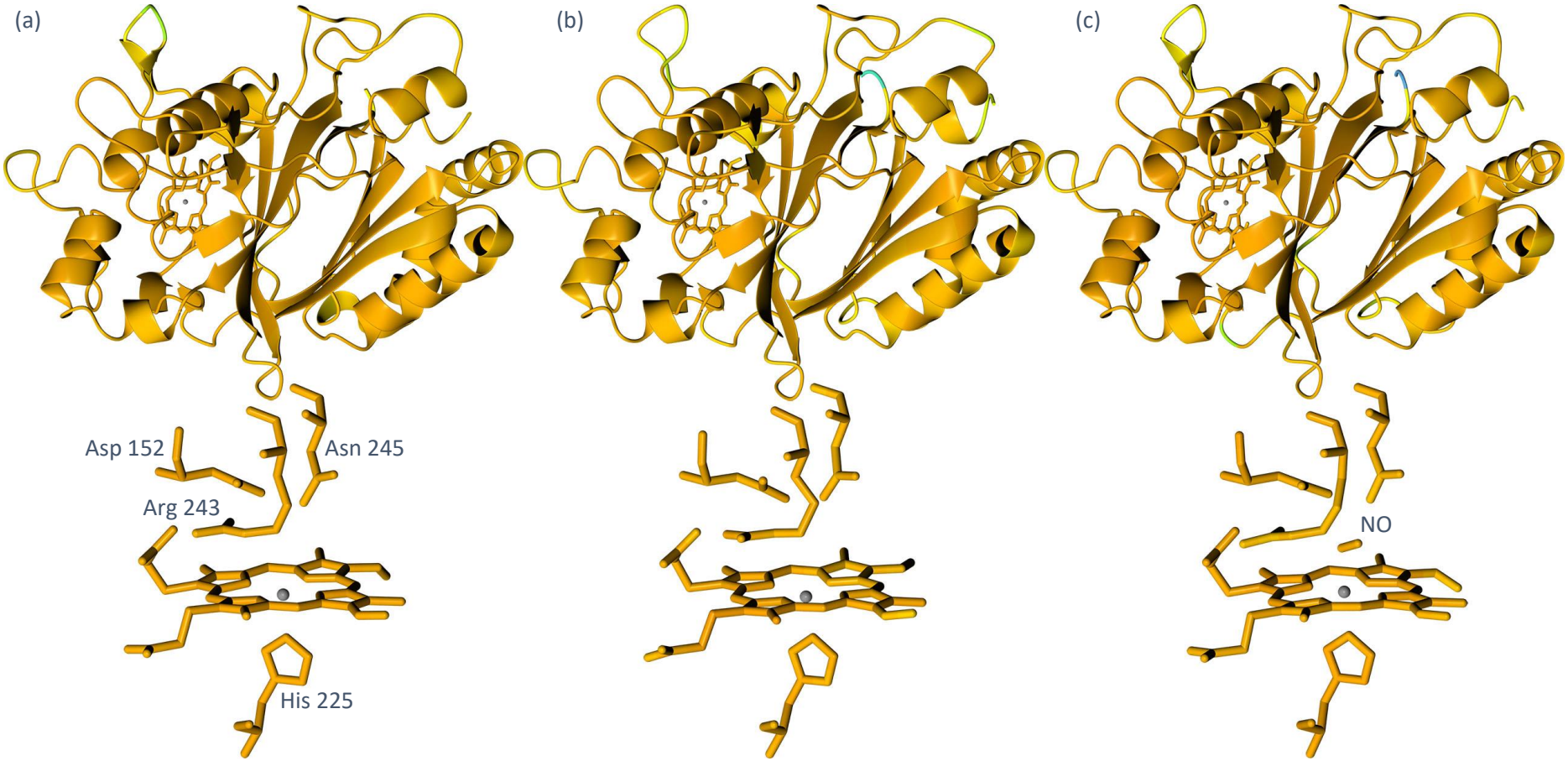

Figure 5.4: DtpB structures coloured by B-factor. (a) SFX (PDB: 6yrj), (b) resting state SSX, (c) NO-bound TR-SSX with 32 µJ laser energy. Orange indicates low B-factor, increasing through green to blue.

changes to the beamline configuration. Laser initiation occurred with 250 pulses over a 5 ms window while operating at 50 kHz, with intensity being changed by addition of UV reflective neutral density (ND) filters into the beam. This used the excite and visit again mode of the fixed target stages (see §3.10), where each position on the chip was visited once to initiate NO release from the photocage, then again to collect diffraction data after a delay. The delay time was 1240 ms, corresponding to visiting 80 positions (four rows of the chip) before returning, with each position including a 10 ms dwell time (during which the 5 ms laser exposure occurred) and 5.5 ms to move between positions.

All NO-bound structures were refined to 2.40 Å resolution, although two structures, at 16 µJ and 8.1 µJ, are less reliable due to lower hit rates in the chips, and therefore data being merged from a lower number (<5000)

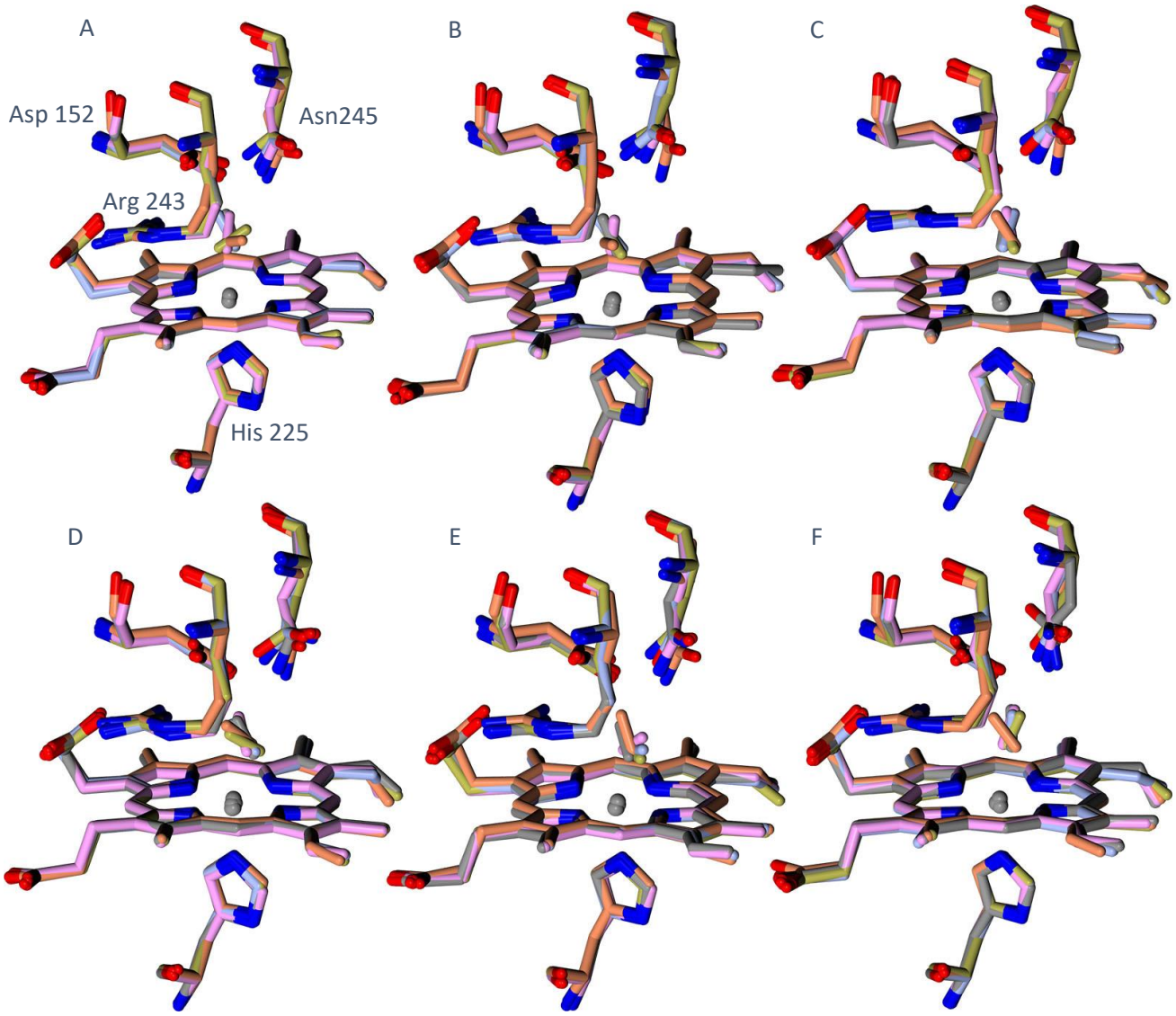

Figure 5.5: NO bound in haem sites of chains A–F of DtpB, from structures collected with different laser energies. Yellow: 64.4 µJ; coral: 32.2 µJ; blue: 16.1 µJ; grey: 8.05 µJ; pink: 0.81 µJ.

of diffraction crystals (§3.11), see the "merged crystals" line in Table 5.1. All structures in the laser energy series show NO bound in all six haem sites, with a consistent Fe–N distance averaging 2.0 Å with a standard deviation (SD) of 0.03 Å. The Fe–NO angle is significantly non-linear, varying between chains and structures in the range 70–160°, and featuring no clear trend between chains or laser energies, shown in Table 5.2. The orientation of the NO around the His–Fe–N axis varies between chains and structures, as shown by superpositions for each chain of structures with different laser energies in Figure 5.5. It is possible that the modelled positions of the NO are sampling different alternate conformers that both exist in the same chains across the crystal, as seen in a high resolution (1.56 Å) structure of McCP-β collected at 100 K (Adams *et al.*, 2023), but at the limited resolution of the structures (2.40 Å), it is more likely that the NO position cannot be accurately determined. The acute (<90°) bond angles identified in some chains are likely also artefacts of the low resolution of the structures. The distal NO-bound structures of McCP show a similar Fe–N bond length, averaging 2.0 Å at room temperature, compared to the 2.1 Å seen in DtpB. The Fe–NO bond angle averaging 120° for DtpB is also similar for McCP-β (130°), although both have large uncertainties (SD 23° and 24°). In the NO bound to the proximal face of the haem in ferrous AxCP, measured at 100 K, the bond length and angle are 1.8 Å and 140°. This bond length in AxCP is the same as that observed in McCP-β at 100 K (1.8 Å), so at

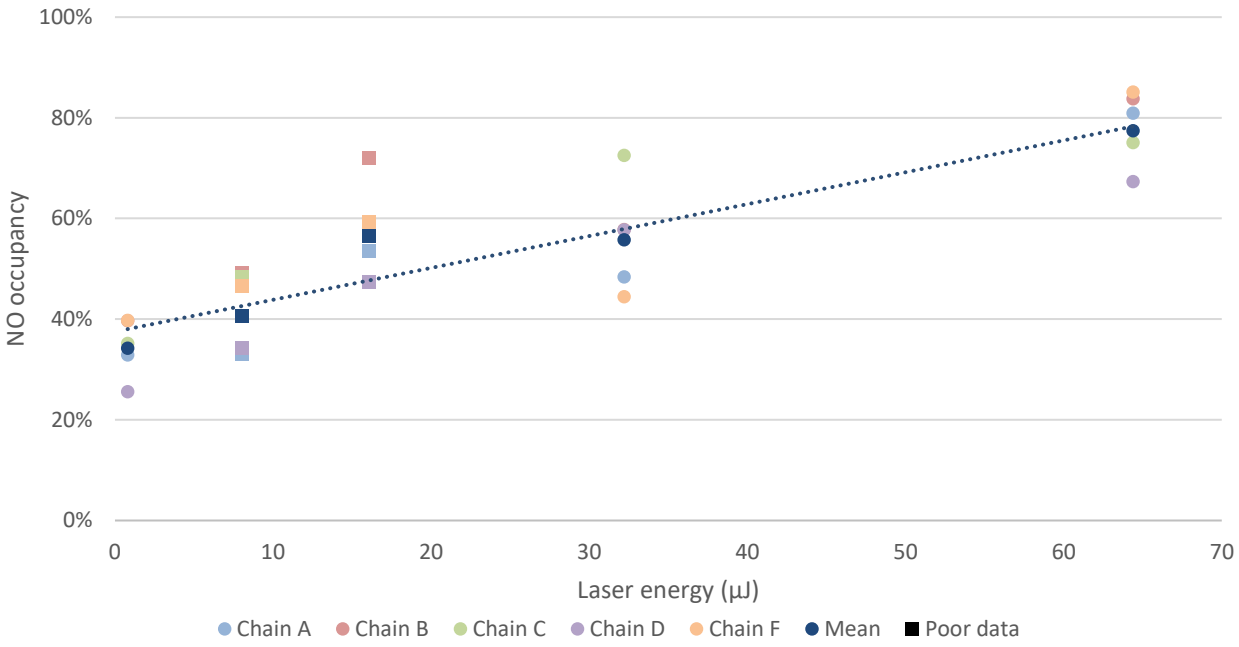

Figure 5.6: NO occupancies as estimated from comparing $F_{obs} - F_{calc}$ difference map peak heights from DtpB structures with different laser energies. Data from less reliable structures is indicated with squares (▪).

room temperature the bond lengths in all three proteins may also be the same as each other. The Fe–N bond lengths observed in McCP and DtpB imply that the local environment or redox state of the iron do not play a significant role bond length, but temperature does. The Fe–NO bond angle in AxCP is larger than those seen in the McCP-β and DtpB structures, likely an effect of the different environment in the proximal pocket compared to the distal pockets of the other proteins.

The occupancy of the NO within each site has been estimated via two methods B-factor analysis and electron density peak height, which are described in sections §5.2.3 and §5.2.4 respectively. Results from the B-factor analysis for each structure are shown in Figure 5.7. The resting state structure is also included to demonstrate the B-factor differences when NO is modelled despite the true occupancy being zero. The oscillation in B-factor difference seen in the range 50–80 % occupancy for the resting state structure is due to noise in the data, causing values for the NO B-factors at each occupancy to refine to one of two values. This behaviour is only seen in the resting state data (Figure 5.7 and Figure 5.16), so is likely an artefact of modelling in a ligand that is not present in the active site. Occupancy varies with laser energy (Table 5.3), in an approximately linear trend from 39 % occupancy with 0.81 μJ laser energy, up to 84 % at 64 μJ, illustrated in Figure 5.8. The graph illustrates that occupancies obtained from both B-factor and electron density peak height analysis are in broad agreement but have a consistent 5 % difference between values. The datapoints from the less reliable structures (indicated with squares) are excluded from the trendlines.

Using similar calculations to those described for McCP-β in §3.5, the number of photons available per photocage molecule could be determined. For these experiments, the laser focal spot was measured at 50 μm (FWHM), and the crystals were approximately 10 × 10 × 10 μm in size, with a solvent content of 53 %. The final concentration of photocage in the crystals was calculated to be 16 mM, providing 20 photocage molecules per unit cell. In space group $P2_12_12_1$, there are four asymmetric units per unit cell. As the asymmetric unit contains a DtpB hexamer with one haem group per monomer, there are 24 haem groups per unit cell. With a 10 μm crystal depth, the absorbance is 0.22, equating to approximately 40 % of the light being absorbed in the crystal. The 10 μm crystal occupies 5.1 % of the beam FWHM area, so 2.9 % of the laser energy hits the crystal, and

	Fe–N bond length (Å)								Fe–NO bond angle (°)							
Laser energy (μJ)	A	B	C	D	E	F	Mean	SD	A	B	C	D	E	F	Mean	SD
64.4	1.96	1.95	1.98	2.02	2.00	2.00	1.99	0.02	109	110	111	97	75	150	109	22
32.2	1.99	2.02	1.99	2.01	2.01	2.02	2.01	0.01	101	96	133	110	141	124	117	16
16.1	2.01	1.99	2.01	2.01	2.01	2.01	2.01	0.01	145	131	135	112	99	147	128	17
8.05	2.03	2.02	2.03	2.04	2.02	20.7	2.04	0.02	156	150	142	130	117	163	143	16
0.81	1.97	1.98	2.00	2.03	1.98	2.05	2.00	0.03	105	104	101	94	74	138	103	19
Mean	1.99	1.99	2.00	2.02	2.00	20.3	2.01	0.03	123	118	124	109	101	144	120	23

Table 5.2: Bond lengths and angles for NO bound in the haem sites of DtpB, in structures with different laser energies.

	NO occupancy from B-factors (%)								NO occupancy from peak heights (%)								Method
Laser energy (μJ)	A	B	C	D	E	F	Mean	SD	A	B	C	D	E	F	Mean	SD	difference
64.4	68	86	75	100	77	100	84	12	81	84	75	67	72	85	77	6	7
32.2	58	71	69	60	51	40	58	11	48	58	73	58	54	44	56	9	2
16.1	51	67	71	58	61	79	65	9	54	72	58	47	48	59	56	8	8
8.05	46	57	60	45	37	53	50	8	33	49	48	34	32	47	41	7	9
0.81	33	43	43	41	26	50	39	8	33	40	35	26	32	40	34	5	5

Table 5.3: NO occupancy as a percentage, for NO bound in the haem sites of DtpB, in structures with different laser energies. Occupancy was estimated in two ways; by analysing B-factors, and by comparing electron density peaks. The differences between the two methods for each structure is shown in the final column, with the occupancy estimated from B-factors being higher for all structures.

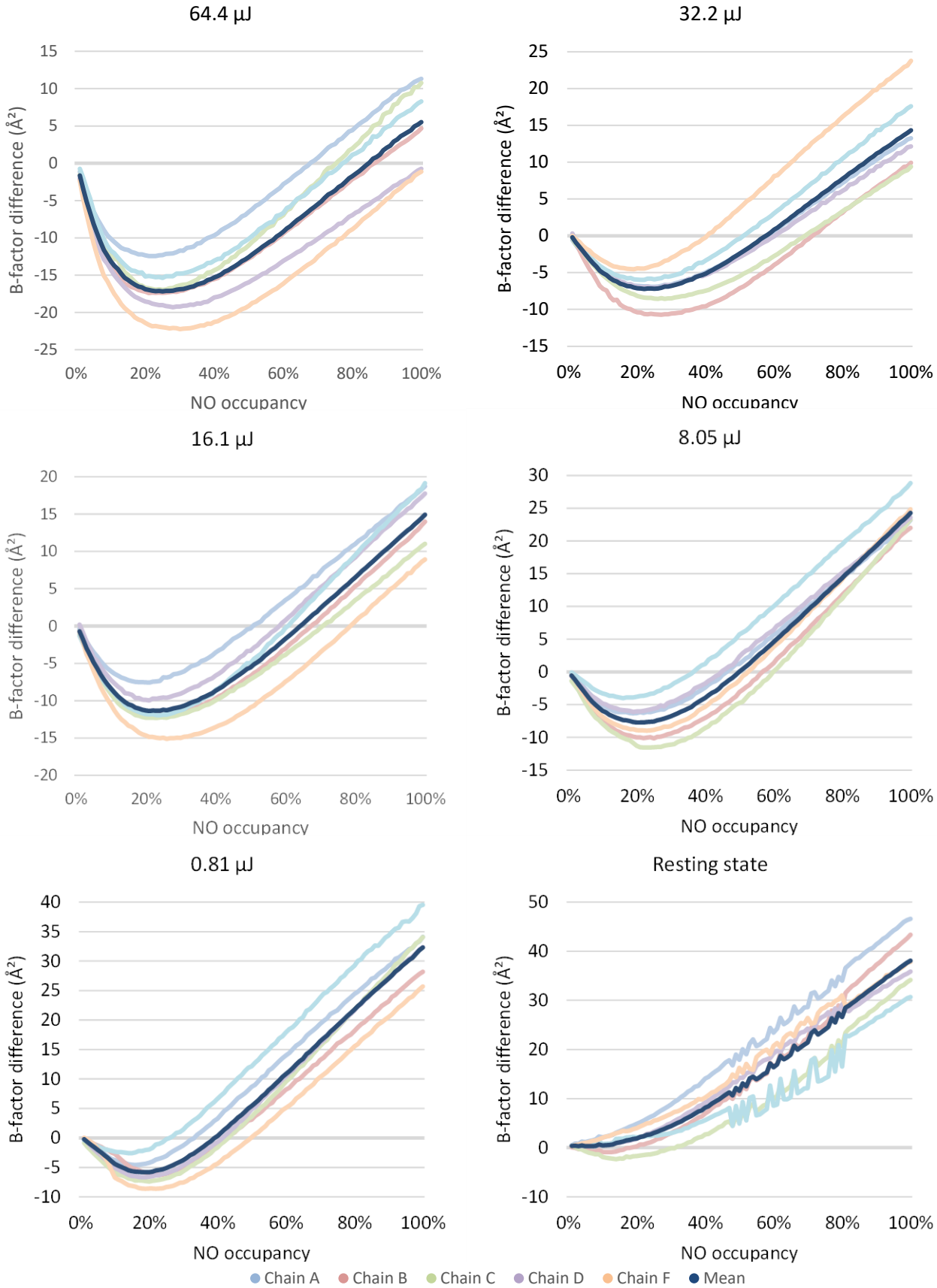

Figure 5.7: Graphs of B-factor difference, comparing the haem iron and the NO nitrogen, against NO occupancy in a DtpB structure. Positive values represent the nitrogen B-factor being higher than the iron. Data is from TR-SSX structures with different laser energies, and a resting state structure.

1.2 % of the total laser energy is absorbed. This equates to 1.8×10^{10} photons being absorbed per microjoule of laser energy provided, or 38 photons per unit cell. This value is a factor of two greater than the number of photocage molecules per cell, so is at or very near single photon absorption. As there are 20 photocage molecules per unit cell, and the quantum yield of the photocage is 1.4, with sufficient photons NO becomes available in a slight excess of approximately 29 molecules to the 24 haem groups in the unit cell unit. The slight excess of photocage in the DtpB crystals allows for less than full photolysis of the photocage while still achieving high NO occupancy in the structures.

5.2.2 Timeseries data collection

To investigate changes in NO occupancy over time (but also increasing X-ray dose) in a manner that is very efficient on sample consumption, a timeseries experiment was performed, where consecutive structures were collected from the same crystals (as described in §3.10). This data was collected from two chips, using a laser energy of 16 µJ, in the middle of the range used for the previous experiments. Structures were collected using

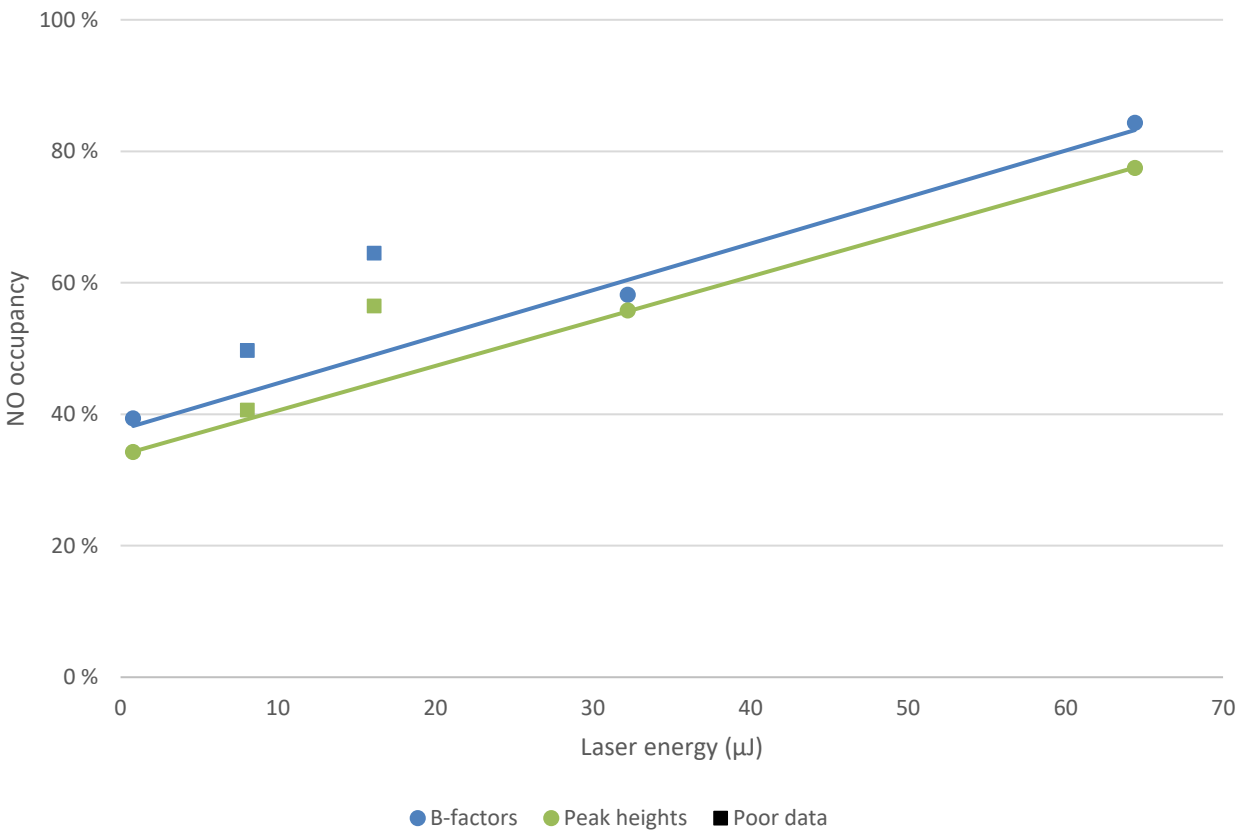

Figure 5.8 : Comparison of the two methods of estimating NO occupancy for DtpB, averaged from chains A–F. The less reliable datapoints (▪) are excluded from the trendlines.

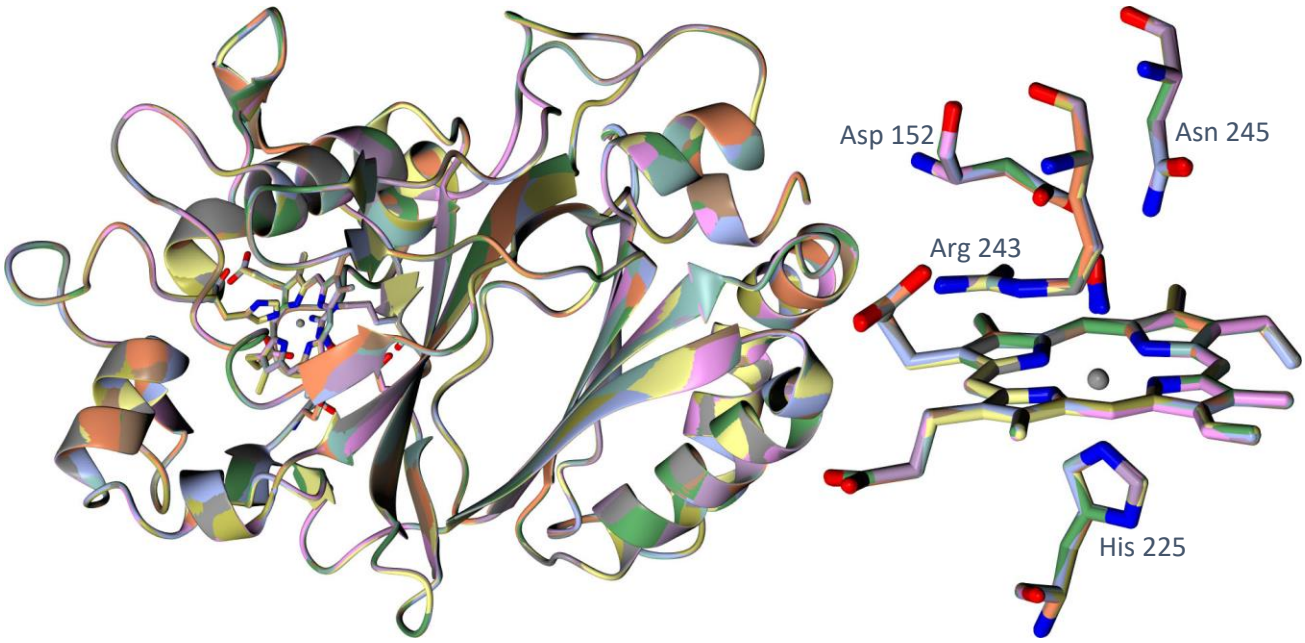

Figure 5.9: Ten DtpB structures at 1.95 Å resolution from a timeseries data collection superposed. No significant differences are visible between the structures, including the position of the NO within the haem pocket. Arg 243, which moves away from the haem upon NO binding, is stationary here, also indicating no change in NO presence across the timeseries.

ten consecutive images with 10 ms exposure time during a single long X-ray exposure. These were at intervals corresponding to 10 ms before laser initiation, during initiation, and up to 80 ms after initiation. The structure of DtpB at the first timepoint (−10 ms) was solved by molecular replacement, again using the SFX structure (PDB: 6yrj) as a search model (Lučić *et al.*, 2020b), and all subsequent structures were solved by refinement of the same model. All structures were solved to 1.95 Å resolution using a $CC_{1/2}$ cutoff of 0.3 and multiplicity of greater than 10 (see §2.9). Data processing and refinement statistics are shown in Table 5.5.

Unexpectedly, NO was shown bound in the haem site in all structures, including the structure collected before the laser exposure that was intended to release NO from the photocage. Figure 5.9 shows no significant changes in the protein structure occurring during the data collection, including in the haem sites. The same occupancy estimation by analysing B-factors (see §5.2.3) was performed as with the structures with different laser energies, shown in Figure 5.10 and Table 5.4. The occupancies are plotted on a graph against time in Figure 5.11, showing occupancy does not change over time, staying at approximately 36 % across all timepoints. Interestingly, the occupancy is significantly lower in chain E compared to the rest of the chains (16 % against 39°%). The Fe–N bond lengths and Fe–NO bond angles are listed in Table 5.6, and plotted in Figure 5.12 and Figure 5.13. The bond lengths in all chains of all structures are not significantly different,

averaging 2.1 Å (SD 0.0 Å), and the bond angles, averaging 150° (SD 12°), do not show any association with either chain or timepoint. The bond angles show less variation than in the structures collected at different laser energies (SD 23°), likely due to the higher resolution of this data more precisely showing the position of the NO.

	NO occupancy (%)						
Timepoint (ms)	A	B	C	D	E	F	Mean
-10	23	40	44	35	10	35	31
0	37	45	36	39	10	46	36
10	32	38	37	52	22	42	37
20	43	35	47	41	20	45	39
30	23	44	31	48	33	43	37
40	35	29	55	44	10	44	36
50	35	45	48	45	12	46	39
60	25	38	43	43	10	41	33
70	28	35	36	43	10	40	32
80	37	39	41	38	15	49	37
Mean	33	39	42	44	16	44	36
SD	6	5	7	5	7	4	11

Table 5.4: NO occupancy as a percentage for DtpB structures in a timeseries dataset collected before and after laser exposure. Occupancy was estimated by analysing B-factors. Values shown are for each chain, and a mean for each chain and each structure.

Timepoint (ms)	-10	0	10	20	30	40	50	60	70	80
Resolution (Å)	71.25–1.95 (1.98–1.95)	71.26–1.95 (1.98–1.95)	71.25–1.95 (1.98–1.95)	71.25–1.95 (1.98–1.95)	71.26–1.95 (1.98–1.95)	71.26–1.95 (1.98–1.95)	71.26–1.95 (1.98–1.95)	71.26–1.95 (1.98–1.95)	71.26–1.95 (1.98–1.95)	71.26–1.95 (1.98–1.95)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell (Å), *a*, *b*, *c*	87.2, 123.1, 195.6	87.2, 123.1, 195.6	87.2, 123.1, 195.6	87.2, 123.1, 195.6	87.2, 123.1, 195.7	87.2, 123.1, 195.7	87.2, 123.1, 195.7	87.2, 123.1, 195.7	87.2, 123.1, 195.7	87.2, 123.1, 195.7
Merged crystals	20 298	20 409	21 164	21 164	21 262	21 506	21 569	216 76	21 618	21 572
Unique reflections	153 563 (7634)	153 635 (7653)	153 585 (7640)	153 598 (7643)	153 620 (7653)	153 630 (7648)	153 663 (7653)	153 663 (7653)	153 665 (7649)	153 675 (7647)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Multiplicity	93.9 (78.1)	97.3 (80.1)	98.1 (81.4)	98.2 (81.4)	100.2 (83.1)	99.2 (82.2)	99.5 (82.2)	100.1 (82.6)	99.7 (82.3)	99.6 (82.2)
$CC_{1/2}$	0.995 (0.335)	0.995 (0.323)	0.996 (0.326)	0.996 (0.349)	0.996 (0.332)	0.996 (0.343)	0.996 (0.346)	0.996 (0.336)	0.996 (0.334)	0.996 (0.312)
R_{split}	0.110 (1.273)	0.110 (1.324)	0.106 (1.249)	0.105 (1.207)	0.104 (1.241)	0.104 (1.231)	0.104 (1.277)	0.104 (1.293)	0.104 (1.322)	0.104 (1.312)
Mean I	26.8 (1.4)	26.6 (1.4)	27.3 (1.4)	27.5 (1.4)	27.4 (1.4)	27.6 (1.4)	27.6 (1.3)	27.4 (1.3)	27.3 (1.3)	27.4 (1.3)
I/σI	10.8 (0.9)	10.9 (0.8)	11.3 (0.9)	11.4 (0.9)	11.5 (0.9)	11.5 (0.9)	11.6 (0.9)	11.5 (0.8)	11.5 (0.8)	11.5 (0.8)
N_{obs}	14 414 248 (596 091)	14 944 441 (612 975)	15 069 349 (622 012)	15 083 037 (621 809)	15 399 659 (635 847)	15 237 157 (628 373)	15 289 695 (629 342)	15 374 259 (632 310)	15 326 353 (629 611)	15 302 978 (628 839)
R_{work}	0.155	0.155	0.153	0.153	0.153	0.153	0.152	0.153	0.154	0.153
R_{free}	0.206	0.203	0.202	0.201	0.203	0.203	0.202	0.202	0.203	0.202
RMSD bond lengths (Å)	0.0145	0.0146	0.0143	0.0142	0.0143	0.0144	0.0145	0.0144	0.0144	0.0146
RMSD bond angles (°)	2.182	2.171	2.042	2.164	2.153	2.131	2.163	2.155	2.137	2.151
Ramachandran favoured (%)	95.94	96.00	96.11	96.38	96.11	96.33	96.00	96.55	96.27	96.60
Mean NO occupancy (%)	31	36	37	39	37	36	39	33	32	37

Table 5.5: Data processing and refinement statistics for NO-bound DtpB structures collected in a timeseries before and after laser exposure. Values in parentheses refer to the outermost resolution shell.

	Fe–N bond length (Å)								Fe–NO bond angle (°)							
Timepoint (ms)	A	B	C	D	E	F	Mean	SD	A	B	C	D	E	F	Mean	SD
-10	2.05	2.05	2.04	2.06	2.03	2.06	2.05	0.01	160	169	157	149	134	150	153	11
0	2.07	2.03	2.04	2.05	2.02	2.04	2.04	0.02	151	172	141	160	171	160	159	11
10	2.08	2.06	2.06	2.06	2.05	2.03	2.06	0.01	146	171	137	151	127	154	148	14
20	2.07	2.03	2.04	2.05	2.06	2.04	2.05	0.01	148	163	154	145	143	154	151	7
30	2.06	2.06	2.04	2.05	2.05	2.06	2.05	0.01	158	177	135	150	162	146	155	13
40	2.06	2.02	2.07	2.05	2.04	2.03	2.05	0.02	144	155	140	145	143	156	147	6
50	2.05	2.06	2.05	2.04	2.04	2.05	2.05	0.01	153	158	142	151	160	149	152	6
0	2.06	2.04	2.05	2.05	2.04	2.04	2.05	0.01	152	159	135	143	154	161	150	9
70	2.05	2.03	2.04	2.06	2.03	2.04	2.04	0.01	153	150	137	147	118	167	145	15
80	2.04	2.04	2.04	2.05	2.04	2.06	2.05	0.01	169	172	144	155	171	160	162	10
Mean	2.06	2.04	2.05	2.05	2.04	2.05	2.05	0.01	153	165	142	150	148	156	152	12

Table 5.6: Bond lengths and angles for NO bound in the haem sites of DtpB, for different points in a timeseries.

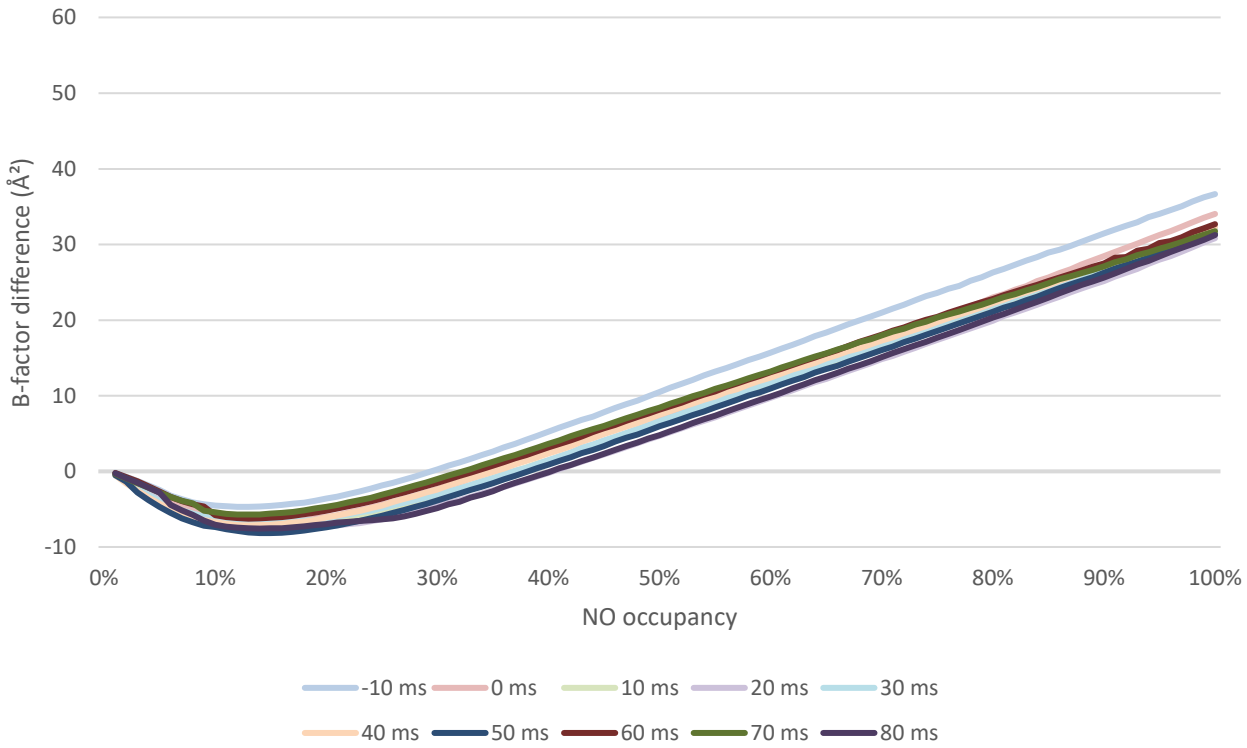

Figure 5.10: B-factor differences between the haem iron and the NO nitrogen plotted against NO occupancy. Averaged across chains A–F for a timeseries of photocaged NO binding to DtpB. This indicates an NO occupancy of approximately 36 % across all timepoints, where the curves cross the X axis.

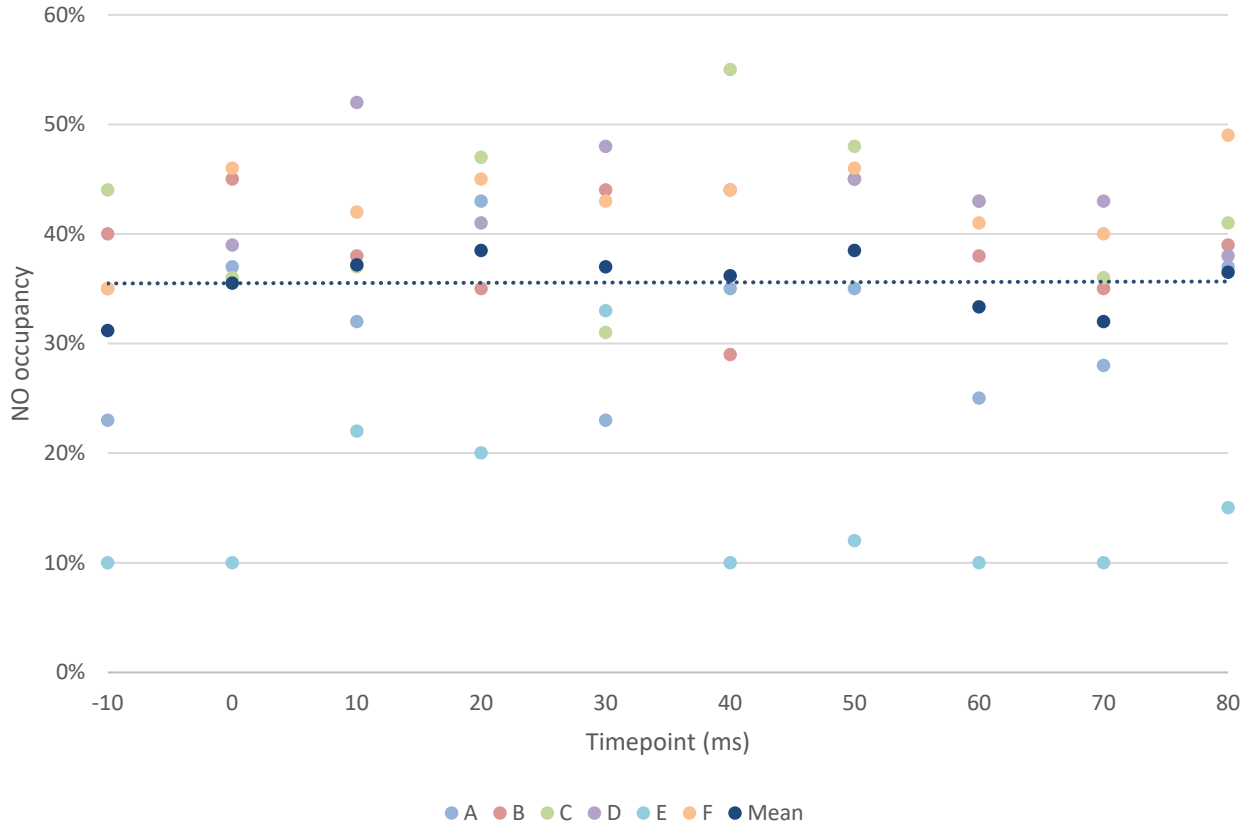

Figure 5.11: NO occupancy of DtpB structures over time, from 10 ms before to 80 ms after 16 µJ laser illumination (at 0 ms). The occupancy averages 36 % (SD 11 %) across all timepoints.

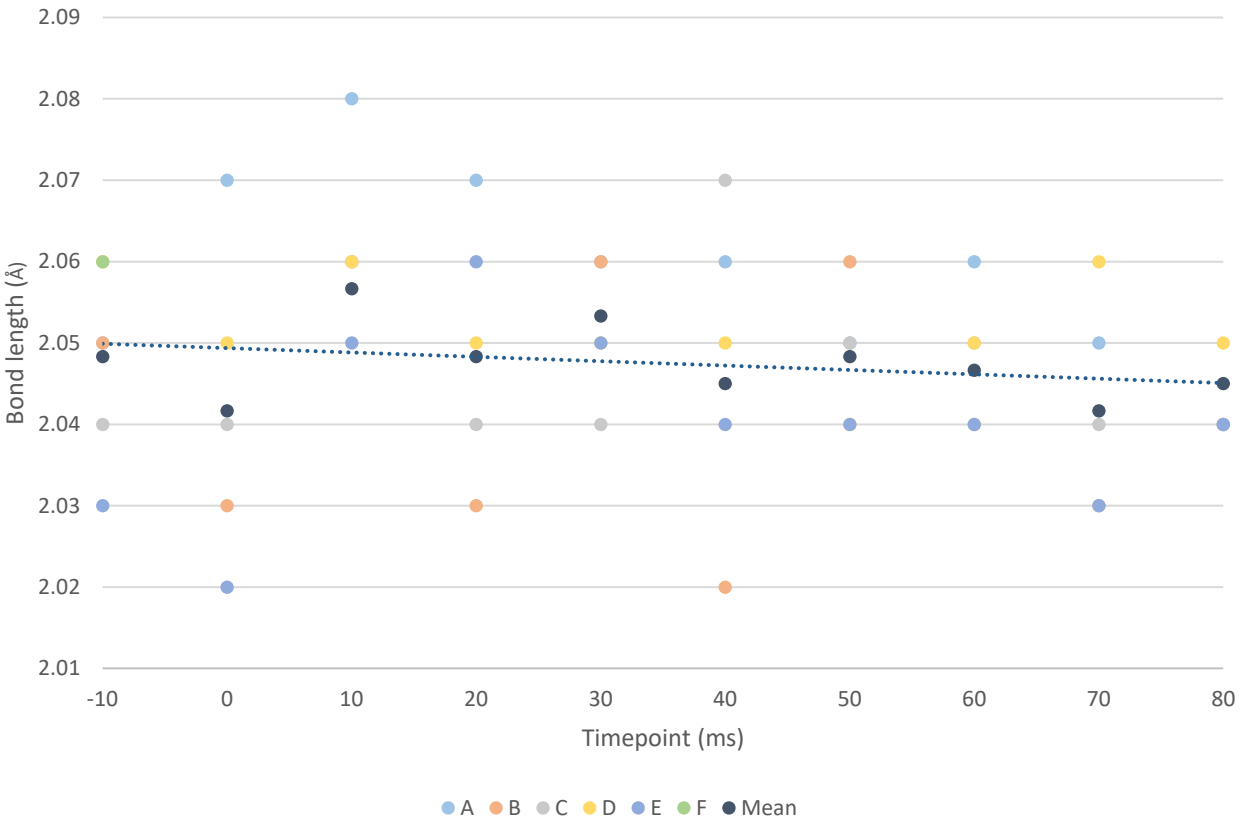

Figure 5.12: Fe-N bond lengths of NO-bound DtpB structures over time, from 10 ms before to 80 ms after 16 µJ laser illumination. The average bond length across the structures is 2.05 Å (SD 0.01 Å).

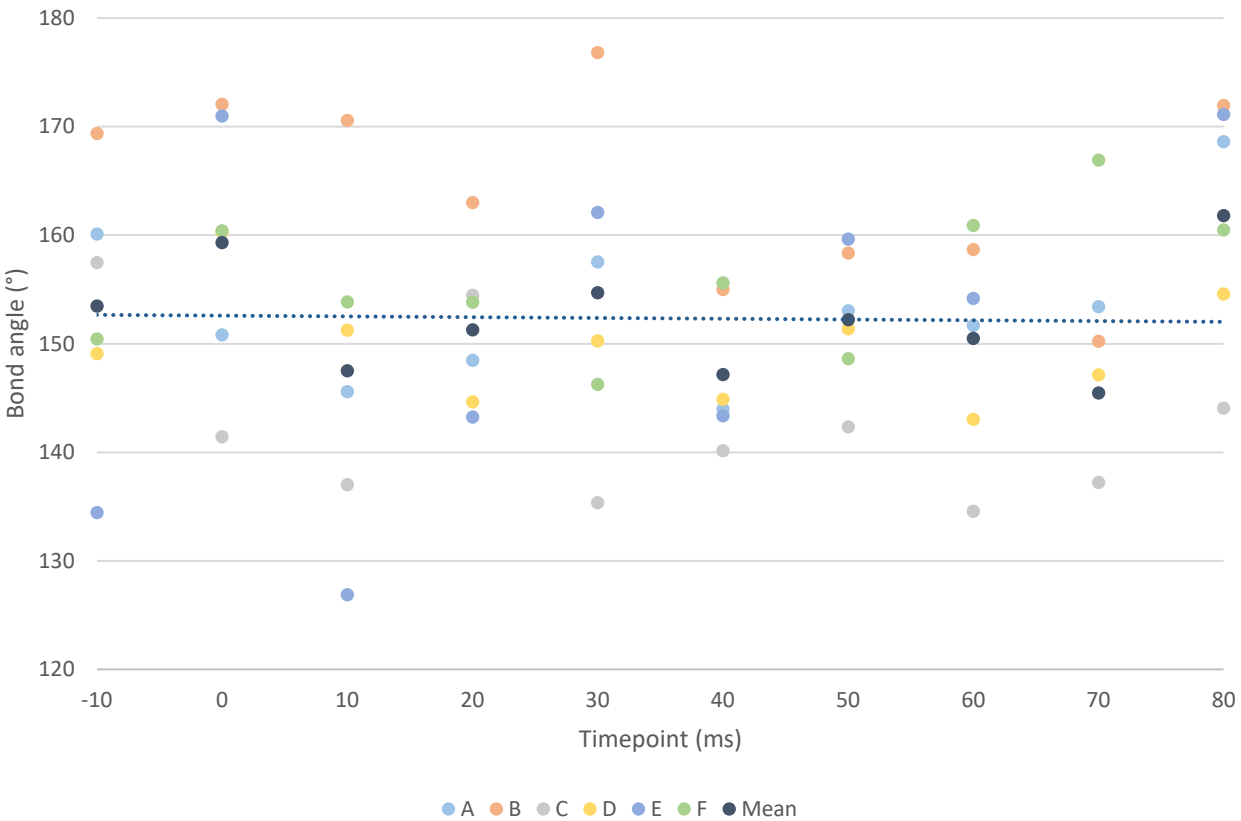

Figure 5.13: Fe-N–O bond angle of NO-bound DtpB structures over time, from 10 ms before to 80 ms after 16 µJ laser illumination. The average bond angle across the structures is 152° (SD 12°).

5.2.3 Estimating occupancy by analysing B-factors

To best estimate the occupancy of the bound NO ligand in the distal pocket of both McCP-β and DtpB structures, two different approaches were used. Analysis with the occupancy refinement within Phenix was found to be unreliable when used on these datasets, reflecting the incomplete understanding of the best method of occupancy estimation. The current approaches of tackling this problem, and their limitations, are discussed in detail in §5.3.1. The first analysis performed uses comparison of atomic B-factors between the haem iron and the nitrogen atom of the bound NO. By refining a model of the protein and bound ligand with the occupancy of the ligand set to different, fixed, values, systematic variation in the real space B-factors between the independent refinement runs can be used to estimate the correct occupancy of the ligand in the dataset. Both occupancy and B-factor are included as fields in ATOM and HETATM (heterogen atom, for any non-peptide atoms) records in the text based PDB file format, so analysis can be carried out in an automated manner using basic text manipulation tools in Bash, alongside crystallographic refinement software. For the results presented here, a script was written to automate analysis of the data collected from the proteins McCP-β and DtpB. The script takes an input PDB file, modifies the NO occupancy, runs REFMAC5 for refinement (Murshudov *et al.*, 2011), and tabulates the relevant B-factor values from the refined models into a comma-separated values (CSV) file.

The script, included as Appendix 1, takes as input: a PDB file for the model (including the NO ligand); the experimental electron density map as MTZ file; and a REFMAC5 ligand dictionary from previous refinement (Vagin *et al.*, 2004). Copies of the input PDB file are made for each occupancy in the range 0.00–1.00 in 0.01 increments, and the occupancy of the atoms in all NO molecules are set by searching for the HETATM records and replacing the values in the corresponding occupancy columns. REFMAC5 jobs with 10 cycles are run in parallel for each of these 101 models, in batches of 10 to prevent system crashes. After refinement, the B-factors for the haem iron and the NO nitrogen in each chain are taken from the REFMAC5 output structures and saved to a CSV file along with the NO occupancy values. This CSV file can be imported into a spreadsheet for comparing B-factors and identifying the ligand occupancy. As well as the CSV file, the modified copies of the input model, and the REFMAC5 log and output files, the script saves a basic log file, an example of which

from analysis of a DtpB structure is included as Appendix 2. For a DtpB model containing six chains, the script takes approximately 30 min to run on a workstation PC, primarily due to the duration of the refinement steps.

The output from the script shows that the B-factor of the haem iron stays consistent at close to 27 Å² across the complete range of occupancies for the NO, whereas the B-factor for the NO nitrogen changes as a function of occupancy. This is demonstrated with a line graph of data from chain A in the 32 µJ laser energy DtpB structure, in Figure 5.14. The difference (green, right-hand axis) between the B-factors for the iron (orange) and the nitrogen (blue) in each chain are a function of variable B-factor for the NO nitrogen, referenced to the stable value for the iron. The B-factor difference is clearly dependent on occupancy, as can be seen from the data forming a smooth curve. To the right of the point where the two B-factors are equal (a), and the right-hand X axis is crossed, at 58 % occupancy, the B-factor difference increases linearly as NO occupancy increases towards 100 %, representing increasing B-factor of the nitrogen. This indicates disorder of the modelled ligand to compensate for the excessive occupancy as it strays further above the value suggested by the experimental data. Immediately to the left of the point where the X axis is crossed, the B-factor for the nitrogen decreases with decreasing occupancy. This point where the occupancies are equal is taken to be the actual occupancy of the NO in the crystal.

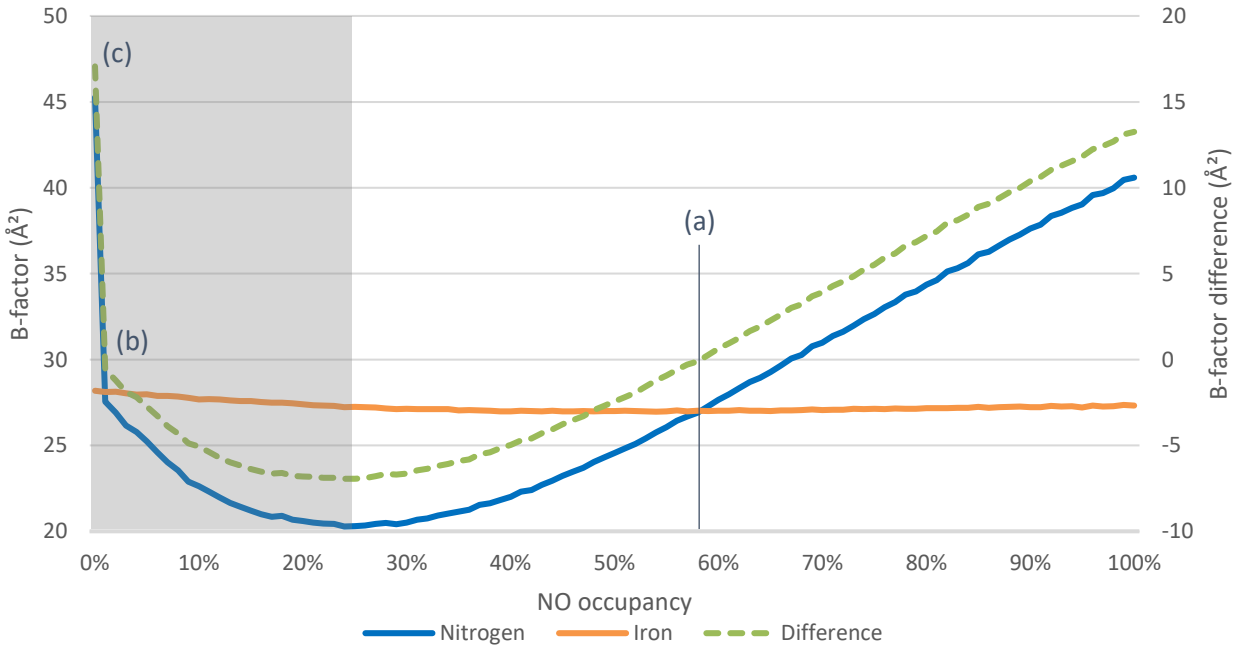

Figure 5.14: A graph of B-factor for the haem iron (orange) and the NO nitrogen (blue) against NO occupancy of a DtpB structure refined with different ligand occupancies. The difference between the two values (green) is plotted on the right-hand axis. Data represents chain A of the TR-SSX structure with 32. µJ of laser energy.

As occupancy decreases further, below 25 % occupancy, this trend reverses, returning to the B-factors being equal (on the X axis) at 1 % occupancy (b). These models and derived B-factor differences at low occupancy, in the grey area of the graph, are unreliable due to the difficulty of modelling ligands with occupancy below 20 %. Where the NO is modelled but completely absent from the structure, at 0 % occupancy, the ligand is not refined, so the nitrogen B-factor remains at its starting value of 45 Å² (c), from when the ligand was added and refined with 100 % occupancy. This makes the ligand B-factors for the 0 % occupancy model very close to those for the 100 % occupancy model at the opposite end of the graph. Therefore, the point for 0 % occupancy has not been plotted on the graphs shown here, except for the example in Figure 5.14. The B-factor difference is plotted for all chains in the 32 µJ structure, along with the mean value (dark blue) in Figure 5.15. For each chain in the model, the line crosses the X axis, indicating the B-factors of the nitrogen and the iron are equal, at a point in the NO occupancy range 40–71 %, averaging 58 % with SD of 11 %. Equivalent graphs for each laser energy are included earlier as Figure 5.7.

In Figure 5.16, the average B-factor differences across chains for all structures are plotted on one graph. Besides 16 and 32 µJ, the lines cross the X axis in order of laser energy. For demonstration purposes, NO has also been modelled into the resting state structure, despite it not being present in the crystals or evidenced

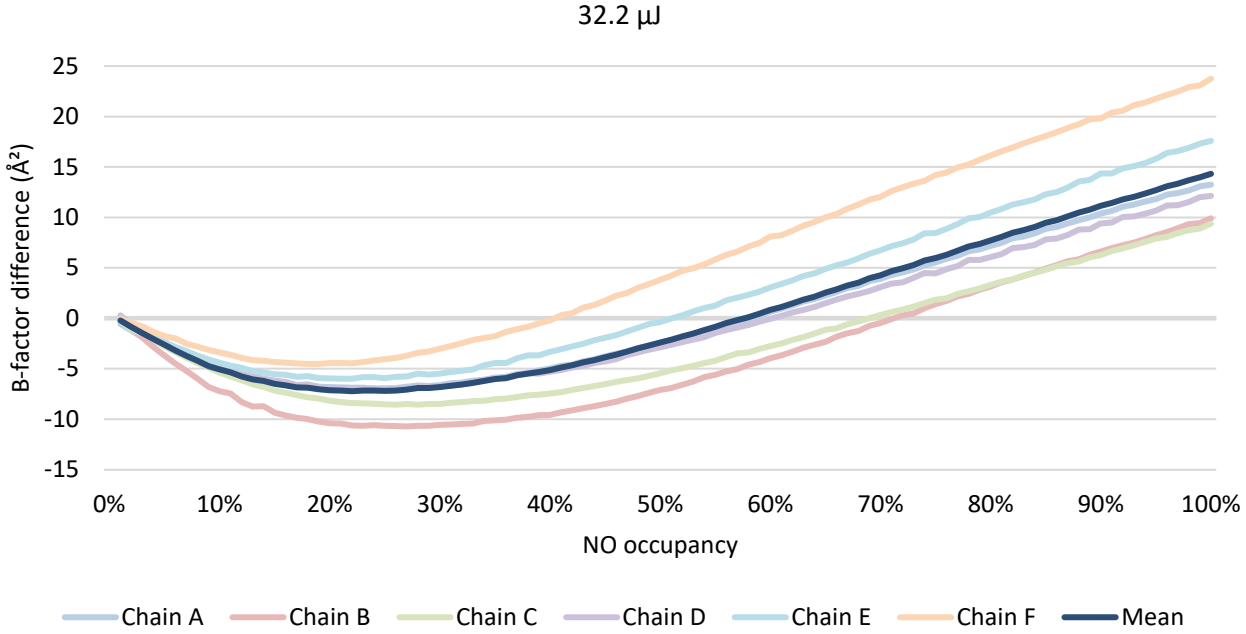

Figure 5.15: Difference in B-factors, comparing the haem iron and the NO nitrogen, plotted against NO occupancy in a DtpB structure. Positive values represent the nitrogen B-factor being higher than the iron. Data represents the TR-SSX structure with 32 µJ of laser energy.

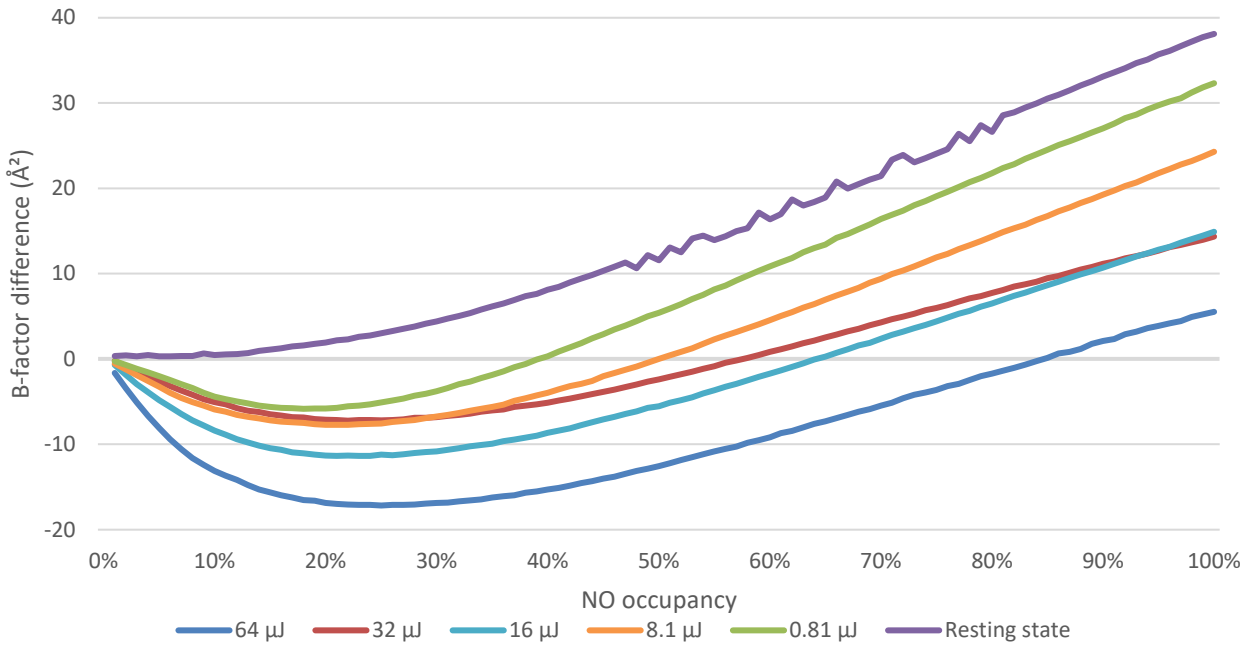

Figure 5.16: Difference in B-factors between haem iron and NO nitrogen, plotted against NO occupancy, for DtpB structures with different laser energies. A resting state structure with NO modelled is also shown.

in the electron density maps. The refinement of NO in the resting state structure shows the B-factor for the nitrogen being equal to the iron at very low occupancy (3 %, purple trace), expectedly indicating that modelling its presence is incorrect.

5.2.4 Estimation of occupancy from electron density peaks

Providing fast feedback on experimental results during complex data collection such as time-resolved experiments is vital for making efficient use of beamtime and getting the best results. The script described above, written for using B-factors to analyse NO binding to McCP-β and DtpB, takes up to 30 min to run on a workstation PC or the Diamond compute cluster, and requires a refined structure with the ligand modelled. The proteins studied are not especially large proteins (Research Collaboratory for Structural Bioinformatics, 2023b), so other structures could take much longer to process. To solve this, an alternative simpler method was sought for estimating occupancies in the NO-bound structures. Getting the same occupancy estimates by a different means also provides validation for the B-factor analysis. To this end, the relative occupancies of the NO in structures were compared using the electron density peaks corresponding to the ligands, in units of RMSD (σ), which can be converted to electrons per cubic ångstrom ($e \cdot Å^{-3}$).

To compare the density peaks, structures were solved with molecular replacement (Vagin and Teplyakov, 2010) using a resting state structure collected previously, and the models were refined without the NO (Murshudov *et al.*, 2011). The heights of the positive difference peaks corresponding to NO in the $F_{obs} - F_{calc}$ maps, shown in Figure 5.17a were then recorded using the difference map peak analysis feature in *Coot* (Emsley *et al.*, 2010). These values from each chain (A–F in DtpB) were collected from structures with different laser energies, shown in Table 5.7. All the datasets were collected under the same conditions, including using crystals from the same batch, during the same beamtime shift, and with the same data collection and processing parameters. This means the electron densities can be reliably compared with each other. Although the $F_{obs} - F_{calc}$ peak height does not provide an absolute occupancy, It does provide relative values for occupancies in the different structures, since the comparison is between the same atom types and the ligands are likely to display similar amounts of disorder. Plotting these peak heights on a graph against laser energy in Figure 5.18, along with the mean values from the chains of each structure, shows an approximately linear

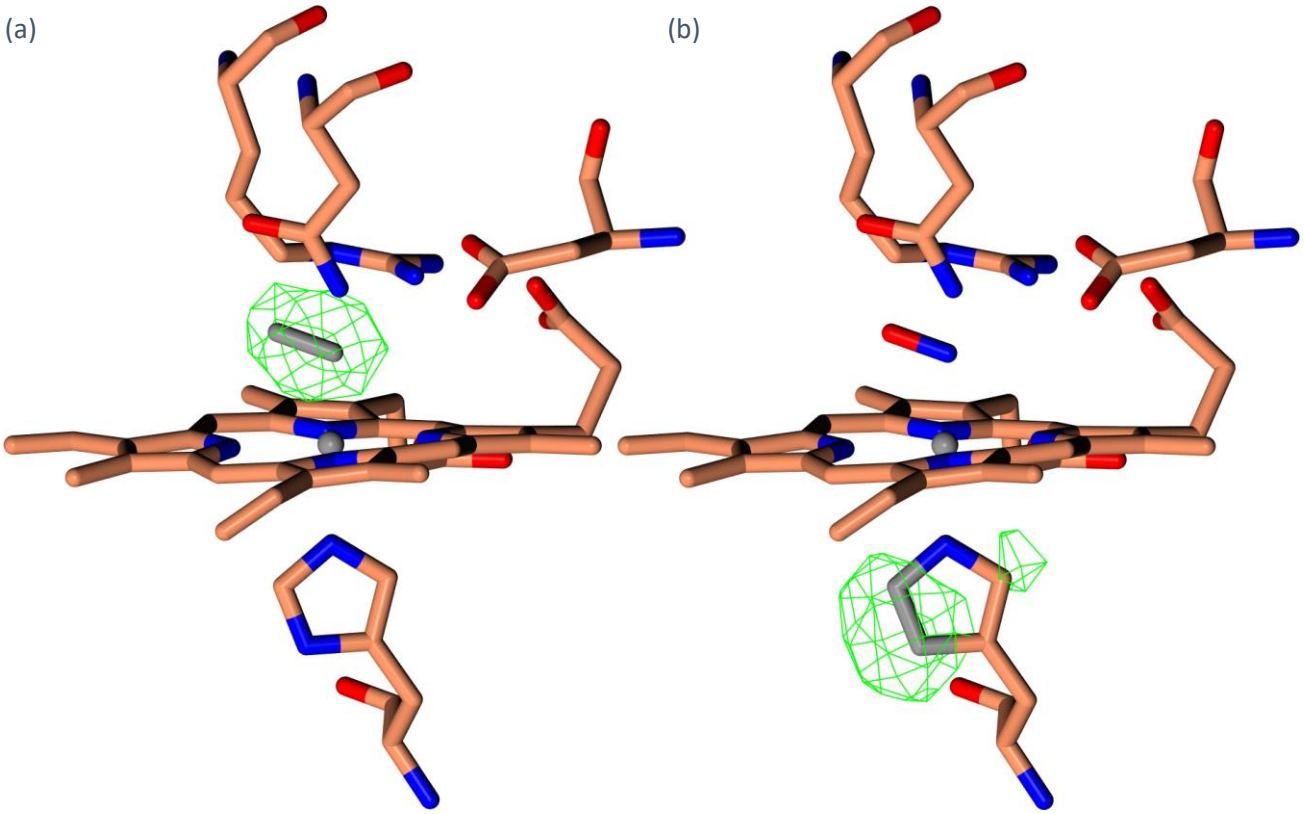

Figure 5.17: Omit maps of a haem site in an NO-bound DtpB structure. (a) Positive difference density for unmodelled NO bound in the distal pocket. (b) Positive difference density for the Nδ¹=Cε¹ atoms deleted from the proximal histidine residue. Atoms missing in the omit maps are indicated in grey. Omit maps are contoured to 3σ. Data taken from the structure with 32 μJ of laser energy.

trend, conforming with expected results from changing laser energy. Data from less reliable structures (explained in §3.11) is indicated with square datapoints (▪).

	Peak heights (σ)							
Laser energy (μJ)	A	B	C	D	E	F	Mean	SD
64.4	10.35	10.26	8.39	9.20	8.65	11.06	9.65	0.97
32.2	7.68	6.95	8.90	8.62	7.73	6.51	7.73	0.84
16.1	7.89	7.84	7.68	7.38	6.40	7.20	7.40	0.51
8.05	5.79	6.87	6.24	5.69	4.08	6.65	5.89	0.91
0.81	4.74	5.43	4.84	4.46	4.09	5.96	4.92	0.62

Table 5.7: Peak heights in RMSD (σ) corresponding to NO in the $F_{obs} - F_{calc}$ difference map from DtpB structures with different laser energies.

To convert these relative occupancies to absolute values in the range 0–100 %, peak heights can be compared to electron density from similar atoms with known occupancy. The protein from which the crystals are formed is present in 100 % occupancy, although some parts of the molecule have a weak or missing electron density and are therefore not modelled to full occupancy. These can include flexible or disordered regions, such as terminal residues or poorly defined sidechains, or atoms likely to exhibit X-ray damage, such as carboxyl groups on aspartate or glutamate sidechains. A suitable comparison for comparing densities of the NO ligand in McCP-β and DtpB is the proximal histidine residue on the opposite side of the haem. This residue is close to the haem and the NO site, so shares the same local environment, and is ordered within the structures. As the proximal histidine is conserved across many haem proteins, it would also be applicable to ligand binding in other proteins.

To compare the difference density peak for the NO with those of the proximal histidine, the Nδ¹=Cε¹ atom pair was deleted from the sidechain, and an omit map ($F_{obs} - F_{calc}$ difference map) was generated, shown in Figure 5.17b. The peak heights for the positive difference density for each chain were recorded in the same way as those for the NO. By taking the NO density peak height as a proportion of the histidine Nδ¹=Cε¹ density peak height for each chain, values could be obtained for the occupancies of the NO, shown in Table 5.3 and plotted in Figure 5.6, including mean values for each laser energy. To validate the two methods for occupancy estimation, this data was compared with the values from analysing B-factors by plotting them together on the same graph, shown in Figure 5.8. As can be seen from the graph, the two methods provide very similar results,

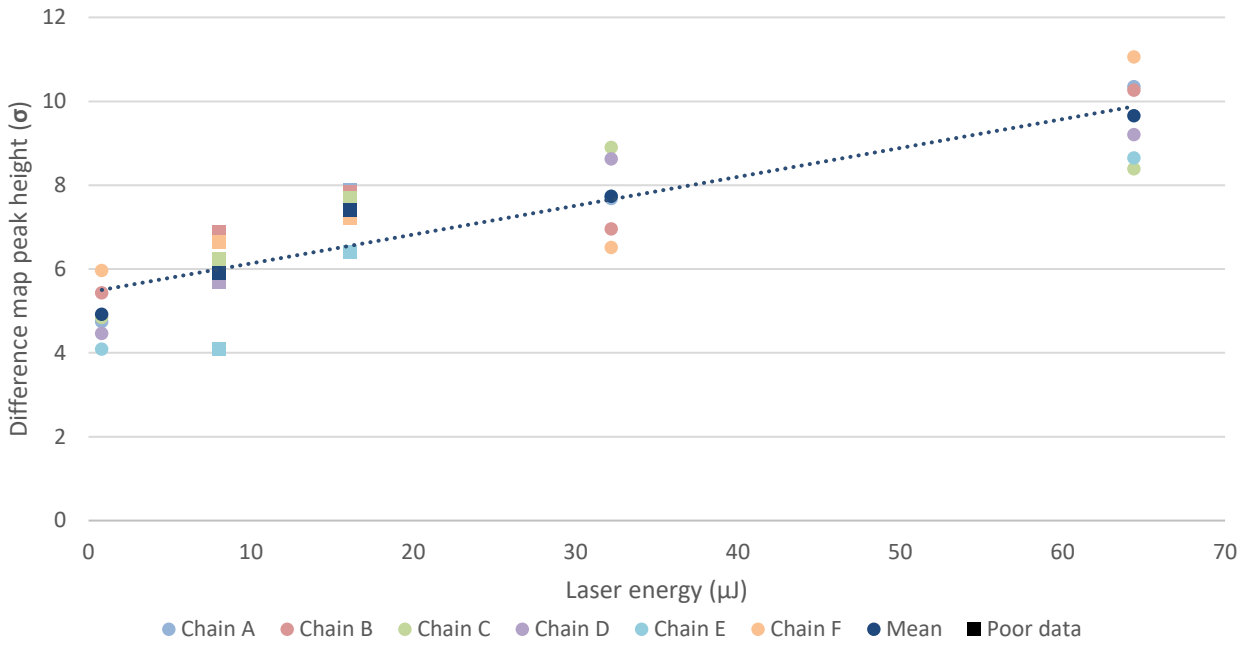

Figure 5.18: Peak heights in RMSD (σ) corresponding to NO in the $F_{obs} - F_{calc}$ difference map from DtpB structure with different laser energies. Data shown for chains A–F and the mean. Data from less reliable structures is indicated with squares (▪).

with the occupancies derived from B-factor analysis being on average 5 % higher than those from the density peak heights, and the linear trendlines of both methods are almost parallel (slope of 1.14 for B-factor method versus 1.09 for the peak height method).

5.3 Discussion

The structures collected here provide information on the binding of NO to the peroxidase DtpB. The Arg 243 residue that is seen to move upon NO binding is known to be involved in the catalytic cycle. It has access to bulk solvent via a hydrogen bonding network, facilitating proton transfer to the haem site, shown in Figure 5.19. This solvent accessibility is likely the cause of its shift in respect to the resting state structures, as it lies in the access channel to the haem site. Although binding of NO is not involved in the peroxidase catalytic cycle of the enzyme, NO is known to have signalling and regulatory roles in *Streptomyces* species via an endogenous NO cycle (Chater, 2016). NO has been shown to control the activity of two haem-based proteins, DevS and DevR, in *Streptomyces coelicolor*, which regulate genes involved in secondary metabolism (Honma *et al.*, 2021). Therefore, NO binding to DtpB could modulate the peroxidase activity of the enzyme. This could be investigated with assays of peroxidase activity in the presence of NO but is outside the scope of this thesis.

The binding of NO also is interesting in the context of bond lengths between ligands and the haem iron. Fe=O bond lengths have previously been studied in DtpB in its Compound I state, including changes associated with increasing X-ray dose (Rožman, 2022). This involved collecting a zero-dose SFX structure and an SSX dose series. The Fe=O bond lengths have an average of 1.8 Å (SD 0.1 Å) in the SFX structure, with the length being considerably less in chain A (1.7 Å). These bond lengths increase with dose, up to the point where the oxygen dissociates, which varies between chains. Oxygen in chain B dissociates at a dose less than 145 kGy, but in chain C it remains bound at 387 kGy. Dose-dependent migration of a bound water molecule away from the haem iron is also seen in the resting state of dye-type peroxidase Aa (DtpAa), also from *Streptomyces lividans*, from 2.4 Å to 3.8 Å (Ebrahim *et al.*, 2019b). As the NO-bound timeseries data presented here (§5.2.2) is also convoluted with dose, it is interesting that no changes are seen across the structures over time. The Fe–N bond lengths presented here are consistent across all chains and timepoints in a small range, 2.0–2.1 Å (Figure 5.12), and no correlation with chain or timepoint is seen between Fe–NO bond angles either (Figure 5.13). The only difference in the structures is the NO occupancy in chain E, which is consistently lower than the other chains, at 16 % compared to 40 %. This is possibly explained by lower solvent accessibility of chain E within the crystal, preventing access of NO after laser initiation. The previous work with DtpB featured no analysis of oxygen occupancy since the Compound I state was generated by an excess of hydrogen peroxide, so is likely at or near full occupancy.

Collecting several structures of NO bound to DtpB has provided another case for demonstrating laser initiation for TR-SSX at I24. These hexameric structures each feature six haem sites with NO bound, providing more datapoints to base estimates of ligand occupancy on compared to McCP-β. By collecting a series of five structures with different laser energies used to lyse the NO photocage, a clear linear relationship is shown between laser energy and ligand occupancy in the active site of the enzyme. This data can be used to guide illumination for future time-resolved studies at I24 and elsewhere. The theoretical energy required to provide single photon excitation and initiate full NO binding in DtpB is only 0.5 μJ, but, in part due to the slight excess of photocage, experimental results suggest approximately 60 μJ is required to get 80 % NO occupancy.

Although this is much higher than predicted, a moderate occupancy of 40 % is still seen near 1 µJ, which is more safely within the single excitation regime.

The required energy to initiate a reaction is greatly dependent on the system being studied and any photocage being used, however an approximate starting energy of 50 µJ is suggested, which is high enough for 75 % ligand occupancy in this system, while remaining low to avoid beam damage. As well, this data demonstrates a methodology for determining the correct laser energy to use for sufficient reaction initiation and occupancy of the ligand in the active site. As was performed during the experiments with McCP-β, starting with a higher laser energy to confirm successful reaction is advisable, after which lower laser energies can be used for repeat experiments, including collecting multiple timepoints as required for capturing the relevant reaction in the crystals. As NO occupancy of approximately 40 % is seen in DtpB even at 1 µJ, then there is latitude to be far below the optimal energy, while still seeing evidence of the ligand to indicate successful reaction initiation. This suggested starting energy agrees with the energy used in the previously published TR-SFX experiments with the NO photocage, which used a pump energy of approximately 40 µJ to illuminate the crystals (Tosha *et al.*, 2017). Even for different photoactive systems, illumination energies of a similar order of magnitude have been used, for example 200 µJ pump illumination for 100 % photodissociation of carbon monoxide from

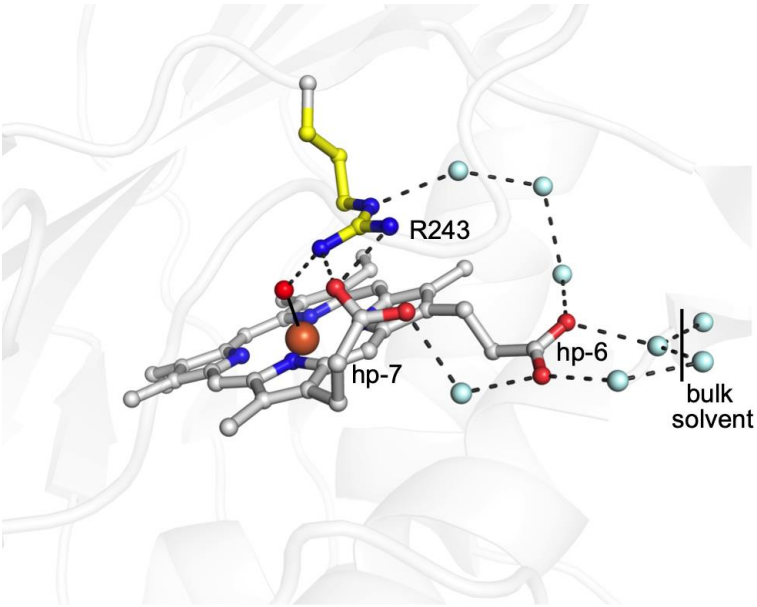

Figure 5.19: The haem site of DtpB in its compound I state (oxygen shown as red sphere), showing the hydrogen-bonded water network (pale blue) connecting Arg 243 (yellow) to the bulk solvent. Reproduced with permission from Lučić *et al.* (2020b).

cytochrome *c* oxidase (Shimada *et al.*, 2017), or 15 μJ for initiating a reaction photocycle of photoactive yellow protein (Tenboer *et al.*, 2014).

For these experiments, the light-sensitive component is a separate photocage molecule soaked into the crystals, not the protein itself. As such, the precise amount of light, in microjoules, used to initiate ligand binding is not vitally important, since excess light does not affect the protein directly. There are two main constraints to laser energy, proving upper and lower limits. The first constraint is having enough light to release the ligand in sufficient concentration within the crystals and reach a high occupancy in the active sites. At lower occupancies, elucidation of ligand binding modes may be harder, as the electron density for them is weaker. For reactions with transient intermediates, having the ligand present at higher concentrations ensures the intermediates are present in higher occupancies. Higher ligand concentrations may also aid in initiating the reaction faster once the ligand is released, thereby increasing the time resolution possible. The second upper bound on laser energy is prevention of an intense beam from damaging the crystals due to heating of the sample. The laser used for these experiments, given longer exposure times, can melt the thin films used to seal the chips, so also has potential to heat the samples themselves (Hasenaka *et al.*, 2009; Kiefersauer *et al.*, 2014; Moffat, 1998). With proteins that contain a light-sensitive cofactor, such as photosystems (Kupitz *et al.*, 2014a) or rhodopsins (Weinert *et al.*, 2019), controlling the laser energy used to initiate reactions is more important. Most light-sensitive protein reactions in their native environments occur with the absorption of a single photon per chromophore, so that is the ideal amount of light to provide to the sample. With excess light, so-called two photon processes can occur, including alternative side reactions, which can prevent a proper understanding of the protein function (Grünbein *et al.*, 2020).

Optimisation of light intensity is complicated by the optical density of the crystal, which means the intensity seen by different parts of the crystal varies. The front side of the crystal where the laser light enters receives full intensity, but as light is absorbed on its path through the crystal, the intensity at the far side of the crystal is greatly reduced (Monteiro *et al.*, 2021; Grünbein *et al.*, 2020). This problem can be reduced with crystals (including soaked photocages) that have lower optical density at the excitation wavelength. The required

illumination intensity would be higher, but proportionally less of the beam would be absorbed from front to back within the crystal. Reduced crystal size would also provide more even illumination, as the light path through the crystal to reach the rearmost unit cells would be shorter. Both measures for providing more even illumination are seen when comparing DtpB and McCP-β. The smaller crystals (10 µm compared to 20 µm) and slightly lower photocage concentration (16 mM compared to 19 mM) in DtpB mean that the transmission through the crystal is much greater (60 % compared to 16 %), and more even illumination results. Light absorption in the crystal can also be combatted by illuminating it from multiple directions, as was implemented in the previous NO photocage experiments with a viscous extruder, where a split beam illuminated the extrusion from two angles 160° apart (Tosha *et al.*, 2017). This would be complicated to execute with fixed-target chips at I24, since side-on illumination at 90° to the X-ray beam is not possible, and any optical components on the downstream side of the chip are likely to block part of the cone of diffracted X-rays from reaching the detector. Additionally, as described in §3.7, illumination is much more effective from one side of the chip than the other, because of how the crystals sit in tapered apertures.

5.3.1 Occupancy estimation

With the same structures, two methods of occupancy estimation have been demonstrated; comparing B-factors after refining at several fixed occupancies; and looking at the difference electron density peak heights of maps omitting the ligand. Because these methods are different, they can be used to cross validate each other for accuracy. The first method is more automated, relying on several rounds of structural refinement of the B-factors of the ligand atoms, which can then be compared to those of the surrounding protein. This method is useful for analysis of many datasets during model building but takes an extended time, approximately 30 minutes, to run. The second method uses the electron density peak heights corresponding to the ligand to gauge relative occupancies of ligands in several datasets. This requires more manual intervention for examining each ligand site but can quickly be performed on a dataset during beamtime, when feedback is required as to the success of an experiment and to guide the following experiments. The two methods of occupancy estimation presented here show agreement on NO occupancies for the structures with different laser energies (Figure 5.8), demonstrating consistency.

As well as providing a demonstration of the physical ability to initiate ligand binding in a crystal using light, these DtpB structures have provided data for investigation of methods for estimating occupancy of partial states within crystallography datasets. Here work has focused on a ligand-bound state but is also of relevance to other partially occupied triggered states, including activated conformations of light-sensitive proteins. Despite much work being undertaken, accurate determination of occupancies is still an open question.

One approach involves generation of difference electron density maps between observed structure factor amplitudes from a reference and a triggered state, $F_{obs}^{trig} - F_{obs}^{ref}$, which highlights structural changes between the reference and triggered states. This method is used by the program *Xtrapol8* (De Zitter *et al.*, 2022), which also produces a set of extrapolated structure factor amplitudes that represent full occupancy of the triggered state. This extrapolation is repeated in parallel for several assumed occupancy values of the observed triggered data. The extracted amplitudes are then used to generate $2F_{extr} - F_{calc}^{ref}$ and $F_{extr} - F_{calc}^{ref}$ maps, into which a model for the pure triggered state (rather than the a mixed population of states present in the crystals) is optionally built and refined. The first method that *Xtrapol8* uses to estimate occupancy involves examining peaks of the extrapolated difference maps and choosing a set of significant peaks that indicate differences between reference and triggered states. With the extrapolated difference maps from different assumed occupancies, these peaks change in height relative to the noise level of the map. The occupancy value with the highest signal to noise ratio is taken as the correct occupancy of the triggered state in the observed data. In the case that the preferred method fails, *Xtrapol8* can also determine occupancy by comparing distances between pairs of atoms in the structure which have adjacent difference map peaks. Upon refinement of the triggered structure against the extrapolated maps at different assumed occupancies, the interatomic distances change compared to the reference model. The true occupancy is determined as the average point where the changes in interatomic distances have reached to 99 % of their maximum values.

Another method for occupancy estimation involves setting the occupancy of the triggered states to a range of fixed values and refining the models against the $2F_{obs}^{trig} - F_{calc}^{trig}$ and $F_{obs}^{trig} - F_{calc}^{trig}$ maps at each occupancy. The $F_{obs}^{trig} - F_{calc}^{trig}$ difference map after refinement can be examined to determine occupancy. Where the

magnitude of difference density is lowest, the occupancy of the refined model most closely matches that of the observed data. This method is known as multicopy refinement since both the reference and triggered states exist and are refined in one model.

In work published after the original submission of this thesis Barends *et al.* (2024) used and compared several methods of occupancy estimation while examining time-resolved structures of the photodissociation of CO from myoglobin at femtosecond timescales and studying the effect of laser power on photolysis yield. The data shows movement of CO from its position as an axial ligand in the distal haem site, to a photodissociated position further from the haem. As well as analysing their experimental data, they used simulated data from mixtures of the two states to validate occupancy determination methods. Their extensive analysis includes extrapolation of structure factors for full occupancy of the photodissociated (triggered) state, similar to that used by *Xtrapol8* and described above, but did not use it for occupancy determination because it produced unrealistic structures for the extrapolated triggered state, and consistently underestimated the occupancy of the simulated data. Final occupancy estimation in the paper was performed by two methods. The first involved taking the peak height for the CO in its photodissociated position from $F_{obs}^{trig} - F_{calc}^{ref}$ omit maps that were generated without any CO modelled, and then dividing that by the sum of the peak heights for both CO positions. The second method was the multicopy approach described above, where the triggered copy of myoglobin was refined with CO in the dissociated position, while the dark state with bound CO was left unaltered. This was found to be consistent with the true occupancy for the simulated datasets, as well as with the omit map method, so was used for the final values presented in the paper.

Although the reaction Barends *et al.* (2024) were investigating involves a ligand moving between two positions, corresponding to the CO axially ligated to the haem and further from the haem in its photodissociated state, their results are relevant to the NO binding studied here, where triggering results in the ligand appearing in one position. Following the publication of the myoglobin study, additional analysis of the McCP-β and DtpB structures presented here has been performed. The process of refining structures at fixed occupancies, which was used for estimation of occupancy by comparing B-factors (see §5.2.3), is

comparable to the multi-copy refinement endorsed by Barends *et al*. The main difference between the two procedures was the way occupancy was determined after refinement. Considering this, $F_{obs}^{trig} - F_{calc}^{trig}$ maps from the NO-bound structures refined with different occupancies were examined, and the magnitude of the difference density at the NO position was recorded. As expected, magnitude decreased with increasing occupancy in the model, from a positive value at low occupancy, to a negative value above at occupancies above the true occupancy in the observed data. The point at which there is magnitude at the NO position is zero can be taken as the true occupancy, which is where the lines plotted in Figure 5.20 cross the X axis. This is similar to the B-factor analysis performed previously (§5.2.3), where the occupancy was chosen where the magnitude of the difference between B-factors was minimised (illustrated in Figure 5.14). As with the original refinement at multiple occupancies, retrieval of the difference densities was automated using Bash script. The script, included as Appendix 3, invokes *Coot* for each structure to analyse, and provides it with a Python script with the positions of the NO nitrogen. An example of a Python script for *Coot* for the 64 µJ dataset is provided as Appendix 4, and the output CSV file is shown in an abbreviated form in Appendix 5. The script was also used for the resting state structure as a comparison (purple line).

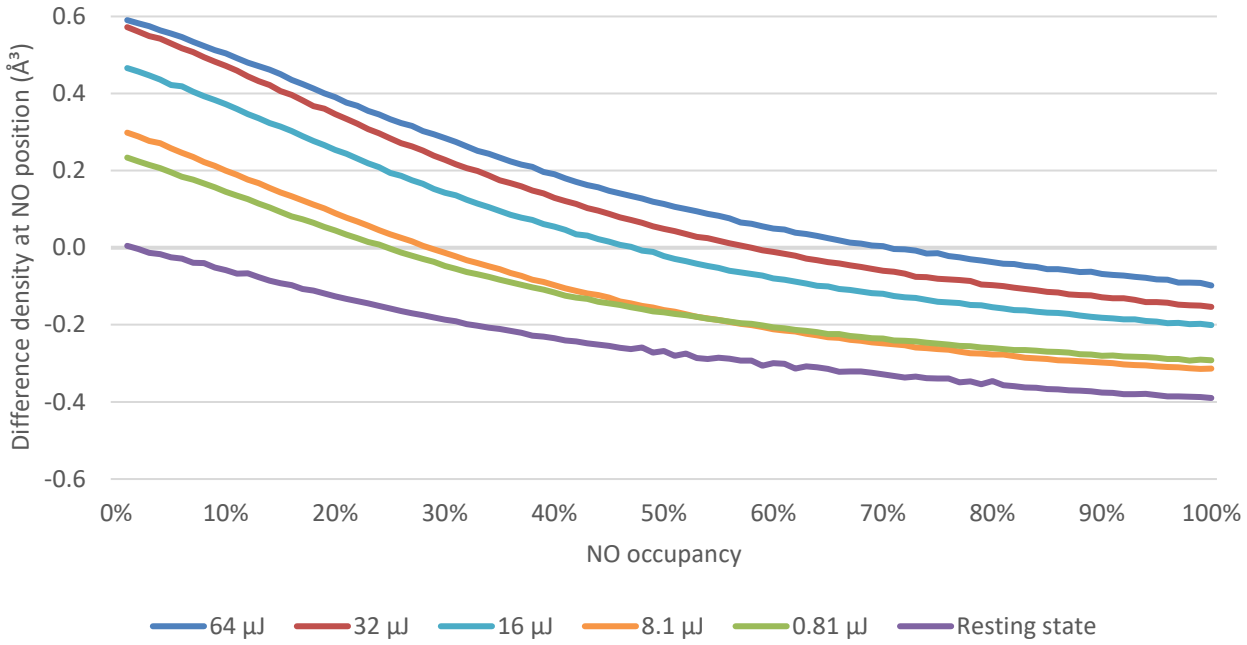

Figure 5.20: Magnitude of the $F_{obs}^{trig} - F_{calc}^{trig}$ difference density maps at the NO nitrogen position, plotted against occupancy of NO in the refined models, for TR-SSX structures with different laser energies and a resting state structure. The true occupancy is taken as the point where the lin crossed the X axis.

The occupancies determined from this method are shown in Table 5.8 and Figure 5.21, displayed alongside the mean values determined previously with the methods described in §5.2.3 and §5.2.4. The new occupancies are comparable with the previously determined values and also agree with expected results. The values are on average 13 % lower than for the B-factor comparison method, and 6 % lower than those from examining the omit map peaks (Table 5.3). The same analysis was also performed on the second set of NO-bound McCP-β structures and added to the graph in Figure 4.13. There, the data also shows agreement with the other methods, but to a lesser degree than with the DtpB data.

	NO occupancy (%)									
Laser energy (μJ)	A	B	C	D	E	F	Mean	SD	B-factors	Peak heights
64.4	78	70	57	62	71	86	71	10	84	77
32.2	55	44	65	100	56	38	60	20	58	56
16.1	61	47	47	45	38	46	47	7	65	56
8.05	29	33	24	32	27	27	29	3	50	41
0.81	29	29	16	22	17	37	25	7	39	34

Table 5.8: NO occupancies of DtpB structures at different laser energies, determined by analysis of $F_{obs}^{trig} - F_{calc}^{trig}$ difference densities after refinement at fixed occupancies. The mean values for two other methods are shown.

The correlation between the three methods for estimating occupancies provides further confidence in the validity of the methods as a way of making relative comparisons between datasets, and for providing approximate absolute occupancy values. Unfortunately, the discrepancy between the three methods prevents use of them for absolute occupancy determination to a high level of precision. Further work could simulate a series of structures at differing occupancies, which could be used to further test the methods in the same way as Barends *et al*. The three methods could also be tested with data featuring larger organic ligands such as chromophores in photoactive protein complex, or drug candidates, to see how applicable they are in other time-resolved studies and ligand binding studies, including with larger changes between states.

5.3.2 Timeseries data collection

The timeseries dataset consisting of 10 structures at different timepoints was intended to demonstrate the ability to collect several timepoints in a reaction in a time and sample efficient manner. As well as timepoints after reaction initiation with a laser, the first structure was supposed to capture a “dark” state of the protein,

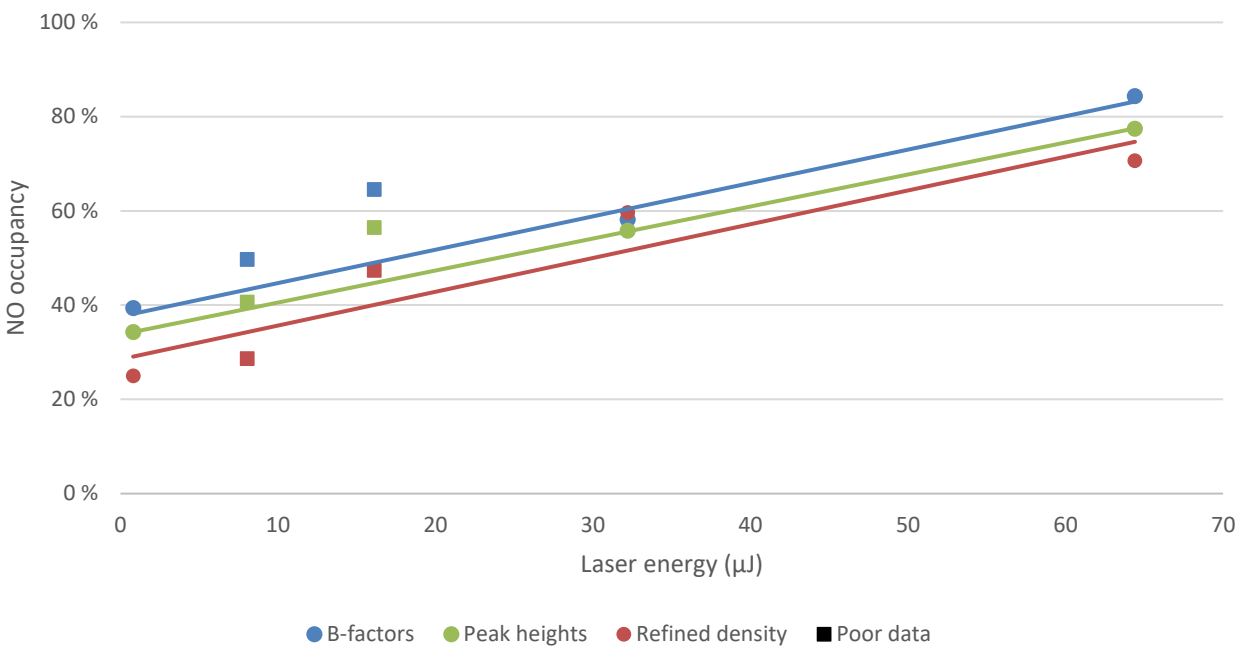

Figure 5.21: Comparison of the three methods for estimating NO occupancy for DtpB, averaged for chains A–F. The new method is shown in red. The less reliable datapoints (▪) are excluded from the trendlines.

where the crystals contain the photocage but NO has not been released and as such the haem sites remain empty. This was not the case, and NO appears in the electron density map for even the first structure. The process of NO release from the photocage, and binding at haem sites within the crystal, is expected to be very fast, taking less than 10 ms. Therefore, with a successful experiment, NO is expected to be present in the 10 ms timepoint (third structure), with occupancy remaining present across the 10–80 ms timepoints. NO may also be present in the 0 ms timepoint (second structure), as the 5 ms laser exposure occurs during the first half of this 10 ms X-ray exposure. However, the intended results would not have shown any NO present at the -10 ms timepoint, which is before the laser initiation.

There are several possible explanations for NO being present in the first structure of the timeseries. One of these is the possibility of the loaded chip containing the crystals and photocage being exposed to light prior to data collection. This could occur while transporting the chip in a box from the darkroom where it is loaded under red light to the beamline for data collection. Although ambient light is much weaker than the laser illumination intended to initiate caged substrate photolysis (especially at 308 nm), prolonged exposure (longer than the 5 ms laser exposure) could cause NO release. Effort is made to keep the beamline dark in regards to wavelengths <400 nm that the NO photocage is sensitive to (described in §1.6.4 and §3.2), including by

constructing an opaque shield to partially surround the sample position. Despite this, guaranteeing samples remain free from light contamination is hard, in part because of the personnel safety system that illuminates the experimental hutch during the search procedure, to ensure that the hutch is free from people before beam is taken.

Another cause of the accidental light exposure could have been contamination from illumination elsewhere on the chip. As the positions on one chip are visited one at a time, most crystals have at least one adjacent position (125 µm away) that has been exposed to light. If there is a discrepancy with laser alignment to the sample position where the X-rays are focused, or the laser beam is not adequately focused onto the plane the sample sits on, then light from one exposure could illuminate adjacent positions prematurely. With NO binding, this would not be detected from the datasets consisting of a single timepoint after laser exposure, as the binding is a single step process to a stable endpoint, and there are no intermediate steps. This means that there is no difference between the correct timepoint and a crystal that has been exposed to light previously, so only the dark exposure from a timeseries would allow such a light exposure error to be detected. It is possible that a mirror or focusing lens that directs the beam to the sample was shifted slightly from its correct position while changing the ND filters in the beam to adjust laser power reaching the sample.

Another possibility for unexpected presence of NO in the first timepoint is diffusion from adjacent wells. Despite the excess mother liquor being removed by suction through the apertures in the chips, a small amount of it remains around the crystals in the tapered wells. If there is enough liquid remaining to bridge the gap between wells (125 µm on centres), then NO released in solution from illumination of one well could potentially bind a crystal in an adjacent well. With wells in adjacent rows of the same block, the time delay between reaction initiation in the two wells can be up to 4.2 s, depending on their position in the row, providing time for diffusion of between wells. Diffusion of NO through the gas phase in the small space between the crystals and the sealing film is also a possibility, but this is unlikely due to the small amount of NO released, and its poor solubility hindering entry into the mother liquor surrounding the adjacent crystals.

As has been performed with McCP-β, a simple “dark” structure of DtpB could be collected, where the crystals have photocage added, but there is no laser exposure. If this remains free from NO, then ambient light exposure is not responsible for the presence of NO prior to the laser exposure in the timeseries. After this, careful repetition of the timeseries experiment may be informative as to why light was present in the first timepoint. If with careful checking of laser focus and alignment, the first timepoint remains free from NO, then an alignment issue was the likely cause. This underscores the importance of checking the laser beam position regularly during a beamtime shift, especially when adjusting the power.

6. Summary

In this thesis, the development of a method for time-resolved serial synchrotron crystallography (TR-SSX) at beamline I24 at Diamond Light Source has been documented. The time-resolved methodology builds on the fixed-target chip system previously developed at I24 (Owen *et al.*, 2017; Horrell *et al.*, 2021), which is already capable of reliable delivery of microcrystals to the X-ray beam, and of holding them there indefinitely. The chip system enables multiple types of time-resolved data collection, including efficient collection of long timepoints using the excite and visit again regime (Mehrabi *et al.*, 2020), and collecting a series of timepoints from one set of crystals via the multiple serial structures (MSS) method, which has been used for studying X-ray damage to crystals (Ebrahim, 2020; Ebrahim *et al.*, 2019a; Rožman, 2022). The fixed target chip system is portable (Sherrell *et al.*, 2015), which allows methods developed at I24 to be easily adapted to XFEL data collection.

Chapter 3 features an explanation of the rationale for using light as the method of reaction initiation (§3.2, and describes the first time-resolved experiments at I24 (§3.6). The chapter goes on to explore several areas where suboptimal conditions were identified in the initial experiments, and how improvements have been made to the methodology. The improvements have mainly focused on optimising the laser energy required for initiating reactions within the crystals, including consideration of chip orientation (§3.7), the thin film used to seal the chips (§3.9), and the process of accurately aligning the laser to the sample position. From the initial experiments, a reduction in required laser energy to get a structure with high (80 %) occupancy has been achieved. This reduction is greater than three orders of magnitude, from 1600 µJ to 10 µJ, taking the number of photons to safely within the single photon regime (Grünbein *et al.*, 2020). The capabilities developed at I24 during this work have been used for investigating other light-sensitive systems, as published by Baxter *et al.* (2022) and Hill *et al.* (2024).

With light initiation of a reaction, light intensity is an important parameter that requires controlling. Too weak illumination results in not initiating a reaction in all molecules in the crystal, and therefore collecting data from a mixed population of states, which is harder to interpret when modelling the intermediate state of the protein

during the reaction. Too strong illumination can cause multiphoton effects, which may initiate alternative reactions that are not part of the usual pathway (Grünbein *et al.*, 2020; Mehrabi and Schulz, 2023). In addition to light activation for time-resolved crystallography, parts of the same methodology described here have been used to develop a method of reaction initiation by mixing. This involves firing droplets of ligand solution at individual wells of the chip (with one side without sealing film), following the same timing scheme as the laser initiation (Shilova *et al.*, 2022; Glerup, 2022). This has been implemented at I24 and also used for experiments at PAL-XFEL (South Korea).

In chapters 4 and 5, the time-resolved methods developed for I24 have been demonstrated with the study of bacterial haem proteins that bind nitric oxide. These have made use of a photocage that releases nitric oxide upon UV illumination, which was soaked into the crystals prior to the experiments (Namiki *et al.*, 1997; Tosha *et al.*, 2017). Serial structures have been collected of the cyt *c*'-β from *Methylococcus capsulatus* (McCP-β) in its resting and NO-bound states. This provides data on the protein structure at room temperature, which has been compared to structures collected using conventional crystallography at 100 K (Adams *et al.*, 2019; Adams *et al.*, 2023). Comparisons have also been made between the resting state structure of the protein as collected with SSX at I24, and SFX at the SACLA XFEL. The first room temperature structures of the cyt *c*'-α from *Alcaligenes xylosoxidans* (AxCP) have also been collected (§4.2.4), although not in an NO-bound state. Another experiment that could be performed to further investigate NO binding to these haem proteins would flash photolysis of the Fe–NO bond using a laser pulse (Silkstone *et al.*, 2012), and capturing the structure of the resulting dissociated state using TR-SFX, as described in §4.3.1.

Time-resolved experiments have also investigated NO binding in the dye-type peroxidase DtpB from *Streptomyces lividans* (Lučić *et al.*, 2020b), including confirming NO binding by collecting the first NO-bound structures. As part of the TR-SSX methods development, a series of NO-bound DtpB structures at different laser energies have been collected, to help guide choices in laser energy when using illumination for reaction initiation. The five DtpB structures have also been used in exploration of methods for estimating ligand occupancy in protein structures. Three methods have been compared, based on difference maps omitting the

ligand, or refining the structures with ligands at fixed occupancies then either comparing B-factors or residual difference map density at the ligand sites, and broad agreement has been shown between all methods.

DtpB has been used to test timeseries data collection, where several structures are collected over time, before and after reaction initiation, in a very sample-efficient manner. Although in this case the experiment was unsuccessful, this first attempt at timeseries data collection has provided useful experience that will help with future time-resolved methods development at I24. Future work on the method development at I24 would aim to collect a successful timeseries, where the first timepoint shows the protein in its resting state, with a light-induced change shown in the later timepoints. With more sample availability, this could be applied to AxCP, which has true intermediate states as part of its distal to proximal binding mechanism (Kekilli *et al.*, 2017b).

In summary, a new methodology for time-resolved crystallography at I24 has been developed, involving laser illumination to initiate reactions. This has been demonstrated with NO binding in two bacterial haem proteins but opens the door for studying a wide range of protein systems, including both proteins that are naturally light activated, and with photocages to initiate reactions.

References

Abendroth, J., Buchko, G. W., Liew, F. N., Nguyen, J. N. and Kim, H. J. (2022) 'Structural Characterization of Cytochrome $c'_{\beta\text{-Met}}$ from an Ammonia-Oxidizing Bacterium', *Biochemistry,* 61(7), pp. 563-574.

Adams, H. R., Krewson, C., Vardanega, J. E., Fujii, S., Moreno, T., Sambongi, Y., Svistunenko, D., Paps, J., Andrew, C. R. and Hough, M. A. (2019) 'One fold, two functions: cytochrome P460 and cytochrome c′-β from the methanotroph *Methylococcus capsulatus* (Bath)', *Chemical Science,* 10(10), pp. 3031-3041.

Adams, H. R., Svistunenko, D. A., Wilson, M. T., Fujii, S., Strange, R. W., Hardy, Z. A., Vazquez, P. A., Dabritz, T., Streblow, G. J., Andrew, C. R. and Hough, M. A. (2023) 'A heme pocket aromatic quadrupole modulates gas binding to cytochrome *c*′-β: Implications for NO sensors', *Journal of Biological Chemistry,* 299(6), pp. 1-17.

Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Basle, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Marmol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J. E., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jimenez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Maly, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millan, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sanchez Rodriguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Uson, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M. and Yamashita, K. (2023) 'The CCP4 suite: integrative software for macromolecular crystallography', *Acta Crystallographica Section D,* 79(6), pp. 449-461.

Aherne, M., Lyons, J. A. and Caffrey, M. (2012) 'A fast, simple and robust protocol for growing crystals in the lipidic cubic phase', *Journal of Applied Crystallography,* 45(6), pp. 1330-1333.

Allen, J. W. A., Leach, N. and Ferguson, S. J. (2005) 'The histidine of the c-type cytochrome CXXCH haem-binding motif is essential for haem attachment by the *Escherichia coli* cytochrome c maturation (Ccm) apparatus', *The Biochemical Journal,* 389(Pt 2), pp. 587-592.

Als-Nielsen, J. and McMorrow, D. (2011) *Elements of Modern X-ray Physics.* John Wiley & Sons.

Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L. and Thornton, J. M. (2008) 'Metal ions in biological catalysis: from enzyme databases to general principles', *JBIC Journal of Biological Inorganic Chemistry,* 13(8), pp. 1205-1218.

Andrew, C. R., Petrova, O. N., Lamarre, I., Lambry, J.-C., Rappaport, F. and Negrerie, M. (2016) 'The dynamics behind the affinity: controlling heme-gas affinity via geminate recombination and heme propionate conformation in the NO carrier cytochrome *c*'', *ACS Chemical Biology,* 11(11), pp. 3191-3201.

Antonyuk, S. V., Rustage, N., Petersen, C. A., Arnst, J. L., Heyes, D. J., Sharma, R., Berry, N. G., Scrutton, N. S., Eady, R. R. and Andrew, C. R. (2011) 'Carbon monoxide poisoning is prevented by the energy costs of conformational changes in gas-binding haemproteins', *Proceedings of the National Academy of Sciences,* 108(38), pp. 15780-15785.

Aquila, A., Hunter, M. S., Doak, R. B., Kirian, R. A., Fromme, P., White, T. A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T. R. M., Barthelmess, M., Bogan, M. J., Bostedt, C., Bottin, H., Bozek, J. D., Caleman, C., Coppola, N., Davidsson, J., DePonte, D. P., Elser, V., Epp, S. W., Erk, B., Fleckenstein, H., Foucar, L., Frank, M., Fromme, R., Graafsma, H., Grotjohann, I., Gumprecht, L., Hajdu, J., Hampton, C. Y., Hartmann, A., Hartmann, R., Hau-Riege, S., Hauser, G., Hirsemann, H., Holl, P., Holton, J. M., Hömke, A., Johansson, L., Kimmel, N., Kassemeyer, S., Krasniqi, F., Kühnel, K.-U., Liang, M., Lomb, L., Malmerberg, E., Marchesini, S., Martin, A. V., Maia, F. R. N. C., Messerschmidt, M., Nass, K., Reich, C., Neutze, R., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schmidt, C., Schmidt, K. E., Schulz, J., Seibert, M. M., Shoeman, R. L., Sierra, R., Soltau, H., Starodub, D., Stellato, F., Stern, S., Strüder, L., Timneanu, N., Ullrich, J., Wang, X., Williams, G. J., Weidenspointner, G., Weierstall, U., Wunderer, C., Barty, A., Spence, J. C. H. and Chapman, H. N. (2012) 'Time-resolved protein nanocrystallography using an X-ray free-electron laser', *Optics Express,* 20(3), pp. 2706-2716.

Arciero, D. M., Hooper, A. B., Cai, M. and Timkovich, R. (1993) 'Evidence for the structure of the active site heme P460 in hydroxylamine oxidoreductase of *Nitrosomonas*', *Biochemistry,* 32(36), pp. 9370-9378.

Arslan, E., Schulz, H., Zufferey, R., Künzler, P. and Thöny-Meyer, L. (1998) 'Overproduction of the *Bradyrhizobium japonicum* c-Type Cytochrome Subunits of the cbb3Oxidase in *Escherichia coli*', *Biochemical and Biophysical Research Communications,* 251(3), pp. 744-747.

Asherie, N. (2004) 'Protein crystallization and phase diagrams', *Methods,* 34(3), pp. 266-272.

Barbarin-Bocahu, I. and Graille, M. (2022) 'The X-ray crystallography phase problem solved thanks to AlphaFold and RoseTTAFold models: a case-study report', *Acta Crystallographica Section D,* 78(4), pp. 517-531.

Barbieri, S., Murphy, L. M., Sawers, R. G., Eady, R. R. and Hasnain, S. S. (2008) 'Modulation of NO binding to cytochrome c′ by distal and proximal haem pocket residues', *JBIC Journal of Biological Inorganic Chemistry,* 13(4), pp. 531-540.

Barends, T. R. M., Foucar, L., Ardevol, A., Nass, K., Aquila, A., Botha, S., Doak, R. B., Falahati, K., Hartmann, E., Hilpert, M., Heinz, M., Hoffmann, M. C., Köfinger, J., Koglin, J. E., Kovacsova, G., Liang, M., Milathianaki, D., Lemke, H. T., Reinstein, J., Roome, C. M., Shoeman, R. L., Williams, G. J., Burghardt, I., Hummer, G., Boutet, S. and Schlichting, I. (2015) 'Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation', *Science,* 350(6259), pp. 445-450.

Barends, T. R. M., Gorel, A., Bhattacharyya, S., Schirò, G., Bacellar, C., Cirelli, C., Colletier, J.-P., Foucar, L., Grünbein, M. L., Hartmann, E., Hilpert, M., Holton, J. M., Johnson, P. J. M., Kloos, M., Knopp, G., Marekha, B., Nass, K., Nass Kovacs, G., Ozerov, D., Stricker, M., Weik, M., Doak, R. B., Shoeman, R. L., Milne, C. J., Huix-Rotllant, M., Cammarata, M. and Schlichting, I. (2024) 'Influence of pump laser fluence on ultrafast myoglobin structural dynamics', *Nature*.

Barty, A., Caleman, C., Aquila, A., Timneanu, N., Lomb, L., White, T. A., Andreasson, J., Arnlund, D., Bajt, S. and Barends, T. R. (2012) 'Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements', *Nature Photonics,* 6(1), pp. 35-40.

Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A. and Chapman, H. (2014) '*Cheetah*: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data', *Journal of Applied Crystallography,* 47(3), pp. 1118-1131.

Basa, P. N., Barr, C. A., Oakley, K. M., Liang, X. and Burdette, S. C. (2019) 'Zinc Photocages with Improved Photophysical Properties and Cell Permeability Imparted by Ternary Complex Formation', *Journal of the American Chemical Society,* 141(30), pp. 12100-12108.

Bax, A. and Clore, G. M. (2019) 'Protein NMR: Boundless opportunities', *Journal of Magnetic Resonance,* 306, pp. 187-191.

Baxter, J. M., Hutchison, C. D. M., Maghlaoui, K., Cordon-Preciado, V., Morgan, R. M. L., Aller, P., Butryn, A., Axford, D., Horrell, S., Owen, R. L., Storm, S. L. S., Devenish, N. E. and van Thor, J. J. (2022) 'Observation of Cation Chromophore Photoisomerization of a Fluorescent Protein Using Millisecond Synchrotron Serial Crystallography and Infrared Vibrational and Visible Spectroscopy', *The Journal of Physical Chemistry B,* 126(45), pp. 9288-9296.

Becker, D. F., Zhu, W. and Moxley, M. A. (2010) 'Flavin Redox Switching of Protein Functions', *Antioxidants & Redox Signaling,* 14(6), pp. 1079-1091.

Beharry, A. A. and Woolley, G. A. (2011) 'Azobenzene photoswitches for biomolecules', *Chemical Society Reviews,* 40(8), pp. 4422-4437.

Beitlich, T., Kuhnel, K., Schulze-Briese, C., Shoeman, R. L. and Schlichting, I. (2007) 'Cryoradiolytic reduction of crystalline heme proteins: analysis by UV-Vis spectroscopy and X-ray crystallography', *Journal of Synchrotron Radiation,* 14(1), pp. 11-23.

Bergmann, D. J. and Hooper, A. B. (2003) 'Cytochrome P460 of Nitrosomonas europaea', *European Journal of Biochemistry,* 270(9), pp. 1935-1941.

Bergmann, D. J., Zahn, J. A. and DiSpirito, A. A. (2000) 'Primary structure of cytochrome *c*′ of *Methylococcus capsulatus* Bath: evidence of a phylogenetic link between P460 and *c*′-type cytochromes', *Archives of Microbiology,* 173(1), pp. 29-34.

Beyerlein, K. R., Dierksmeyer, D., Mariani, V., Kuhn, M., Sarrou, I., Ottaviano, A., Awel, S., Knoska, J., Fuglerud, S., Jonsson, O., Stern, S., Wiedorn, M. O., Yefanov, O., Adriano, L., Bean, R., Burkhardt, A., Fischer, P., Heymann, M., Horke, D. A., Jungnickel, K. E. J., Kovaleva, E., Lorbeer, O., Metz, M., Meyer, J., Morgan, A., Pande, K., Panneerselvam, S., Seuring, C., Tolstikova, A., Lieske, J., Aplin, S., Roessle, M., White, T. A., Chapman, H. N., Meents, A. and Oberthuer, D. (2017) 'Mix-and-diffuse serial synchrotron crystallography', *IUCrJ,* 4(6), pp. 769-777.

Bhattacharyya, R., Dhar, J., Ghosh Dastidar, S., Chakrabarti, P. and Weiss, M. S. (2020) 'The susceptibility of disulfide bonds towards radiation damage may be explained by S...O interactions', *IUCrJ,* 7(5), pp. 825-834.

Blakeley, M. P., Langan, P., Niimura, N. and Podjarny, A. (2008) 'Neutron crystallography: opportunities, challenges, and limitations', *Current Opinion in Structural Biology,* 18(5), pp. 593-600.

Bordeleau, L. M. and Bartha, R. (1972) 'Biochemical transformations of herbicide-derived anilines: purification and characterization of causative enzymes', *Canadian Journal of Microbiology,* 18(12), pp. 1865-1871.

Botha, S., Nass, K., Barends, T. R. M., Kabsch, W., Latz, B., Dworkowski, F., Foucar, L., Panepucci, E., Wang, M., Shoeman, R. L., Schlichting, I. and Doak, R. B. (2015) 'Room-temperature serial crystallography at synchrotron X-ray sources using slowly flowing free-standing high-viscosity microstreams', *Acta Crystallographica Section D,* 71(2), pp. 387-397.

Boutet, S., Lomb, L., Williams, G. J., Barends, T. R., Aquila, A., Doak, R. B., Weierstall, U., DePonte, D. P., Steinbrener, J. and Shoeman, R. L. (2012) 'High-resolution protein structure determination by serial femtosecond crystallography', *Science,* 337(6092), pp. 362-364.

Bowler, M. W., Guijarro, M., Petitdemange, S., Baker, I., Svensson, O., Burghammer, M., Mueller-Dieckmann, C., Gordon, E. J., Flot, D., McSweeney, S. M. and Leonard, G. A. (2010) 'Diffraction cartography: applying microbeams to macromolecular crystallography sample evaluation and data collection', *Acta Crystallographica Section D: Biological Crystallography,* 66(8), pp. 855-864.

Bowman, S. E. J. and Bren, K. L. (2008) 'The chemistry and biochemistry of heme *c*: functional bases for covalent attachment', *Natural Product Reports,* 25(6), pp. 1118-1130.

Bowman, S. E. J., Bridwell-Rabb, J. and Drennan, C. L. (2016) 'Metalloprotein Crystallography: More than a Structure', *Accounts of Chemical Research,* 49(4), pp. 695-702.

Bragg, W. H. and Bragg, W. L. (1913) 'The reflection of X-rays by crystals', *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character,* 88(605), pp. 428-438.

Brändén, G. and Neutze, R. (2021) 'Advances and challenges in time-resolved macromolecular crystallography', *Science,* 373(6558), pp. eaba0954.

Brewster, A., Young, I., Lyubimov, A., Bhowmick, A. and Sauter, N. (2019) 'Processing serial crystallographic data from XFELs or synchrotrons using the *cctbx.xfel* GUI', *Computational Crystallography News,* 10(2), pp. 22-39.

Brewster, A. S., Waterman, D. G., Parkhurst, J. M., Gildea, R. J., Michels-Clark, T. M., Young, I. D., Bernstein, H. J., Winter, G., Evans, G. and Sauter, N. K. (2016) 'Processing XFEL data with cctbx.xfel and DIALS'.

Brivati, J. A., Symons, M. C. R., Tinling, D. J. A. and Williams, D. O. (1969) 'Unstable intermediates. Part LIX. Electron spin resonance studies from 4 to 77°K of hydrogen-bonded hydroxyl radicals in γ-irradiated ice', *Journal of the Chemical Society A: Inorganic, Physical, Theoretical,* (0), pp. 719-720.

Burgie, E. S., Clinger, J. A., Miller, M. D., Brewster, A. S., Aller, P., Butryn, A., Fuller, F. D., Gul, S., Young, I. D. and Pham, C. C. (2020) 'Photoreversible interconversion of a phytochrome photosensory module in the crystalline state', *Proceedings of the National Academy of Sciences,* 117(1), pp. 300-307.

Burmeister, W. P. (2000) 'Structural changes in a cryo-cooled protein crystal owing to radiation damage', *Acta Crystallographica Section D: Biological Crystallography,* 56(3), pp. 328-341.

Burnley, B. T., Afonine, P. V., Adams, P. D. and Gros, P. (2012) 'Modelling dynamics in protein crystal structures by ensemble refinement', *eLife,* 1, pp. e00311.

Butterbach-Bahl, K., Baggs, E. M., Dannenmann, M., Kiese, R. and Zechmeister-Boltenstern, S. (2013) 'Nitrous oxide emissions from soils: how well do we understand the processes and their controls?', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 368(1621), pp. 20130122.

Cabail, M. Z., Lace, P. J., Uselding, J. and Pacheco, A. A. (2002) 'Kinetic studies of the photoinitiated NO-releasing reactions of N, N'-bis-(carboxymethyl)-N, N'-dinitroso-1, 4-phenylenediamine', *Journal of Photochemistry and Photobiology A: Chemistry,* 152(1-3), pp. 109-121.

Cadenas, E. (1989) 'Biochemistry of oxygen toxicity', *Annual Review of Biochemistry,* 58(1), pp. 79-110.

Calvey, G. D., Katz, A. M., Schaffer, C. B. and Pollack, L. (2016) 'Mixing injector enables time-resolved crystallography with high hit rate at X-ray free electron lasers', *Structural Dynamics,* 3(5), pp. 054301.

Cameron, K. C., Di, H. J. and Moir, J. L. (2013) 'Nitrogen losses from the soil/plant system: a review', *Annals of Applied Biology,* 162(2), pp. 145-173.

Caranto, J. D. and Lancaster, K. M. (2017) 'Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase', *Proceedings of the National Academy of Sciences,* 114(31), pp. 8217-8222.

Carugo, O. (2018) 'How large B-factors can be in protein crystal structures', *BMC Bioinformatics,* 19(1), pp. 61.

Carugo, O. and Carugo, K. D. (2005) 'When X-rays modify the protein structure: radiation damage at work', *Trends in Biochemical Sciences,* 30(4), pp. 213-219.

Chapman, H. N., Barty, A., Bogan, M. J., Boutet, S., Frank, M., Hau-Riege, S. P., Marchesini, S., Woods, B. W., Bajt, S. and Benner, W. H. (2006) 'Femtosecond diffractive imaging with a soft-X-ray free-electron laser', *Nature Physics,* 2(12), pp. 839.

Chapman, H. N., Caleman, C. and Timneanu, N. (2014) 'Diffraction before destruction', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 369(1647), pp. 20130313.

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.-U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R. M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. and Spence, J. C. H. (2011) 'Femtosecond X-ray protein nanocrystallography', *Nature,* 470(7332), pp. 73-77.

Chater, K. F. (2016) 'Recent advances in understanding Streptomyces', *F1000Research,* 5.

Chayen, N. E. (2004) 'Turning protein crystallisation from an art into a science', *Current Opinion in Structural Biology,* 14(5), pp. 577-583.

Cheng, R. K. (2020) 'Towards an Optimal Sample Delivery Method for Serial Crystallography at XFEL', *Crystals,* 10(3), pp. 215.

Choi, P. S., Grigoryants, V. M., Abruna, H. D., Scholes, C. P. and Shapleigh, J. P. (2005) 'Regulation and function of cytochrome c' in *Rhodobacter sphaeroides* 2.4.3', *Journal of Bacteriology,* 187(12), pp. 4077-85.

Cipriani, F., Felisaz, F., Launer, L., Aksoy, J.-S., Caserotto, H., Cusack, S., Dallery, M., di-Chiaro, F., Guijarro, M., Huet, J., Larsen, S., Lentini, M., McCarthy, J., McSweeney, S., Ravelli, R., Renier, M., Taffut, C., Thompson, A., Leonard, G. A. and Walsh, M. A. (2006) 'Automation of sample mounting for macromolecular crystallography', *Acta Crystallographica Section D,* 62(10), pp. 1251-1259.

Congreve, M., de Graaf, C., Swain, N. A. and Tate, C. G. (2020) 'Impact of GPCR Structures on Drug Discovery', *Cell,* 181(1), pp. 81-91.

Conrad, C. E., Basu, S., James, D., Wang, D., Schaffer, A., Roy-Chowdhury, S., Zatsepin, N. A., Aquila, A., Coe, J. and Gati, C. (2015) 'A novel inert crystal delivery medium for serial femtosecond crystallography', *IUCrJ,* 2(4), pp. 421-430.

Cooper, C. E. (1999) 'Nitric oxide and iron proteins', *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1411(2), pp. 290-309.

Copeland, D. M., Soares, A. S., West, A. H. and Richter-Addo, G. B. (2006) 'Crystal structures of the nitrite and nitric oxide complexes of horse heart myoglobin', *Journal of Inorganic Biochemistry,* 100(8), pp. 1413-1425.

Cowan, J. A. and Nave, C. (2008) 'The optimum conditions to collect X-ray data from very small samples', *Journal of Synchrotron Radiation,* 15(5), pp. 458-462.

Cross, R., Lloyd, D., Poole, R. K. and Moir, J. W. B. (2001) 'Enzymatic Removal of Nitric Oxide Catalyzed by Cytochrome *c*' in *Rhodobacter capsulatus*', *Journal of Bacteriology,* 183(10), pp. 3050-3054.

Cusack, S., Belrhali, H., Bram, A., Burghammer, M., Perrakis, A. and Riekel, C. (1998) 'Small is beautiful: protein micro-crystallography', *Nature Structural & Molecular Biology,* 5(8s), pp. 634.

D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W. and Marsh, M. (2014) 'Microseed matrix screening for optimization in protein crystallization: what have we learned?', *Acta Crystallographica Section F: Structural Biology Communications,* 70(9), pp. 1117-1126.

Dasgupta, M., Budday, D., de Oliveira, S. H. P., Madzelan, P., Marchany-Rivera, D., Seravalli, J., Hayes, B., Sierra, R. G., Boutet, S., Hunter, M. S., Alonso-Mori, R., Batyuk, A., Wierman, J., Lyubimov, A., Brewster, A. S., Sauter, N. K., Applegate, G. A., Tiwari, V. K., Berkowitz, D. B., Thompson, M. C., Cohen, A. E., Fraser, J. S., Wall, M. E., van den Bedem, H. and Wilson, M. A. (2019) 'Mix-and-inject XFEL crystallography reveals gated conformational dynamics during enzyme catalysis', *Proceedings of the National Academy of Sciences,* 116(51), pp. 25634-25640.

de la Mora, E., Coquelle, N., Bury, C. S., Rosenthal, M., Holton, J. M., Carmichael, I., Garman, E. F., Burghammer, M., Colletier, J.-P. and Weik, M. (2020) 'Radiation damage and dose limits in serial synchrotron crystallography at cryo- and room temperatures', *Proceedings of the National Academy of Sciences,* 117(8), pp. 4142-4151.

De Zitter, E., Coquelle, N., Oeser, P., Barends, T. R. M. and Colletier, J.-P. (2022) 'Xtrapol8 enables automatic elucidation of low-occupancy intermediate-states in crystallographic studies', *Communications Biology,* 5(1), pp. 640.

Decking, W. and Abeghyan, S. and Abramian, P. and Abramsky, A. and Aguirre, A. and Albrecht, C. and Alou, P. and Altarelli, M. and Altmann, P. and Amyan, K. and Anashin, V. and Apostolov, E. and Appel, K. and Auguste, D. and

Ayvazyan, V. and Baark, S. and Babies, F. and Baboi, N. and Bak, P. and Balandin, V. and Baldinger, R. and Baranasic, B. and Barbanotti, S. and Belikov, O. and Belokurov, V. and Belova, L. and Belyakov, V. and Berry, S. and Bertucci, M. and Beutner, B. and Block, A. and Blöcher, M. and Böckmann, T. and Bohm, C. and Böhnert, M. and Bondar, V. and Bondarchuk, E. and Bonezzi, M. and Borowiec, P. and Bösch, C. and Bösenberg, U. and Bosotti, A. and Böspflug, R. and Bousonville, M. and Boyd, E. and Bozhko, Y. and Brand, A. and Branlard, J. and Briechle, S. and Brinker, F. and Brinker, S. and Brinkmann, R. and Brockhauser, S. and Brovko, O. and Brück, H. and Brüdgam, A. and Butkowski, L. and Büttner, T. and Calero, J. and Castro-Carballo, E. and Cattalanotto, G. and Charrier, J. and Chen, J. and Cherepenko, A. and Cheskidov, V. and Chiodini, M. and Chong, A. and Choroba, S. and Chorowski, M. and Churanov, D. and Cichalewski, W. and Clausen, M. and Clement, W. and Cloué, C. and Cobos, J. A. and Coppola, N. and Cunis, S. and Czuba, K. and Czwalinna, M. and D'Almagne, B. and Dammann, J. and Danared, H. and de Zubiaurre Wagner, A. and Delfs, A. and Delfs, T. and Dietrich, F. and Dietrich, T. and Dohlus, M. and Dommach, M. and Donat, A. and Dong, X. and Doynikov, N. and Dressel, M. and Duda, M. and Duda, P. and Eckoldt, H. and Ehsan, W. and Eidam, J. and Eints, F. and Engling, C. and Englisch, U. and Ermakov, A. and Escherich, K. and Eschke, J. and Saldin, E. and Faesing, M. and Fallou, A. and Felber, M. and Fenner, M. and Fernandes, B. and Fernández, J. M. and Feuker, S. and Filippakopoulos, K. and Floettmann, K. and Fogel, V. and Fontaine, M. and Francés, A. and Martin, I. F. and Freund, W. and Freyermuth, T. and Friedland, M. and Fröhlich, L. and Fusetti, M. and Fydrych, J. and Gallas, A. and García, O. and Garcia-Tabares, L. and Geloni, G. and Gerasimova, N. and Gerth, C. and Geßler, P. and Gharibyan, V. and Gloor, M. and Głowinkowski, J. and Goessel, A. and Gołębiewski, Z. and Golubeva, N. and Grabowski, W. and Graeff, W. and Grebentsov, A. and Grecki, M. and Grevsmuehl, T. and Gross, M. and Grosse-Wortmann, U. and Grünert, J. and Grunewald, S. and Grzegory, P. and Feng, G. and Guler, H. and Gusev, G. and Gutierrez, J. L. and Hagge, L. and Hamberg, M. and Hanneken, R. and Harms, E. and Hartl, I. and Hauberg, A. and Hauf, S. and Hauschildt, J. and Hauser, J. and Havlicek, J. and Hedqvist, A. and Heidbrook, N. and Hellberg, F. and Henning, D. and Hensler, O. and Hermann, T. and Hidvégi, A. and Hierholzer, M. and Hintz, H. and Hoffmann, F. and Hoffmann, M. and Hoffmann, M. and Holler, Y. and Hüning, M. and Ignatenko, A. and Ilchen, M. and Iluk, A. and Iversen, J. and Iversen, J. and Izquierdo, M. and Jachmann, L. and Jardon, N. and Jastrow, U. and Jensch, K. and Jensen, J. and Jeżabek, M. and Jidda, M. and Jin, H. and Johansson, N. and Jonas, R. and Kaabi, W. and Kaefer, D. and Kammering, R. and Kapitza, H. and Karabekyan, S. and Karstensen, S. and Kasprzak, K. and Katalev, V. and Keese, D. and Keil, B. and Kholopov, M. and Killenberger, M. and Kitaev, B. and Klimchenko, Y. and Klos, R. and Knebel, L. and Koch, A. and Koepke, M. and Köhler, S. and Köhler, W. and Kohlstrunk, N. and Konopkova, Z. and Konstantinov, A. and Kook, W. and Koprek, W. and Körfer, M. and Korth, O. and Kosarev, A. and Kosiński, K. and Kostin, D. and Kot, Y. and Kotarba, A. and Kozak, T. and Kozak, V. and Kramert, R. and Krasilnikov, M. and Krasnov, A. and Krause, B. and Kravchuk, L. and Krebs, O. and Kretschmer, R. and Kreutzkamp, J. and Kröplin, O. and Krzysik, K. and Kube, G. and Kuehn, H. and Kujala, N. and Kulikov, V. and Kuzminych, V. and La Civita, D. and Lacroix, M. and Lamb, T. and Lancetov, A. and Larsson, M. and Le Pinvidic, D. and Lederer, S. and Lensch, T. and Lenz, D. and Leuschner, A. and Levenhagen, F. and Li, Y. and Liebing, J. and Lilje, L. and Limberg, T. and Lipka, D. and List, B. and Liu, J. and Liu, S. and Lorbeer, B. and Lorkiewicz, J. and Lu, H. H. and Ludwig, F. and Machau, K. and Maciocha, W. and Madec, C. and Magueur, C. and Maiano, C. and Maksimova, I. and Malcher, K. and Maltezopoulos, T. and Mamoshkina, E. and Manschwetus, B. and Marcellini, F. and Marinkovic, G. and Martinez, T. and Martirosyan, H. and Maschmann, W. and Maslov, M. and Matheisen, A. and Mavric, U. and Meißner, J. and Meissner, K. and Messerschmidt, M. and Meyners, N. and Michalski, G. and Michelato, P. and Mildner, N. and Moe, M. and Moglia, F. and Mohr, C. and Mohr, S. and Möller, W. and Mommerz, M. and Monaco, L. and Montiel, C. and Moretti, M. and Morozov, I. and Morozov, P. and Mross, D. (2020) 'A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator', *Nature Photonics,* 14(6), pp. 391-397.

Deiters, A., Groff, D., Ryu, Y., Xie, J. and Schultz, P. G. (2006) 'A genetically encoded photocaged tyrosine', *Angewandte Chemie,* 118(17), pp. 2794-2797.

DePonte, D., Weierstall, U., Schmidt, K., Warner, J., Starodub, D., Spence, J. and Doak, R. (2008) 'Gas dynamic virtual nozzle for generation of microscopic droplet streams', *Journal of Physics D: Applied Physics,* 41(19), pp. 195505.

Dickerson, J. L. and Garman, E. F. (2019) 'The potential benefits of using higher X-ray energies for macromolecular crystallography', *Journal of Synchrotron Radiation,* 26(4), pp. 922-930.

Doak, R. B., Nass Kovacs, G., Gorel, A., Foucar, L., Barends, T. R. M., Grünbein, M. L., Hilpert, M., Kloos, M., Roome, C. M., Shoeman, R. L., Stricker, M., Tono, K., You, D., Ueda, K., Sherrell, D. A., Owen, R. L. and Schlichting, I. (2018) 'Crystallography on a chip - without the chip: sheet-on-sheet sandwich', *Acta Crystallographica Section D: Structural Biology,* 74(Pt 10), pp. 1000-1007.

Donath, T., Brandstetter, S., Cibik, L., Commichau, S., Hofer, P., Krumrey, M., Lüthi, B., Marggraf, S., Müller, P. and Schneebeli, M. (2013) 'Characterization of the PILATUS photon-counting pixel detector for X-ray energies from 1.75 keV to 60 keV', *Journal of Physics: Conference Series,* 425(6), pp. 062001.

Donath, T., Sisak Jung, D., Burian, M., Radicci, V., Zambon, P., Fitch, A. N., Dejoie, C., Zhang, B., Ruat, M., Hanfland, M., Kewish, C. M., van Riessen, G. A., Naumenko, D., Amenitsch, H., Bourenkov, G., Bricogne, G., Chari, A. and Schulze-Briese, C. (2023) 'EIGER2 hybrid-photon-counting X-ray detectors for advanced synchrotron diffraction experiments', *Journal of Synchrotron Radiation,* 30(4), pp. 723-738.

Drulyte, I., Johnson, R. M., Hesketh, E. L., Hurdiss, D. L., Scarff, C. A., Porav, S. A., Ranson, N. A., Muench, S. P. and Thompson, R. F. (2018) 'Approaches to altering particle distributions in cryo-electron microscopy sample preparation', *Acta Crystallographica Section D: Structural Biology,* 74(6), pp. 560-571.

Ebrahim, A. (2020) *Development of Serial Crystallography Methods for Synchrotrons and X-ray Free-Electron Lasers.* University of Essex.

Ebrahim, A., Appleby, M. V., Axford, D., Beale, J., Moreno-Chicano, T., Sherrell, D. A., Strange, R. W., Hough, M. A. and Owen, R. L. (2019a) 'Resolving polymorphs and radiation-driven effects in microcrystals using fixed-target serial synchrotron crystallography', *Acta Crystallographica Section D: Structural Biology,* 75(2), pp. 151-159.

Ebrahim, A., Moreno-Chicano, T., Appleby, M. V., Chaplin, A. K., Beale, J. H., Sherrell, D. A., Duyvesteyn, H. M., Owada, S., Tono, K. and Sugimoto, H. (2019b) 'Dose-resolved serial synchrotron and XFEL structures of radiation-sensitive metalloproteins', *IUCrJ,* 6(4), pp. 543-551.

Ebrahim, A., Riley, B. T., Kumaran, D., Andi, B., Fuchs, M. R., McSweeney, S. and Keedy, D. A. (2022) 'The temperature-dependent conformational ensemble of SARS-CoV-2 main protease (Mpro)', *IUCrJ,* 9(5), pp. 682-694.

Einsle, O., Andrade, S. L., Dobbek, H., Meyer, J. and Rees, D. C. (2007) 'Assignment of individual metal redox states in a metalloprotein by crystallographic refinement at multiple X-ray wavelengths', *Journal of the American Chemical Society,* 129(8), pp. 2210-2211.

Elmore, B. O., Bergmann, D. J., Klotz, M. G. and Hooper, A. B. (2007) 'Cytochromes P460 and *c*'-beta; A new family of high-spin cytochromes *c*', *FEBS Letters,* 581(5), pp. 911-916.

Elsliger, M. A., Deacon, A. M., Godzik, A., Lesley, S. A., Wooley, J., Wuthrich, K. and Wilson, I. A. (2010) 'The JCSG high-throughput structural biology pipeline', *Acta Crystallographica Section F: Structural Biology and Crystallization Communications,* 66(Pt 10), pp. 1137-1142.

Emma, P. 'First lasing of the LCLS X-ray FEL at 1.5 Å'. *Proceedings of PAC09, Vancouver*, Vancouver CA: JACoW.

Emma, P., Akre, R., Arthur, J., Bionta, R., Bostedt, C., Bozek, J., Brachmann, A., Bucksbaum, P., Coffee, R., Decker, F. J., Ding, Y., Dowell, D., Edstrom, S., Fisher, A., Frisch, J., Gilevich, S., Hastings, J., Hays, G., Hering, P., Huang, Z., Iverson, R., Loos, H., Messerschmidt, M., Miahnahri, A., Moeller, S., Nuhn, H. D., Pile, G., Ratner, D., Rzepiela, J., Schultz, D., Smith, T., Stefan, P., Tompkins, H., Turner, J., Welch, J., White, W., Wu, J., Yocky, G. and Galayda, J. (2010) 'First lasing and operation of an ångstrom-wavelength free-electron laser', *Nature Photonics,* 4(9), pp. 641-647.

Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010) 'Features and development of Coot', *Acta Crystallographica Section D: Biological Crystallography,* 66(Pt 4), pp. 486-501.

Evans, G., Axford, D. and Owen, R. L. (2011) 'The design of macromolecular crystallography diffraction experiments', *Acta Crystallographica Section D: Biological Crystallography,* 67(Pt 4), pp. 261-270.

Farah, C., Michel, L. Y. M. and Balligand, J.-L. (2018) 'Nitric oxide signalling in cardiovascular health and disease', *Nature Reviews Cardiology,* 15(5), pp. 292-316.

Fischer, M., Shoichet, B. K. and Fraser, J. S. (2015) 'One Crystal, Two Temperatures: Cryocooling Penalties Alter Ligand Binding to Transient Protein Sites', *ChemBioChem,* 16(11), pp. 1560-1564.

Flot, D., Mairs, T., Giraud, T., Guijarro, M., Lesourd, M., Rey, V., Van Brussel, D., Morawe, C., Borel, C. and Hignette, O. (2010) 'The ID23-2 structural biology microfocus beamline at the ESRF', *Journal of Synchrotron Radiation,* 17(1), pp. 107-118.

Franklin, R. E. and Gosling, R. G. (1953a) 'Molecular Configuration in Sodium Thymonucleate', *Nature,* 171(4356), pp. 740-741.

Franklin, R. E. and Gosling, R. G. (1953b) 'The structure of sodium thymonucleate fibres. I. The influence of water content', *Acta Crystallographica,* 6(8-9), pp. 673-677.

Fujii, S., Oki, H., Kawahara, K., Yamane, D., Yamanaka, M., Maruno, T., Kobayashi, Y., Masanari, M., Wakai, S., Nishihara, H., Ohkubo, T. and Sambongi, Y. (2017) 'Structural and functional insights into thermally stable cytochrome c' from a thermophile', *Protein Science,* 26(4), pp. 737-748.

Galayda, J. N. 'The LCLS-II: A high power upgrade to the LCLS', 2018-06: SLAC National Accelerator Lab., Menlo Park, CA (United States), 18-23.

Garman, E. F. (2010) 'Radiation damage in macromolecular crystallography: what is it and why should we care?', *Acta Crystallographica Section D: Biological Crystallography,* 66(4), pp. 339-351.

Garman, E. F. and Schneider, T. R. (1997) 'Macromolecular cryocrystallography', *Journal of Applied Crystallography,* 30(3), pp. 211-237.

Geremia, S., Campagnolo, M., Demitri, N. and Johnson, L. N. (2006) 'Simulation of diffusion time of small molecules in protein crystals', *Structure,* 14(3), pp. 393-400.

Gerstel, M., Deane, C. M. and Garman, E. F. (2015) 'Identifying and quantifying radiation damage at the atomic level', *Journal of Synchrotron Radiation,* 22(2), pp. 201-212.

Giacomo, P. (1984) 'News from the BIPM', *Metrologia,* 20(1), pp. 25.

Giordano, R., Leal, R. M. F., Bourenkov, G. P., McSweeney, S. and Popov, A. N. (2012) 'The application of hierarchical cluster analysis to the selection of isomorphous crystals', *Acta Crystallographica Section D,* 68(6), pp. 649-658.

Glerup, J. (2022) *Development of drop-on-chip time-resolved serial crystallography.* MChem, University of Edinburgh.

Gorel, A., Schlichting, I. and Barends, T. R. M. (2021) 'Discerning best practices in XFEL-based biological crystallography - standards for nonstandard experiments', *IUCrJ,* 8(4), pp. 532-543.

Gosling, R. (1954) *X-ray diffraction studies of Deoxyribose Nucleic Acid.* University of London.

Gotthard, G., Aumonier, S., De Sanctis, D., Leonard, G., von Stetten, D. and Royant, A. (2019) 'Specific radiation damage is a lesser concern at room temperature', *IUCrJ,* 6(4), pp. 665-680.

Gouet, P., Jouve, H.-M., Williams, P. A., Andersson, I., Andreoletti, P., Nussaume, L. and Hajdu, J. (1996) 'Ferryl intermediates of catalase captured by time-resolved Weissenberg crystallography and UV-VIS spectroscopy', *Nature Structural Biology,* 3(11), pp. 951-956.

Grünbein, M. L., Stricker, M., Nass Kovacs, G., Kloos, M., Doak, R. B., Shoeman, R. L., Reinstein, J., Lecler, S., Haacke, S. and Schlichting, I. (2020) 'Illumination guidelines for ultrafast pump–probe experiments by serial femtosecond crystallography', *Nature Methods,* 17(7), pp. 681-684.

Hajdu, J., Acharya, K. R., Stuart, D. I., McLaughlin, P. J., Barford, D., Oikonomakos, N. G., Klein, H. and Johnson, L. N. (1987) 'Catalysis in the crystal: synchrotron radiation studies with glycogen phosphorylase b', *The EMBO Journal,* 6(2), pp. 539-546.

Hamra, G. B., Laden, F., Cohen, A. J., Raaschou-Nielsen, O., Brauer, M. and Loomis, D. (2015) 'Lung Cancer and Exposure to Nitrogen Dioxide and Traffic: A Systematic Review and Meta-Analysis', *Environmental Health Perspectives,* 123(11), pp. 1107-1112.

Hara, T., Fukami, K., Inagaki, T., Kawaguchi, H., Kinjo, R., Kondo, C., Otake, Y., Tajiri, Y., Takebe, H., Togawa, K., Yoshino, T., Tanaka, H. and Ishikawa, T. (2016) 'Pulse-by-pulse multi-beam-line operation for x-ray free-electron lasers', *Physical Review Accelerators and Beams,* 19(2), pp. 020703.

Hargreaves, K. J., Fowler, D., Storeton-West, R. L. and Duyzer, J. H. (1992) 'The exchange of nitric oxide, nitrogen dioxide and ozone between pasture and the atmosphere', *Environmental Pollution,* 75(1), pp. 53-59.

Hasenaka, H., Sugiyama, S., Hirose, M., Shimizu, N., Kitatani, T., Takahashi, Y., Adachi, H., Takano, K., Murakami, S., Inoue, T., Mori, Y. and Matsumura, H. (2009) 'Femtosecond laser processing of protein crystals grown in agarose gel', *Journal of Crystal Growth,* 312(1), pp. 73-78.

Heinrich, T. A., da Silva, R. S., Miranda, K. M., Switzer, C. H., Wink, D. A. and Fukuto, J. M. (2013) 'Biological nitric oxide signalling: chemistry and terminology', *British Journal of Pharmacology,* 169(7), pp. 1417-1429.

Henderson, R. (1990) 'Cryo-protection of protein crystals against radiation damage in electron and X-ray diffraction', *Proceedings of the Royal Society of London. Series B: Biological Sciences,* 241(1300), pp. 6-8.

Hill, J. A., Nyathi, Y., Horrell, S., von Stetten, D., Axford, D., Owen, R. L., Beddard, G. S., Pearson, A. R., Ginn, H. M. and Yorke, B. A. (2024) 'An ultraviolet-driven rescue pathway for oxidative stress to eye lens protein human gamma-D crystallin', *Communications Chemistry,* 7(1), pp. 81.

Hill, M. A. and Smith, F. A. (1994) 'Calculation of initial and primary yields in the radiolysis of water', *Radiation Physics and Chemistry,* 43(3), pp. 265-280.

Hirata, K., Shinzawa-Itoh, K., Yano, N., Takemura, S., Kato, K., Hatanaka, M., Muramoto, K., Kawahara, T., Tsukihara, T. and Yamashita, E. (2014) 'Determination of damage-free crystal structure of an X-ray–sensitive protein using an XFEL', *Nature Methods,* 11(7), pp. 734-736.

Holton, J. M. (2009) 'A beginner's guide to radiation damage', *Journal of Synchrotron Radiation,* 16(2), pp. 133-142.

Honma, S., Ito, S., Yajima, S. and Sasaki, Y. (2021) 'Nitric Oxide Signaling for Actinorhodin Production in Streptomyces coelicolor A3(2) via the DevS/R Two-Component System', *Applied and Environmental Microbiology,* 87(14), pp. e00480-21.

Horrell, S., Antonyuk, S. V., Eady, R. R., Hasnain, S. S., Hough, M. A. and Strange, R. W. (2016) 'Serial crystallography captures enzyme catalysis in copper nitrite reductase at atomic resolution from one crystal', *IUCrJ,* 3(4), pp. 271-281.

Horrell, S., Axford, D., Devenish, N. E., Ebrahim, A., Hough, M. A., Sherrell, D. A., Storm, S. L. S., Tews, I., Worrall, J. A. R. and Owen, R. L. (2021) 'Fixed Target Serial Data Collection at Diamond Light Source', *Journal of Visualized Experiments,* (168), pp. e62200.

Horrell, S., Kekilli, D., Sen, K., Owen, R. L., Dworkowski, F. S. N., Antonyuk, S. V., Keal, T. W., Yong, C. W., Eady, R. R., Hasnain, S. S., Strange, R. W. and Hough, M. A. (2018) 'Enzyme catalysis captured using multiple structures from one crystal at varying temperatures', *IUCrJ,* 5(Pt 3), pp. 283-292.

Hough, M. A. and Andrew, C. R. (2015) 'Cytochromes c': structure, reactivity and relevance to haem-based gas sensing', *Advances in Microbial Physiology*: Elsevier, pp. 1-84.

Hough, M. A., Antonyuk, S. V., Barbieri, S., Rustage, N., McKay, A. L., Servid, A. E., Eady, R. R., Andrew, C. R. and Hasnain, S. S. (2011) 'Distal-to-proximal NO conversion in hemoproteins: the role of the proximal pocket', *Journal of Molecular Biology,* 405(2), pp. 395-409.

Hough, M. A., Antonyuk, S. V., Strange, R. W., Eady, R. R. and Hasnain, S. S. (2008) 'Crystallography with Online Optical and X-ray Absorption Spectroscopies Demonstrates an Ordered Mechanism in Copper Nitrite Reductase', *Journal of Molecular Biology,* 378(2), pp. 353-361.

Hough, M. A., Kekilli, D., Horrell, S., Sen, K., Yong, C., Keal, T. W., Antonyuk, S. V., Eady, R. R., Hasnain, S. S. and Strange, R. W. 'MSOX crystallography and simulations to capture redox enzyme catalysis', *Acta Crystallographica Section A: Foundations and Advances*, C651.

Hough, M. A. and Owen, R. L. (2021) 'Serial synchrotron and XFEL crystallography for studies of metalloprotein catalysis', *Current Opinion in Structural Biology,* 71, pp. 232-238.

Howard, A. J. 'Data processing in macromolecular crystallography', *Crystallographic Computing 7: Proceedings from the Macromolecular Crystallographic Computing School*.

Hunter, M. S., Segelke, B., Messerschmidt, M., Williams, G. J., Zatsepin, N. A., Barty, A., Benner, W. H., Carlson, D. B., Coleman, M. and Graf, A. (2014) 'Fixed-target protein serial microcrystallography with an x-ray free electron laser', *Scientific Reports,* 4, pp. 6026.

Hwang, C. S., Jan, J. C., Chang, C. S., Chen, S. D., Chang, C. H. and Uen, T. M. (2011) 'Development trends for insertion devices of future synchrotron light sources', *Physical Review Special Topics - Accelerators and Beams,* 14(4), pp. 044801.

Illava, G., Jayne, R., Finke, A. D., Closs, D., Zeng, W., Milano, S. K., Huang, Q., Kriksunov, I., Sidorenko, P., Wise, F. W., Zipfel, W. R., Apker, B. A. and Thorne, R. E. (2021) 'Integrated sample-handling and mounting system for fixed-target serial synchrotron crystallography', *Acta Crystallographica Section D: Structural Biology,* 77(5), pp. 628-644.

Inoue, I., Inubushi, Y., Sato, T., Tono, K., Katayama, T., Kameshima, T., Ogawa, K., Togashi, T., Owada, S. and Amemiya, Y. (2016) 'Observation of femtosecond X-ray interactions with matter using an X-ray–X-ray pump–probe scheme', *Proceedings of the National Academy of Sciences,* 113(6), pp. 1492-1497.

Ishigami, I., Lewis-Ballester, A., Echelmeier, A., Brehm, G., Zatsepin, N. A., Grant, T. D., Coe, J. D., Lisova, S., Nelson, G., Zhang, S., Dobson, Z. F., Boutet, S., Sierra, R. G., Batyuk, A., Fromme, P., Fromme, R., Spence, J. C. H., Ros, A., Yeh, S. R. and Rousseau, D. L. (2019) 'Snapshot of an oxygen intermediate in the catalytic reaction of cytochrome c oxidase', *Proceedings of the National Academy of Sciences,* 116(9), pp. 3572-3577.

Jiang, Y., Niu, Y., Xia, Y., Liu, C., Lin, Z., Wang, W., Ge, Y., Lei, X., Wang, C., Cai, J., Chen, R. and Kan, H. (2019) 'Effects of personal nitrogen dioxide exposure on airway inflammation and lung function', *Environmental Research,* 177, pp. 108620.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. and Hassabis, D. (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature,* 596(7873), pp. 583-589.

Kantardjieff, K. A. and Rupp, B. (2003) 'Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals', *Protein Science,* 12(9), pp. 1865-1871.

Kekilli, D. (2015) *Ligand discrimination mechanisms in cytochromes c prime.* Doctor of Philosophy, University of Essex.

Kekilli, D., Dworkowski, F. S., Pompidor, G., Fuchs, M. R., Andrew, C. R., Antonyuk, S., Strange, R. W., Eady, R. R., Hasnain, S. S. and Hough, M. A. (2014) 'Fingerprinting redox and ligand states in haemprotein crystal structures using resonance Raman spectroscopy', *Acta Crystallographica Section D: Biological Crystallography,* 70(5), pp. 1289-1296.

Kekilli, D., Moreno-Chicano, T., Chaplin, A. K., Horrell, S., Dworkowski, F. S. N., Worrall, J. A. R., Strange, R. W. and Hough, M. A. (2017a) 'Photoreduction and validation of haem-ligand intermediate states in protein crystals by in situ single-crystal spectroscopy and diffraction', *IUCrJ,* 4(3), pp. 263-270.

Kekilli, D., Petersen, C. A., Pixton, D. A., Ghafoor, D. D., Abdullah, G. H., Dworkowski, F. S., Wilson, M. T., Heyes, D. J., Hardman, S. J. and Murphy, L. M. (2017b) 'Engineering proximal vs. distal heme–NO coordination via dinitrosyl dynamics: implications for NO sensor design', *Chemical Science,* 8(3), pp. 1986-1994.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R., Wyckoff, H. and Phillips, D. C. (1958) 'A three-dimensional model of the myoglobin molecule obtained by x-ray analysis', *Nature,* 181(4610), pp. 662-666.

Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. and Shore, V. (1960) 'Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å. resolution', *Nature,* 185(4711), pp. 422-427.

Kern, J., Alonso-Mori, R., Tran, R., Hattne, J., Gildea, R. J., Echols, N., Glöckner, C., Hellmich, J., Laksmono, H., Sierra, R. G., Lassalle-Kaiser, B., Koroidov, S., Lampe, A., Han, G., Gul, S., DiFiore, D., Milathianaki, D., Fry, A. R., Miahnahri, A., Schafer, D. W., Messerschmidt, M., Seibert, M. M., Koglin, J. E., Sokaras, D., Weng, T.-C., Sellberg, J., Latimer, M. J., Grosse-Kunstleve, R. W., Zwart, P. H., White, W. E., Glatzel, P., Adams, P. D., Bogan, M. J., Williams, G. J., Boutet, S., Messinger, J., Zouni, A., Sauter, N. K., Yachandra, V. K., Bergmann, U. and Yano, J. (2013) 'Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature', *Science,* 340(6131), pp. 491-495.

Kiefersauer, R., Grandl, B., Krapp, S. and Huber, R. (2014) 'IR laser-induced protein crystal transformation', *Acta Crystallographica Section D,* 70(5), pp. 1224-1232.

Kim, S. J. and Shoda, M. (1999) 'Purification and Characterization of a Novel Peroxidase from *Geotrichum candidum* Dec 1 Involved in Decolorization of Dyes', *Applied and Environmental Microbiology,* 65(3), pp. 1029-1035.

Kirian, R. A., Wang, X., Weierstall, U., Schmidt, K. E., Spence, J. C. H., Hunter, M., Fromme, P., White, T., Chapman, H. N. and Holton, J. (2010) 'Femtosecond protein nanocrystallography—data analysis methods', *Optics Express,* 18(6), pp. 5713-5723.

Kirian, R. A., White, T. A., Holton, J. M., Chapman, H. N., Fromme, P., Barty, A., Lomb, L., Aquila, A., Maia, F. R. N. C., Martin, A. V., Fromme, R., Wang, X., Hunter, M. S., Schmidt, K. E. and Spence, J. C. H. (2011) 'Structure-factor analysis of femtosecond microdiffraction patterns from protein nanocrystals', *Acta Crystallographica Section A,* 67(2), pp. 131-140.

Kraft, P., Bergamaschi, A., Broennimann, C., Dinapoli, R., Eikenberry, E. F., Henrich, B., Johnson, I., Mozzanica, A., Schleputz, C. M., Willmott, P. R. and Schmitt, B. (2009) 'Performance of single-photon-counting PILATUS detector modules', *Journal of Synchrotron Radiation,* 16(3), pp. 368-375.

Krauss, U., Drepper, T. and Jaeger, K.-E. (2011) 'Enlightened Enzymes: Strategies to Create Novel Photoresponsive Proteins', *Chemistry – A European Journal,* 17(9), pp. 2552-2560.

Krissinel, E. and Henrick, K. (2004) 'Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions', *Acta Crystallographica Section D,* 60(12 Part 1), pp. 2256-2268.

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N. A., Rendek, K. N., Hunter, M. S., Shoeman, R. L., White, T. A., Wang, D., James, D., Yang, J.-H., Cobb, D. E., Reeder, B., Sierra, R. G., Liu, H., Barty, A., Aquila, A. L., Deponte, D., Kirian, R. A., Bari, S., Bergkamp, J. J., Beyerlein, K. R., Bogan, M. J., Caleman, C., Chao, T.-C., Conrad, C. E., Davis, K. M., Fleckenstein, H., Galli, L., Hau-Riege, S. P., Kassemeyer, S., Laksmono, H., Liang, M., Lomb, L., Marchesini, S., Martin, A. V., Messerschmidt, M., Milathianaki, D., Nass, K., Ros, A., Roy-Chowdhury, S., Schmidt, K., Seibert, M., Steinbrener, J., Stellato, F., Yan, L., Yoon, C., Moore, T. A., Moore, A. L., Pushkar, Y., Williams, G. J., Boutet, S., Doak, R. B., Weierstall, U., Frank, M., Chapman, H. N., Spence, J. C. H. and Fromme, P. (2014a) 'Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser', *Nature,* 513(7517), pp. 261-265.

Kupitz, C., Grotjohann, I., Conrad, C. E., Roy-Chowdhury, S., Fromme, R. and Fromme, P. (2014b) 'Microcrystallization techniques for serial femtosecond crystallography using photosystem II from *Thermosynechococcus elongatus* as a model system', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 369(1647), pp. 20130316.

Kupitz, C., Olmos, J. L., Jr., Holl, M., Tremblay, L., Pande, K., Pandey, S., Oberthür, D., Hunter, M., Liang, M., Aquila, A., Tenboer, J., Calvey, G., Katz, A., Chen, Y., Wiedorn, M. O., Knoska, J., Meents, A., Majriani, V., Norwood, T., Poudyal, I., Grant, T., Miller, M. D., Xu, W., Tolstikova, A., Morgan, A., Metz, M., Martin-Garcia, J. M., Zook, J. D., Roy-Chowdhury, S., Coe, J., Nagaratnam, N., Meza, D., Fromme, R., Basu, S., Frank, M., White, T., Barty, A., Bajt, S., Yefanov, O., Chapman, H. N., Zatsepin, N., Nelson, G., Weierstall, U., Spence, J., Schwander, P., Pollack, L., Fromme, P., Ourmazd, A., Phillips, G. N., Jr. and Schmidt, M. (2016) 'Structural enzymology using X-ray free electron lasers', *Structural Dynamics,* 4(4), pp. 044003.

Lasek, J. A. and Lajnert, R. (2022) 'On the Issues of NOx as Greenhouse Gases: An Ongoing Discussion…', *Applied Sciences,* 12(20), pp. 10429.

LaVerne, J. A. and Pimblott, S. M. (1997) 'Effect of elastic collisions on energy deposition by electrons in water', *The Journal of Physical Chemistry A,* 101(25), pp. 4504-4510.

Lawson, D. M., Stevenson, C. E. M., Andrew, C. R. and Eady, R. R. (2000) 'Unprecedented proximal binding of nitric oxide to heme: implications for guanylate cyclase', *The EMBO Journal,* 19(21), pp. 5661-5671.

Lee, D., Baek, S., Park, J., Lee, K., Kim, J., Lee, S. J., Chung, W. K., Lee, J.-L., Cho, Y. and Nam, K. H. (2019) 'Nylon mesh-based sample holder for fixed-target serial femtosecond crystallography', *Scientific Reports,* 9(1), pp. 6971.

Lee, H.-M., Larson, D. R. and Lawrence, D. S. (2009) 'Illuminating the Chemistry of Life: Design, Synthesis, and Applications of "Caged" and Related Photoresponsive Compounds', *ACS Chemical Biology,* 4(6), pp. 409-427.

Lee, K., Lee, D., Baek, S., Park, J., Lee, S. J., Park, S., Chung, W. K., Lee, J.-L., Cho, H.-S., Cho, Y. and Nam, K. H. (2020) 'Viscous-medium-based crystal support in a sample holder for fixed-target serial femtosecond crystallography', *Journal of Applied Crystallography,* 53(4), pp. 1051-1059.

Lehnert, N., Galinato, M. G. I., Paulat, F., Richter-Addo, G. B., Sturhahn, W., Xu, N. and Zhao, J. (2010) 'Nuclear Resonance Vibrational Spectroscopy Applied to [Fe (OEP)(NO)]: The Vibrational Assignments of Five-Coordinate Ferrous Heme–Nitrosyls and Implications for Electronic Structure', *Inorganic chemistry,* 49(9), pp. 4133-4148.

Leiros, H.-K. S., Timmins, J., Ravelli, R. B. G. and McSweeney, S. M. (2006) 'Is radiation damage dependent on the dose rate used during macromolecular crystallography data collection?', *Acta Crystallographica Section D,* 62(2), pp. 125-132.

Leman, J. K. and Weitzner, B. D. and Lewis, S. M. and Adolf-Bryfogle, J. and Alam, N. and Alford, R. F. and Aprahamian, M. and Baker, D. and Barlow, K. A. and Barth, P. and Basanta, B. and Bender, B. J. and Blacklock, K. and Bonet, J. and Boyken, S. E. and Bradley, P. and Bystroff, C. and Conway, P. and Cooper, S. and Correia, B. E. and Coventry, B. and Das, R. and De Jong, R. M. and DiMaio, F. and Dsilva, L. and Dunbrack, R. and Ford, A. S. and Frenz, B. and Fu, D. Y. and Geniesse, C. and Goldschmidt, L. and Gowthaman, R. and Gray, J. J. and Gront, D. and Guffy, S. and Horowitz, S. and Huang, P.-S. and Huber, T. and Jacobs, T. M. and Jeliazkov, J. R. and Johnson, D. K. and Kappel, K. and Karanicolas, J. and Khakzad, H. and Khar, K. R. and Khare, S. D. and Khatib, F. and Khramushin, A. and King, I. C. and Kleffner, R. and Koepnick, B. and Kortemme, T. and Kuenze, G. and Kuhlman, B. and Kuroda, D. and Labonte, J. W. and Lai, J. K. and Lapidoth, G. and Leaver-Fay, A. and Lindert, S. and Linsky, T. and London, N. and Lubin, J. H. and Lyskov, S. and Maguire, J. and Malmström, L. and Marcos, E. and Marcu, O. and Marze, N. A. and Meiler, J. and Moretti, R. and Mulligan, V. K. and Nerli, S. and Norn, C. and Ó'Conchúir, S. and Ollikainen, N. and Ovchinnikov, S. and Pacella, M. S. and Pan, X. and Park, H. and Pavlovicz, R. E. and Pethe, M. and Pierce, B. G. and Pilla, K. B. and Raveh, B. and Renfrew, P. D. and Burman, S. S. R. and Rubenstein, A. and Sauer, M. F. and Scheck, A. and Schief, W. and Schueler-Furman, O. and Sedan, Y. and Sevy, A. M. and Sgourakis, N. G. and Shi, L. and Siegel, J. B. and Silva, D.-A. and Smith, S. and Song, Y. and Stein, A. and Szegedy, M. and Teets, F. D. and Thyme, S. B. and Wang, R. Y.-R. and Watkins, A. and Zimmerman, L. and Bonneau, R. (2020) 'Macromolecular modeling and design in Rosetta: recent methods and frameworks', *Nature Methods,* 17(7), pp. 665-680.

Leonarski, F., Redford, S., Mozzanica, A., Lopez-Cuenca, C., Panepucci, E., Nass, K., Ozerov, D., Vera, L., Olieric, V., Buntschu, D., Schneider, R., Tinti, G., Froejdh, E., Diederichs, K., Bunk, O., Schmitt, B. and Wang, M. (2018) 'Fast and accurate data collection for macromolecular crystallography using the JUNGFRAU detector', *Nature Methods,* 15(10), pp. 799-804.

Li, B., Huang, S., Pan, Q.-Y., Li, M.-J., Zhou, H., Wang, Q.-S., Yu, F., Sun, B., Chen, J.-Q. and He, J.-H. (2018) 'New design for multi-crystal data collection at SSRF', *Nuclear Science and Techniques,* 29(2).

Li, T., Bonkovsky, H. L. and Guo, J.-t. (2011) 'Structural analysis of heme proteins: implications for design and prediction', *BMC Structural Biology,* 11(1), pp. 13.

Liew, F. N., Brandys, M. A., Biswas, S., Nguyen, J., Rahmawati, M., Nevala, M., Elmore, B. O., Hendrich, M. P. and Kim, H. (2019) 'Cytochrome c' β-Met is a Variant in the P460 Superfamily Lacking the Heme-Lysyl Crosslink: A Peroxidase Mimic Generating a Ferryl Intermediate', *Biochemistry,* 59(5), pp. 704-706.

Liu, J., Chakraborty, S., Hosseinzadeh, P., Yu, Y., Tian, S., Petrik, I., Bhagi, A. and Lu, Y. (2014) 'Metalloproteins containing cytochrome, iron–sulfur, or copper redox centers', *Chemical Reviews,* 114(8), pp. 4366-4469.

Loeliger, T., Brönnimann, C., Donath, T., Schneebeli, M., Schnyder, R. and Trüb, P. (2012) 'The new PILATUS3 ASIC with instant retrigger capability', *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, pp. 610-615.

Lučić, M., Chaplin, A. K., Moreno-Chicano, T., Dworkowski, F., Wilson, M., Svistunenko, D., Hough, M. and Worrall, J. A. (2020a) 'A subtle structural change in the distal haem pocket has a remarkable effect on tuning hydrogen peroxide reactivity in dye decolourising peroxidases from *Streptomyces lividans*', *Dalton Transactions,* 49(5), pp. 1620-1636.

Lučić, M., Svistunenko, D., Wilson, M., Chaplin, A., Davy, B., Ebrahim, A., Axford, D., Tosha, T., Sugimoto, H. and Owada, S. (2020b) 'Serial femtosecond zero dose crystallography captures a water-free distal heme site in a dye-decolourising peroxidase to reveal a catalytic role for an arginine in FeIV= O formation', *Angewandte Chemie,* 59(48), pp. 21656-21662.

Lučić, M., Wilson, M. T., Svistunenko, D. A., Owen, R. L., Hough, M. A. and Worrall, J. A. R. (2021) 'Aspartate or arginine? Validated redox state X-ray structures elucidate mechanistic subtleties of FeIV = O formation in bacterial dye-decolorizing peroxidases', *JBIC Journal of Biological Inorganic Chemistry,* 26(7), pp. 743-761.

Ludovici, C., Fröhlich, R., Vogtt, K., Mamat, B. and Lübben, M. (2002) 'Caged O2: Reaction of cytochrome *bo*$_3$ oxidase with photochemically released dioxygen from a cobalt peroxo complex', *European Journal of Biochemistry,* 269(10), pp. 2630-2637.

Lyubimov, A. Y., Murray, T. D., Koehl, A., Araci, I. E., Uervirojnangkoorn, M., Zeldin, O. B., Cohen, A. E., Soltis, S. M., Baxter, E. L. and Brewster, A. S. (2015) 'Capture and X-ray diffraction studies of protein microcrystals in a microfluidic trap array', *Acta Crystallographica Section D: Biological Crystallography,* 71(4), pp. 928-940.

Mancuso, A. P., Aquila, A., Batchelor, L., Bean, R. J., Bielecki, J., Borchers, G., Doerner, K., Giewekemeyer, K., Graceffa, R., Kelsey, O. D., Kim, Y., Kirkwood, H. J., Legrand, A., Letrun, R., Manning, B., Lopez Morillo, L., Messerschmidt, M., Mills, G., Raabe, S., Reimers, N., Round, A., Sato, T., Schulz, J., Signe Takem, C., Sikorski, M., Stern, S., Thute, P., Vagovic, P., Weinhausen, B. and Tschentscher, T. (2019) 'The Single Particles, Clusters and Biomolecules and Serial Femtosecond Crystallography instrument of the European XFEL: initial installation', *Journal of Synchrotron Radiation,* 26(Pt 3), pp. 660-676.

Mangubat-Medina, A. E. and Ball, Z. T. (2021) 'Triggering biological processes: methods and applications of photocaged peptides and proteins', *Chemical Society Reviews,* 50(18), pp. 10403-10421.

Manole, A., Kekilli, D., Svistunenko, D. A., Wilson, M. T., Dobbin, P. S. and Hough, M. A. (2015) 'Conformational control of the binding of diatomic gases to cytochrome c'', *JBIC Journal of Biological Inorganic Chemistry,* 20(4), pp. 675-686.

Martin-Garcia, J. M., Conrad, C. E., Coe, J., Roy-Chowdhury, S. and Fromme, P. (2016) 'Serial femtosecond crystallography: A revolution in structural biology', *Archives of Biochemistry and Biophysics,* 602, pp. 32-47.

Mayburd, A. L. and Kassner, R. J. (2002) 'Mechanism and biological role of nitric oxide binding to cytochrome c'', *Biochemistry,* 41(39), pp. 11582-11591.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. (2007) 'Phaser crystallographic software', *Journal of Applied Crystallography,* 40(4), pp. 658-674.

McCray, J. A. and Trentham, D. R. (1989) 'Properties and uses of photoreactive caged compounds', *Annual Review of Biophysics and Biophysical Chemistry,* 18(1), pp. 239-270.

McNicholas, S., Potterton, E., Wilson, K. S. and Noble, M. E. M. (2011) 'Presenting your structures: the CCP4mg molecular-graphics software', *Acta Crystallographica Section D: Biological Crystallography,* 67(4), pp. 386-394.

Mehrabi, P., Müller-Werkmeister, H. M., Leimkohl, J.-P., Schikora, H., Ninkovic, J., Krivokuca, S., Andriček, L., Epp, S. W., Sherrell, D. and Owen, R. L. (2020) 'The HARE chip for efficient time-resolved serial synchrotron crystallography', *Journal of Synchrotron Radiation,* 27(2), pp. 360-370.

Mehrabi, P. and Schulz, E. C. (2023) 'Sample Preparation for Time-Resolved Serial Crystallography: Practical Considerations', in Sousa, Â. & Passarinha, L. (eds.) *Advanced Methods in Structural Biology*. New York, NY: Springer US, pp. 361-379.

Mehrabi, P., Schulz, E. C., Agthe, M., Horrell, S., Bourenkov, G., von Stetten, D., Leimkohl, J.-P., Schikora, H., Schneider, T. R. and Pearson, A. R. (2019) 'Liquid application method for time-resolved analyses by serial synchrotron crystallography', *Nature Methods,* 16(10), pp. 979-982.

Meilleur, F., Myles, D. A. A. and Blakeley, M. P. (2006) 'Neutron Laue macromolecular crystallography', *European Biophysics Journal,* 35(7), pp. 611-620.

Miarzlou, D. A., Leisinger, F., Joss, D., Häussinger, D. and Seebeck, F. P. (2019) 'Structure of formylglycine-generating enzyme in complex with copper and a substrate reveals an acidic pocket for binding and activation of molecular oxygen', *Chemical science,* 10(29), pp. 7049-7058.

Michel, L. O., Sandkvist, M. and Bagdasarian, M. (1995) 'Specificity of the protein secretory apparatus: secretion of the heat-labile enterotoxin B subunit pentamers by different species of gram-bacteria', *Gene,* 152(1), pp. 41-45.

Miller, R. J. D., Paré-Labrosse, O., Sarracini, A. and Besaw, J. E. (2020) 'Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin in the multiphoton regime and biological relevance', *Nature Communications,* 11(1), pp. 1240.

Minhas, R., Bansal, Y. and Bansal, G. (2020) 'Inducible nitric oxide synthase inhibitors: A comprehensive update', *Medicinal Research Reviews,* 40(3), pp. 823-855.

Moffat, K. (1997) '[22] Laue diffraction', *Methods in enzymology*: Elsevier, pp. 433-447.

Moffat, K. (1998) 'Time-resolved crystallography', *Acta Crystallographica Section A: Foundations of Crystallography,* 54(6), pp. 833-841.

Moffatt, H. K., Chen, Y., Ng, K., McRee, D., Getzoff, E. D., Rossmann, M. G., Hajdu, J., Petsko, G. A., Cruickshank, D. W. J., Helliwell, J. R., Johnson, L. N., Moffat, K., Phillips, D. C., Cruickshank, D. W. J., Helliwell, J. R. and Johnson, L. N. (1992) 'Time-resolved crystallography: principles, problems and practice', *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences,* 340(1657), pp. 175-190.

Mokwena, K. K. and Tang, J. (2012) 'Ethylene Vinyl Alcohol: A Review of Barrier Properties for Packaging Shelf Stable Foods', *Critical Reviews in Food Science and Nutrition,* 52(7), pp. 640-650.

Monteiro, D. C. F., Amoah, E., Rogers, C. and Pearson, A. R. (2021) 'Using photocaging for fast time-resolved structural biology studies', *Acta Crystallographica Section D: Biological Crystallography,* 77(10), pp. 1218-1232.

Monteiro, D. C. F., von Stetten, D., Stohrer, C., Sans, M., Pearson, A. R., Santoni, G., van der Linden, P. and Trebbin, M. (2020) '3D-MiXD: 3D-printed X-ray-compatible microfluidic devices for rapid, low-consumption serial synchrotron crystallography data collection in flow', *IUCrJ,* 7(2), pp. 207-219.

Moreno-Chicano, T., Ebrahim, A., Axford, D., Appleby, M. V., Beale, J. H., Chaplin, A. K., Duyvesteyn, H. M. E., Ghiladi, R. A., Owada, S., Sherrell, D. A., Strange, R. W., Sugimoto, H., Tono, K., Worrall, J. A. R., Owen, R. L. and Hough, M. A. (2019) 'High-throughput structures of protein-ligand complexes at room temperature using serial femtosecond crystallography', *IUCrJ,* 6(Pt 6), pp. 1074-1085.

Morris, R., Dye, E., Axford, D., Newton, M. I., Beale, J. and Docker, P. (2019) 'Non-contact universal sample presentation for room temperature macromolecular crystallography using acoustic levitation', *Scientific Reports,* 9(1), pp. 1-10.

Mueller, C., Marx, A., Epp, S. W., Zhong, Y., Kuo, A., Balo, A. R., Soman, J., Schotte, F., Lemke, H. T., Owen, R. L., Pai, E. F., Pearson, A. R., Olson, J. S., Anfinrud, P. A., Ernst, O. P. and Dwayne Miller, R. J. (2015) 'Fixed target matrix for femtosecond time-resolved and in situ serial micro-crystallography', *Structural Dynamics,* 2(5), pp. 054302.

Mueller, M., Wang, M. and Schulze-Briese, C. (2012) 'Optimal fine ϕ-slicing for single-photon-counting pixel detectors', *Acta Crystallographica Section D,* 68(1), pp. 42-56.

Murray, J. W., Garman, E. F. and Ravelli, R. B. (2004) 'X-ray absorption by macromolecular crystals: the effects of wavelength and crystal composition on absorbed dose', *Journal of Applied Crystallography,* 37(4), pp. 513-522.

Murray, T. D., Lyubimov, A. Y., Ogata, C. M., Vo, H., Uervirojnangkoorn, M., Brunger, A. T. and Berger, J. M. (2015) 'A high-transparency, micro-patternable chip for X-ray diffraction analysis of microcrystals under native growth conditions', *Acta Crystallographica Section D: Biological Crystallography,* 71(10), pp. 1987-1997.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. and Vagin, A. A. (2011) 'REFMAC5 for the refinement of macromolecular crystal structures', *Acta Crystallographica Section D: Biological Crystallography,* 67(4), pp. 355-367.

Nakane, T., Joti, Y., Tono, K., Yabashi, M., Nango, E., Iwata, S., Ishitani, R. and Nureki, O. (2016) 'Data processing pipeline for serial femtosecond crystallography at SACLA', *Journal of Applied Crystallography,* 49(3), pp. 1035-1041.

Nam, K. H. (2020) 'Lard Injection Matrix for Serial Crystallography', *International Journal of Molecular Sciences,* 21(17), pp. 5977.

Namiki, S., Arai, T. and Fujimori, K. (1997) 'High-performance caged nitric oxide: a new molecular design, synthesis, and photochemical reaction', *Journal of the American Chemical Society,* 119(16), pp. 3840-3841.

Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A.-N., Standfuss, J., Neutze, R. and Iwata, S. (2016) 'A three-dimensional movie of structural changes in bacteriorhodopsin', *Science,* 354(6319), pp. 1552-1557.

Nass, K. (2019) 'Radiation damage in protein crystallography at X-ray free-electron lasers', *Acta Crystallographica Section D: Biological Crystallography,* 75(2), pp. 211-218.

Nass, K., Foucar, L., Barends, T. R. M., Hartmann, E., Botha, S., Shoeman, R. L., Doak, R. B., Alonso-Mori, R., Aquila, A., Bajt, S., Barty, A., Bean, R., Beyerlein, K. R., Bublitz, M., Drachmann, N., Gregersen, J., Jonsson, H. O., Kabsch, W., Kassemeyer, S., Koglin, J. E., Krumrey, M., Mattle, D., Messerschmidt, M., Nissen, P., Reinhard, L., Sitsel, O., Sokaras, D., Williams, G. J., Hau-Riege, S., Timneanu, N., Caleman, C., Chapman, H. N., Boutet, S. and Schlichting, I. (2015) 'Indications

of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams', *Journal of Synchrotron Radiation,* 22(2), pp. 225-238.

Nave, C. and Garman, E. F. (2005) 'Towards an understanding of radiation damage in cryocooled macromolecular crystals', *Journal of Synchrotron Radiation,* 12(3), pp. 257-260.

Nave, C. and Hill, M. A. (2005) 'Will reduced radiation damage occur with very small crystals?', *Journal of Synchrotron Radiation,* 12(3), pp. 299-303.

Nelson, E. (2019) *Multiple crystallographic dataset analysis to explore crystallographic occupancy and the conformational states of proteins.* University of Oxford.

Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. and Hajdu, J. (2000) 'Potential for biomolecular imaging with femtosecond X-ray pulses', *Nature,* 406(6797), pp. 752-757.

Novo, E. and Parola, M. (2008) 'Redox mechanisms in hepatic chronic wound healing and fibrogenesis', *Fibrogenesis & Tissue Repair,* 1(1), pp. 5.

Oberthuer, D., Knoška, J., Wiedorn, M. O., Beyerlein, K. R., Bushnell, D. A., Kovaleva, E. G., Heymann, M., Gumprecht, L., Kirian, R. A. and Barty, A. (2017) 'Double-flow focused liquid injector for efficient serial femtosecond crystallography', *Scientific Reports,* 7, pp. 44628.

Oghbaey, S., Sarracini, A., Ginn, H. M., Pare-Labrosse, O., Kuo, A., Marx, A., Epp, S. W., Sherrell, D. A., Eger, B. T. and Zhong, Y. (2016) 'Fixed target combined with spectral mapping: approaching 100% hit rates for serial crystallography', *Acta Crystallographica Section D: Structural Biology,* 72(8), pp. 944-955.

Ohana, J., Jacquamet, L., Joly, J., Bertoni, A., Taunier, P., Michel, L., Charrault, P., Pirocchi, M., Carpentier, P., Borel, F., Kahn, R. and Ferrer, J.-L. (2004) 'CATS: a Cryogenic Automated Transfer System installed on the beamline FIP at ESRF', *Journal of Applied Crystallography,* 37(1), pp. 72-77.

Olmos, J. L., Pandey, S., Martin-Garcia, J. M., Calvey, G., Katz, A., Knoska, J., Kupitz, C., Hunter, M. S., Liang, M., Oberthuer, D., Yefanov, O., Wiedorn, M., Heyman, M., Holl, M., Pande, K., Barty, A., Miller, M. D., Stern, S., Roy-Chowdhury, S., Coe, J., Nagaratnam, N., Zook, J., Verburgt, J., Norwood, T., Poudyal, I., Xu, D., Koglin, J., Seaberg, M. H., Zhao, Y., Bajt, S., Grant, T., Mariani, V., Nelson, G., Subramanian, G., Bae, E., Fromme, R., Fung, R., Schwander, P., Frank, M., White, T. A., Weierstall, U., Zatsepin, N., Spence, J., Fromme, P., Chapman, H. N., Pollack, L., Tremblay, L., Ourmazd, A., Phillips, G. N. and Schmidt, M. (2018) 'Enzyme intermediates captured "on the fly" by mix-and-inject serial crystallography', *BMC Biology,* 16(1), pp. 59.

Owen, R. L., Axford, D., Nettleship, J. E., Owens, R. J., Robinson, J. I., Morgan, A. W., Dore, A. S., Lebon, G., Tate, C. G., Fry, E. E., Ren, J., Stuart, D. I. and Evans, G. (2012) 'Outrunning free radicals in room-temperature macromolecular crystallography', *Acta Crystallographica Section D: Biological Crystallography,* 68(7), pp. 810-818.

Owen, R. L., Axford, D., Sherrell, D. A., Kuo, A., Ernst, O. P., Schulz, E. C., Miller, R. J. and Mueller-Werkmeister, H. M. (2017) 'Low-dose fixed-target serial synchrotron crystallography', *Acta Crystallographica Section D: Structural Biology,* 73(Pt 4), pp. 373-378.

Owen, R. L., Juanhuix, J. and Fuchs, M. (2016) 'Current advances in synchrotron radiation instrumentation for MX experiments', *Archives of Biochemistry and Biophysics,* 602, pp. 21-31.

Owen, R. L., Paterson, N., Axford, D., Aishima, J., Schulze-Briese, C., Ren, J., Fry, E. E., Stuart, D. I. and Evans, G. (2014) 'Exploiting fast detectors to enter a new dimension in room-temperature crystallography', *Acta Crystallographica Section D: Biological Crystallography,* 70(5), pp. 1248-1256.

Owen, R. L., Rudiño-Piñera, E. and Garman, E. F. (2006) 'Experimental determination of the radiation dose limit for cryocooled protein crystals', *Proceedings of the National Academy of Sciences,* 103(13), pp. 4912-4917.

Owen, R. L., Yorke, B. A., Gowdy, J. A. and Pearson, A. R. (2011) 'Revealing low-dose radiation damage using single-crystal spectroscopy', *Journal of Synchrotron Radiation,* 18(3), pp. 367-373.

Pacher, P., Beckman, J. S. and Liaudet, L. (2007) 'Nitric Oxide and Peroxynitrite in Health and Disease', *Physiological Reviews,* 87(1), pp. 315-424.

Park, J. and Nam, K. H. (2023) 'Sample Delivery Systems for Serial Femtosecond Crystallography at the PAL-XFEL', *Photonics,* 10(5), pp. 557.

Pearson, A. R., Elmore, B. O., Yang, C., Ferrara, J. D., Hooper, A. B. and Wilmot, C. M. (2007) 'The crystal structure of cytochrome P460 of *Nitrosomonas europaea* reveals a novel cytochrome fold and heme−protein cross-link', *Biochemistry,* 46(28), pp. 8340-8349.

Pearson, A. R. and Mehrabi, P. (2020) 'Serial synchrotron crystallography for time-resolved structural biology', *Current Opinion in Structural Biology,* 65, pp. 168-174.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. and North, A. (1960) 'Structure of hæmoglobin: a three-dimensional Fourier synthesis at 5.5-Å. resolution, obtained by X-ray analysis', *Nature,* 185(4711), pp. 416-422.

Petrus, M. L. C., Vijgenboom, E., Chaplin, A. K., Worrall, J. A. R., van Wezel, G. P. and Claessen, D. (2016) 'The DyP-type peroxidase DtpA is a Tat-substrate required for GlxA maturation and morphogenesis in *Streptomyces*', *Open Biology,* 6(1), pp. 150149.

Pfanzagl, V., Beale, J. H., Michlits, H., Schmidt, D., Gabler, T., Obinger, C., Djinovic-Carugo, K. and Hofbauer, S. (2020) 'X-ray induced photoreduction of heme metal centers rapidly induces active site perturbations in a protein-independent manner', *Journal of Biological Chemistry,* 295(39), pp. 13488-13501.

Pflugrath, J. (1999) 'The finer things in X-ray diffraction data collection', *Acta Crystallographica Section D: Biological Crystallography,* 55(10), pp. 1718-1725.

Phillips, J. C., Wlodawer, A., Yevitz, M. M. and Hodgson, K. O. (1976) 'Applications of synchrotron radiation to protein crystallography: preliminary results', *Proceedings of the National Academy of Sciences,* 73(1), pp. 128-132.

Pixton, D. A., Petersen, C. A., Franke, A., van Eldik, R., Garton, E. M. and Andrew, C. R. (2009) 'Activation Parameters for Heme−NO Binding in Alcaligenes xylosoxidans Cytochrome c': The Putative Dinitrosyl Intermediate Forms via a Dissociative Mechanism', *Journal of the American Chemical Society,* 131(13), pp. 4846-4853.

Pohl, E., Pradervand, C., Schneider, R., Tomizaki, T., Pauluhn, A., Chen, Q., Ingold, G., Zimoch, E. and Schulze-Briese, C. (2006) 'Facility Update: The New Protein Crystallography Beamline X10SA at the Swiss Light Source', *Synchrotron Radiation News,* 19(1), pp. 24-26.

Poulos, T. L. (2014) 'Heme Enzyme Structure and Function', *Chemical Reviews,* 114(7), pp. 3919-3962.

Priego-Capote, F. and de Castro, L. (2006) 'Ultrasound-assisted levitation: Lab-on-a-drop', *Trends in Analytical Chemistry,* 25(9), pp. 856-867.

Ramakrishnan, S., Stagno, J. R., Conrad, C. E., Ding, J., Yu, P., Bhandari, Y. R., Lee, Y.-T., Pauly, G., Yefanov, O., Wiedorn, M. O., Knoska, J., Oberthür, D., White, T. A., Barty, A., Mariani, V., Li, C., Brehm, W., Heinz, W. F., Magidson, V., Lockett, S., Hunter, M. S., Boutet, S., Zatsepin, N. A., Zuo, X., Grant, T. D., Pandey, S., Schmidt, M., Spence, J. C. H., Chapman, H. N. and Wang, Y.-X. (2021) 'Synchronous RNA conformational changes trigger ordered phase transitions in crystals', *Nature Communications,* 12(1), pp. 1762.

Reich, E. S. (2013) 'Ultimate upgrade for US synchrotron', *Nature News,* 501(7466), pp. 148.

Ren, Z. and Moffat, K. (1995) 'Quantitative Analysis of Synchrotron Laue Diffraction Patterns in Macromolecular Crystallography', *Journal of Applied Crystallography,* 28(5), pp. 461-481.

Research Collaboratory for Structural Bioinformatics (2023a) *PDB data distribution by experimental method and molecular type*. Available at: https://www.rcsb.org/stats/summary (Accessed: 2023-06-15].

Research Collaboratory for Structural Bioinformatics (2023b) *PDB Statistics: PDB Data Distribution by Molecular Weight (Entity)*. Available at: https://www.rcsb.org/stats/distribution-molecular-weight-entity (Accessed: 2023-07-22].

Robinson, A. L. (2015) 'History of Synchrotron Radiation', *Synchrotron Radiation News,* 28(4), pp. 4-9.

Rodgers, D. W. (1994) 'Cryocrystallography', *Structure,* 2(12), pp. 1135-1140.

Rossmann, M. G. and Blow, D. M. (1962) 'The detection of sub-units within the crystallographic asymmetric unit', *Acta Crystallographica,* 15(1), pp. 24-31.

Rossmann, M. G. and van Beek, C. G. (1999) 'Data processing', *Acta Crystallographica Section D: Biological Crystallography,* 55(10), pp. 1631-1640.

Rožman, M. (2022) *Mechanistic studies on Compound I and II formation in the dye-decolourising peroxidases from Streptomyces lividans.* University of Essex.

Rupp, B. (2010) *Biomolecular crystallography: principles, practice, and application to structural biology.*

Russell, H. J., Hardman, S. J. O., Heyes, D. J., Hough, M. A., Greetham, G. M., Towrie, M., Hay, S. and Scrutton, N. S. (2013) 'Modulation of ligand–heme reactivity by binding pocket residues demonstrated in cytochrome c' over the femtosecond–second temporal range', *The FEBS Journal,* 280(23), pp. 6070-6082.

Saavedra, J. E., Southan, G. J., Davies, K. M., Lundell, A., Markou, C., Hanson, S. R., Adrie, C., Hurford, W. E., Zapol, W. M. and Keefer, L. K. (1996) 'Localizing antithrombotic and vasodilatory activity with a novel, ultrafast nitric oxide donor', *Journal of Medicinal Chemistry,* 39(22), pp. 4361-4365.

Sadovski, O., Beharry, A. A., Zhang, F. and Woolley, G. A. (2009) 'Spectral Tuning of Azobenzene Photoswitches for Biological Applications', *Angewandte Chemie International Edition,* 48(8), pp. 1484-1486.

Sakaguchi, M., Kimura, T., Nishida, T., Tosha, T., Sugimoto, H., Yamaguchi, Y., Yanagisawa, S., Ueno, G., Murakami, H. and Ago, H. (2016) 'A nearly on-axis spectroscopic system for simultaneously measuring UV–visible absorption and X-ray diffraction in the SPring-8 structural genomics beamline', *Journal of Synchrotron Radiation,* 23(1), pp. 334-338.

Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. and Echols, N. (2013) 'New Python-based methods for data processing', *Acta Crystallographica Section D: Biological Crystallography,* 69(7), pp. 1274-1282.

Schlichting, I. (2015) 'Serial femtosecond crystallography: The first five years', *IUCrJ,* 2(Pt 2), pp. 246-255.

Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. F. and Petsko, G. A. (1990) 'Time-resolved X-ray crystallographic study of the conformational change in Ha-Ras p21 protein on GTP hydrolysis', *Nature,* 345(6273), pp. 309-315.

Schlichting, I., Rapp, G., John, J., Wittinghofer, A., Pai, E. F. and Goody, R. S. (1989) 'Biochemical and crystallographic characterization of a complex of c-Ha-ras p21 and caged GTP with flash photolysis', *Proceedings of the National Academy of Sciences,* 86(20), pp. 7687-7690.

Schmidt, M. (2013) 'Mix and inject: Reaction initiation by diffusion for time-resolved macromolecular crystallography', *Advances in Condensed Matter Physics,* 2013, pp. 167276.

Schmidt, M. (2020) 'Reaction Initiation in Enzyme Crystals by Diffusion of Substrate', *Crystals,* 10(2), pp. 116.

Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N., Wulff, M. and Anfinrud, P. A. (2003) 'Watching a Protein as it Functions with 150-ps Time-Resolved X-ray Crystallography', *Science,* 300(5627), pp. 1944-1947.

Schotte, F., Soman, J., Olson, J. S., Wulff, M. and Anfinrud, P. A. (2004) 'Picosecond time-resolved X-ray crystallography: probing protein function in real time', *Journal of Structural Biology,* 147(3), pp. 235-246.

Schulz, E. C., Mehrabi, P., Müller-Werkmeister, H. M., Tellkamp, F., Jha, A., Stuart, W., Persch, E., De Gasparo, R., Diederich, F. and Pai, E. F. (2018) 'The hit-and-return system enables efficient time-resolved serial synchrotron crystallography', *Nature Methods,* 15(11), pp. 901-904.

Schulz, E. C., Yorke, B. A., Pearson, A. R. and Mehrabi, P. (2022) 'Best practices for time-resolved serial synchrotron crystallography', *Acta Crystallographica Section D: Biological Crystallography,* 78(1), pp. 14-29.

Schulz, H., Fabianek, R. A., Pellicioli, E. C., Hennecke, H. and Thöny-Meyer, L. (1999) 'Heme transfer to the heme chaperone CcmE during cytochrome c maturation requires the CcmC protein, which may function independently of the ABC-transporter CcmAB', *Proceedings of the National Academy of Sciences,* 96(11), pp. 6462-6467.

Shapiro, D. A., Chapman, H. N., DePonte, D., Doak, R., Fromme, P., Hembree, G., Hunter, M., Marchesini, S., Schmidt, K. and Spence, J. (2008) 'Powder diffraction from a continuous microjet of submicrometer protein crystals', *Journal of Synchrotron Radiation,* 15(6), pp. 593-599.

Sharma, S., Ebrahim, A. and Keedy, D. A. (2023) 'Room-temperature serial synchrotron crystallography of the human phosphatase PTP1B', *Acta Crystallographica Section F,* 79(1), pp. 23-30.

Shelley, K. L. and Garman, E. F. (2022) 'Quantifying and comparing radiation damage in the Protein Data Bank', *Nature Communications,* 13(1), pp. 1314.

Sherrell, D. A., Foster, A. J., Hudson, L., Nutter, B., O'Hea, J., Nelson, S., Pare-Labrosse, O., Oghbaey, S., Miller, R. J. D. and Owen, R. L. (2015) 'A modular and compact portable mini-endstation for high-precision, high-speed fixed target serial crystallography at FEL and synchrotron sources', *Journal of Synchrotron Radiation,* 22(6), pp. 1372-1378.

Shilova, A., Aller, P., Owen, R., Bosman, R., Kamps, J., Glerup, J. and Orville, A. M. (2022) 'Sample-delivery techniques to perform serial crystallography at the XFEL Hub of Diamond Light Source', *Acta Crystallographica Section A,* 78(a2), pp. e394.

Shimada, A., Kubo, M., Baba, S., Yamashita, K., Hirata, K., Ueno, G., Nomura, T., Kimura, T., Shinzawa-Itoh, K., Baba, J., Hatano, K., Eto, Y., Miyamoto, A., Murakami, H., Kumasaka, T., Owada, S., Tono, K., Yabashi, M., Yamaguchi, Y., Yanagisawa, S., Sakaguchi, M., Ogura, T., Komiya, R., Yan, J., Yamashita, E., Yamamoto, M., Ago, H., Yoshikawa, S. and Tsukihara, T. (2017) 'A nanosecond time-resolved XFEL analysis of structural changes associated with CO release from cytochrome c oxidase', *Science Advances,* 3(7), pp. e1603042.

Shoun, H., Fushinobu, S., Jiang, L., Kim, S.-W. and Wakagi, T. (2012) 'Fungal denitrification and nitric oxide reductase cytochrome P450nor', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367(1593), pp. 1186-1194.

Sierra, R. G., Weierstall, U., Oberthuer, D., Sugahara, M., Nango, E. and Meents, A. (2018) 'Sample Delivery Techniques for Serial Crystallography', in Boutet, S., Fromme, P. & Hunter, M.S. (eds.) *X-ray Free Electron Lasers: A Revolution in Structural Biology*: Springer, pp. 109-184.

Silkstone, G., Kapetanaki, S. M., Husu, I., Vos, M. H. and Wilson, M. T. (2010) 'Nitric oxide binds to the proximal heme coordination site of the ferrocytochrome c/cardiolipin complex: formation mechanism and dynamics', *Journal of Biological Chemistry,* 285(26), pp. 19785-19792.

Silkstone, G., Kapetanaki, S. M., Husu, I., Vos, M. H. and Wilson, M. T. (2012) 'Nitric Oxide Binding to the Cardiolipin Complex of Ferric Cytochrome c', *Biochemistry,* 51(34), pp. 6760-6766.

Sillman, S. (1999) 'The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments', *Atmospheric Environment,* 33(12), pp. 1821-1845.

Simpkin, A. J., Thomas, J. M. H., Simkovic, F., Keegan, R. M. and Rigden, D. J. (2019) 'Molecular replacement using structure predictions from databases', *Acta Crystallographica Section D: Structural Biology,* 75(12), pp. 1051-1062.

Sliz, P., Harrison, S. C. and Rosenbaum, G. (2003) 'How does radiation damage in protein crystals depend on X-ray dose?', *Structure,* 11(1), pp. 13-19.

Smith, N. and Wilson, M. A. (2022) 'Understanding Cysteine Chemistry Using Conventional and Serial X-ray Protein Crystallography', *Crystals,* 12(11), pp. 1671.

Solem, J. C. (1986) 'Imaging biological specimens with high-intensity soft X-rays', *Journal of the Optical Society of America B,* 3(11), pp. 1551-1565.

Southworth-Davies, R. J., Medina, M. A., Carmichael, I. and Garman, E. F. (2007) 'Observation of decreased radiation damage at higher dose rates in room temperature protein crystallography', *Structure,* 15(12), pp. 1531-1541.

Stagno, J. R., Bhandari, Y. R., Conrad, C. E., Liu, Y. and Wang, Y.-X. (2017) 'Real-time crystallographic studies of the adenine riboswitch using an X-ray free-electron laser', *The FEBS Journal,* 284(20), pp. 3374-3380.

Standfuss, J. (2017) 'Time-resolved serial crystallography of bacteriorhodopsin using synchrotrons and X-ray lasers', *Acta Crystallographica Section A,* 73(a2), pp. C1049.

Stein, L. Y. (2011) 'Surveying N_2O-producing pathways in bacteria', *Methods in Enzymology*: Elsevier, pp. 131-152.

Stoddard, B. L. and Farber, G. K. (1995) 'Direct measurement of reactivity in the protein crystal by steady-state kinetic studies', *Structure,* 3(10), pp. 991-996.

Stohrer, C., Horrell, S., Meier, S., Sans, M., von Stetten, D., Hough, M., Goldman, A., Monteiro, D. C. and Pearson, A. R. (2021) 'Homogeneous batch micro-crystallization of proteins from ammonium sulfate', *Acta Crystallographica Section D: Structural Biology,* 77(2), pp. 194-204.

Storm, S. L. S., Axford, D. and Owen, R. L. (2021) 'Experimental evidence for the benefits of higher X-ray energies for macromolecular crystallography', *IUCrJ,* 8(6), pp. 896-904.

Storm, S. L. S., Crawshaw, A. D., Devenish, N. E., Bolton, R., Hall, D. R., Tews, I. and Evans, G. (2020) 'Measuring energy-dependent photoelectron escape in microcrystals', *IUCrJ,* 7(1), pp. 129-135.

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H. and Shen, J.-R. (2015) 'Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses', *Nature,* 517(7532), pp. 99-103.

Suga, M., Akita, F., Yamashita, K., Nakajima, Y., Ueno, G., Li, H., Yamane, T., Hirata, K., Umena, Y., Yonekura, S., Yu, L.-J., Murakami, H., Nomura, T., Kimura, T., Kubo, M., Baba, S., Kumasaka, T., Tono, K., Yabashi, M., Isobe, H., Yamaguchi, K., Yamamoto, M., Ago, H. and Shen, J.-R. (2019) 'An oxyl/oxo mechanism for oxygen-oxygen coupling in PSII revealed by an x-ray free-electron laser', *Science,* 366(6463), pp. 334-338.

Sugahara, M., Nakane, T., Masuda, T., Suzuki, M., Inoue, S., Song, C., Tanaka, R., Nakatsu, T., Mizohata, E., Yumoto, F., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Numata, K., Nango, E. and Iwata, S. (2017) 'Hydroxyethyl cellulose matrix applied to serial crystallography', *Scientific Reports,* 7(1), pp. 703.

Sun, Z., Liu, Q., Qu, G., Feng, Y. and Reetz, M. T. (2019) 'Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability', *Chemical Reviews,* 119(3), pp. 1626-1665.

Tahirov, T. H., Misaki, S., Meyer, T. E., Cusanovich, M. A., Higuchi, Y. and Yasuoka, N. (1996) 'Concerted movement of side chains in the haem vicinity observed on ligand binding in cytochrome c′ from Rhodobactercapsulatus', *Nature Structural Biology,* 3(5), pp. 459-464.

Taylor, G. L. (2003) 'The phase problem', *Acta Crystallographica Section D: Biological Crystallography,* 59(11), pp. 1881-1890.

Taylor, G. L. (2010) 'Introduction to phasing', *Acta Crystallographica Section D: Biological Crystallography,* 66(4), pp. 325-338.

Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G. J. and Koglin, J. E. (2014) 'Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein', *Science,* 346(6214), pp. 1242-1246.

Togashi, T., Owada, S., Kubota, Y., Sueda, K., Katayama, T., Tomizawa, H., Yabuuchi, T., Tono, K. and Yabashi, M. (2020) 'Femtosecond Optical Laser System with Spatiotemporal Stabilization for Pump-Probe Experiments at SACLA', *Applied Sciences,* 10(21), pp. 7934.

Tomlinson, E. J. and Ferguson, S. J. (2000) 'Conversion of a *c* type cytochrome to a *b* type that spontaneously forms *in vitro* from apo protein and heme: Implications for *c* type cytochrome biogenesis and folding', *Proceedings of the National Academy of Sciences,* 97(10), pp. 5156-5160.

Tono, K., Nango, E., Sugahara, M., Song, C., Park, J., Tanaka, T., Tanaka, R., Joti, Y., Kameshima, T., Ono, S., Hatsui, T., Mizohata, E., Suzuki, M., Shimamura, T., Tanaka, Y., Iwata, S. and Yabashi, M. (2015) 'Diverse application platform for hard X-ray diffraction in SACLA (DAPHNIS): application to serial protein crystallography using an X-ray free-electron laser', *Journal of Synchrotron Radiation,* 22(3), pp. 532-537.

Tono, K., Togashi, T., Inubushi, Y., Sato, T., Katayama, T., Ogawa, K., Ohashi, H., Kimura, H., Takahashi, S., Takeshita, K., Tomizawa, H., Goto, S., Ishikawa, T. and Yabashi, M. (2013) 'Beamline, experimental stations and photon beam diagnostics for the hard x-ray free electron laser of SACLA', *New Journal of Physics,* 15(8), pp. 083035.

Tosha, T., Nomura, T., Nishida, T., Saeki, N., Okubayashi, K., Yamagiwa, R., Sugahara, M., Nakane, T., Yamashita, K. and Hirata, K. (2017) 'Capturing an initial intermediate during the P450nor enzymatic reaction using time-resolved XFEL crystallography and caged-substrate', *Nature Communications,* 8(1), pp. 1585.

Trueblood, K., Bürgi, H.-B., Burzlaff, H., Dunitz, J., Gramaccioli, C., Schulz, H., Shmueli, U. and Abrahams, S. (1996) 'Atomic dispacement parameter nomenclature. Report of a subcommittee on atomic displacement parameter nomenclature', *Acta Crystallographica Section A: Foundations of Crystallography,* 52(5), pp. 770-781.

Uervirojnangkoorn, M., Zeldin, O. B., Lyubimov, A. Y., Hattne, J., Brewster, A. S., Sauter, N. K., Brunger, A. T. and Weis, W. I. (2015) 'Enabling X-ray free electron laser crystallography for challenging biological systems from a limited number of crystals', *eLife,* 4, pp. e05421.

Urzhumtsev, A. G. and Lunin, V. Y. (2019) 'Introduction to crystallographic refinement of macromolecular atomic models', *Crystallography Reviews,* 25(3), pp. 164-262.

Vagin, A. and Teplyakov, A. (1997) 'MOLREP: an Automated Program for Molecular Replacement', *Journal of Applied Crystallography,* 30(6), pp. 1022-1025.

Vagin, A. and Teplyakov, A. (2010) 'Molecular replacement with MOLREP', *Acta Crystallographica Section D: Biological Crystallography,* 66(1), pp. 22-25.

Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. and Murshudov, G. N. (2004) 'REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use', *Acta Crystallographica Section D: Biological Crystallography,* 60(12), pp. 2184-2195.

Warkentin, M. A., Atakisi, H., Hopkins, J. B., Walko, D. and Thorne, R. E. (2017) 'Lifetimes and spatio-temporal response of protein crystals in intense X-ray microbeams', *IUCrJ,* 4(6), pp. 785-794.

Warren, A. J., Axford, D., Paterson, N. G. and Owen, R. L. (2016) 'Exploiting Microbeams for Membrane Protein Structure Determination', in Moraes, I. (ed.) *The Next Generation in Membrane Protein Structure Determination*: Springer International Publishing, pp. 105-117.

Watanabe, A., Hasegawa, K. and Abe, Y. (2018) 'Contactless fluid manipulation in air: Droplet coalescence and active mixing by acoustic levitation', *Scientific Reports,* 8(1), pp. 1-8.

Waterman, D. G., Winter, G., Parkhurst, J. M., Fuentes-Montero, L., Hattne, J., Brewster, A., Sauter, N. K., Evans, G. and Rosenstrom, P. (2013) 'The DIALS framework for integration software', *CCP4 Newslett. Protein Crystallogr,* 49, pp. 13-15.

Watson, J. D. and Crick, F. (1953) 'Molecular structure of nucleic acids', *Nature,* 171(4356), pp. 737-738.

Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C. H., Bruce Doak, R., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Won Han, G., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G. J., Koglin, J. E., Marvin Seibert, M., Klinker, M., Gati, C., Shoeman, R. L., Barty, A., Chapman, H. N., Kirian, R. A., Beyerlein, K. R., Stevens, R. C., Li, D., Shah, S. T. A., Howe, N., Caffrey, M. and Cherezov, V. (2014) 'Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography', *Nature Communications,* 5(1), pp. 3309.

Weik, M., Berges, J., Raves, M. L., Gros, P., McSweeney, S., Silman, I., Sussman, J. L., Houee-Levin, C. and Ravelli, R. B. G. (2002) 'Evidence for the formation of disulfide radicals in protein crystals upon X-ray irradiationPresented at the

`Second International Workshop on Radiation Damage to Crystalline Biological Samples' held at Advanced Photon Source, Chicago, USA, in December 2001', *Journal of Synchrotron Radiation,* 9(6), pp. 342-346.

Weinert, T., Skopintsev, P., James, D., Dworkowski, F., Panepucci, E., Kekilli, D., Furrer, A., Brünle, S., Mous, S. and Ozerov, D. (2019) 'Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography', *Science,* 365(6448), pp. 61-65.

Weise, H. and Decking, W. (2017) 'Commissioning and first lasing of the European XFEL', *Proc. FEL'17*, pp. 9-13.

White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A. and Chapman, H. N. (2012) 'CrystFEL: a software suite for snapshot serial crystallography', *Journal of applied crystallography,* 45(2), pp. 335-341.

White, T. A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K. R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C. H., Diederichs, K. and Chapman, H. N. (2016) 'Recent developments in CrystFEL', *Journal of Applied Crystallography,* 49(2), pp. 680-689.

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J. and Chen, V. B. (2018) 'MolProbity: More and better reference data for improved all-atom structure validation', *Protein Science,* 27(1), pp. 293-315.

Winiwarter, W. and Klimont, Z. (2011) 'The role of N-gases (N2O, NOx, NH3) in cost-effective strategies to reduce greenhouse gas emissions and air pollution in Europe', *Current Opinion in Environmental Sustainability,* 3(5), pp. 438-445.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. and McCoy, A. (2011) 'Overview of the CCP4 suite and current developments', *Acta Crystallographica Section D: Biological Crystallography,* 67(4), pp. 235-242.

Winter, G. (2010) 'xia2: an expert system for macromolecular crystallography data reduction', *Journal of Applied Crystallography,* 43(1), pp. 186-190.

Winter, G., Beilsten-Edmands, J., Devenish, N., Gerstel, M., Gildea, R. J., McDonagh, D., Pascal, E., Waterman, D. G., Williams, B. H. and Evans, G. (2022) 'DIALS as a toolkit', *Protein Science,* 31(1), pp. 232-250.

Winter, G., Lobley, C. M. C. and Prince, S. M. (2013) 'Decision making in xia2', *Acta Crystallographica Section D,* 69(7), pp. 1260-1273.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T. and Young, I. D. (2018) '*DIALS*: implementation and evaluation of a new integration package', *Acta Crystallographica Section D: Structural Biology,* 74(2), pp. 85-97.

Wlodawer, A. (2017) 'Stereochemistry and Validation of Macromolecular Structures', in Wlodawer, A., Dauter, Z. & Jaskolski, M. (eds.) *Protein Crystallography: Methods and Protocols*. New York, NY: Springer New York, pp. 595-610.

Wu, N., Deiters, A., Cropp, T. A., King, D. and Schultz, P. G. (2004) 'A Genetically Encoded Photocaged Amino Acid', *Journal of the American Chemical Society,* 126(44), pp. 14306-14307.

Wulff, M., Schotte, F., Naylor, G., Bourgeois, D., Moffat, K. and Mourou, G. (1997) 'Time-resolved structures of macromolecules at the ESRF: Single-pulse Laue diffraction, stroboscopic data collection and femtosecond flash photolysis', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* 398(1), pp. 69-84.

Yabashi, M., Tanaka, H., Tono, K. and Ishikawa, T. (2017) 'Status of the SACLA Facility', *Applied Sciences,* 7(6), pp. 604.

Yano, J., Kern, J., Irrgang, K.-D., Latimer, M. J., Bergmann, U., Glatzel, P., Pushkar, Y., Biesiadka, J., Loll, B. and Sauer, K. (2005) 'X-ray damage to the Mn_4Ca complex in single crystals of photosystem II: a case study for metalloprotein crystallography', *Proceedings of the National Academy of Sciences,* 102(34), pp. 12047-12052.

Yoshimura, T., Suzuki, S., Nakahara, A., Iwasaki, H., Masuko, M. and Matsubara, T. (1986) 'Spectral properties of nitric oxide complexes of cytochrome *c*' from *Alcaligenes*nassNCIB 11015', *Biochemistry,* 25(9), pp. 2436-2442.

Zakurdaeva, O. A., Nesterov, S. V. and Feldman, V. I. (2005) 'Radiolysis of Aqueous Solutions of Poly(ethylene oxide) at 77 K', *High Energy Chemistry,* 39(4), pp. 201-206.

Zambon, P., Radicci, V., Trueb, P., Disch, C., Rissi, M., Sakhelashvili, T., Schneebeli, M. and Broennimann, C. (2018) 'Spectral response characterization of CdTe sensors of different pixel size with the IBEX ASIC', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* 892, pp. 106-113.

Zarrine-Afsar, A., Barends, T. R., Müller, C., Fuchs, M. R., Lomb, L., Schlichting, I. and Miller, R. D. (2012) 'Crystallography on a chip', *Acta Crystallographica Section D: Biological Crystallography,* 68(3), pp. 321-323.

Zhang, Z., Sauter, N. K., van den Bedem, H., Snell, G. and Deacon, A. M. (2006) 'Automated diffraction image analysis and spot searching for high-throughput crystal screening', *Journal of Applied Crystallography,* 39(1), pp. 112-119.

Appendices

1. Occupancy refinement script

```
#!/bin/bash
#takes input PDB, MTZ, and restraint cif files, makes 101 copies with NO
occupancies in range 0-1, then runs REFMAC using them

#reports time taken to run
start=$(date +%s)

#copies stdout and stderr to occ_refine.log
echo "" > occ_refine.log
exec > >(tee -i occ_refine.log)
exec 2>&1

#runs REFMAC5 for 10 cycles with default parameters
#If testing, change $cyc (number of cycles) to 1 for speed, change back to 10 for
actual processing
refmac () {
   cyc=10
   echo "Refining structure with NO occupancy=$o with $cyc cycle(s) of
refinement."
   refmac5 > "refmac${o}.log" \
      XYZIN $xyzin \
      HKLIN $hklin \
      XYZOUT $xyzout \
      HKLOUT $hklout \
      <<XXX
      MAKE NEWLIGAND CONTINUE
      NCYC $cyc
      @../$rstin
      MONI DIST 100000
      END
XXX
   echo "Refined structure with NO occupancy = $o"
}

#function for creating directories, modifying occupancies, and running refmac
#makes and enters directory for each occupancy
#edits NO occupancy for all chains and creates new PDB file
#runs refmac function defined above
process () {
   mkdir "occ$o"
   cd "occ$o"
   xyzin="occ${o}_$(echo ${pdbin} | sed 's:.*/::')"
        hklin=${mtzin}
        xyzout="REFMAC_${xyzin}"
        hklout="REFMAC_occ${o}_$(echo ${mtzin} | sed 's:.*/::')"
   if [ ! -f "$xyzout" ]
   then
      sed -E "/HETATM/{/NO/{s/(^.{56}).{4}/\1$o/;};}" "$pdbin" > "$xyzin"
      refmac
   fi
   cd ..
}
```

```
#function to use instead of process when testing script
test_process () {
   echo "Running $o"
   sleep 0.1
   echo "Ran $o"
}

#writes column titles to csv result file, then greps B-factors for haem Fe and NO
N in chains A-F for each occupancy structure
#haem code is HEM in dtpb, and HEC in mccp
resultcsv () {
   line="$o"
   for chain in {A..F}
   do
      line_fe="$(grep HETATM occ$o/REFMAC*.pdb | grep "FE   HE[CM] $chain")"
      line_no="$(grep HETATM occ$o/REFMAC*.pdb | grep "N   NO  $chain")"
#     bf_fe=${line_fe:60:6}
#     bf_no=${line_no:60:6}
#     echo $line_fe
#     echo $line_no
      bf_fe=$(echo ${line_fe:60:6} | tr -d '[:space:]')
                bf_no=$(echo ${line_no:60:6} | tr -d '[:space:]')
#     echo $bf_fe
#     echo $bf_no
      line="$line,$bf_fe,$bf_no"
   done
   echo "$line" | tee -a "$resultfile"
}

#################################
#ACTUALLY STARTS PROCESSING HERE#
#################################

#specify input files to refine NO occupancies:
#$1 must be PDB format, $2 must be MTZ format
#readlink shows absolute path to files, -s supresses errors
pdbin=$(readlink -mns $1)
mtzin=$(readlink -mns $2)
rstin=$(readlink -mns $3)

#checks input file types, extension is case insensitive:
pdbinlower=${pdbin,,}
mtzinlower=${mtzin,,}
if [ "${pdbinlower: -4}" == ".pdb" ] && [ "${mtzinlower: -4}" == ".mtz" ]
then
   echo "Input PDB file:      $pdbin"
   echo "Input MTZ file:      $mtzin"
   echo "Input restraint file:   $rstin"
else
   echo -e "Error: Please specify input files in PDB and MTZ formats, and a
restrain file. For example \"./occ_refine.sh model.pdb map.mtz
restraint_dic.txt\"."
   exit 1
fi

module load ccp4

#performs processing in parallel, batches of N
N=10
```

```
for o in
{0.00,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.10,0.11,0.12,0.13,0.14,0.15,
0.16,0.17,0.18,0.19,0.20,0.21,0.22,0.23,0.24,0.25,0.26,0.27,0.28,0.29,0.30,0.31,0
.32,0.33,0.34,0.35,0.36,0.37,0.38,0.39,0.40,0.41,0.42,0.43,0.44,0.45,0.46,0.47,0.
48,0.49,0.50,0.51,0.52,0.53,0.54,0.55,0.56,0.57,0.58,0.59,0.60,0.61,0.62,0.63,0.6
4,0.65,0.66,0.67,0.68,0.69,0.70,0.71,0.72,0.73,0.74,0.75,0.76,0.77,0.78,0.79,0.80
,0.81,0.82,0.83,0.84,0.85,0.86,0.87,0.88,0.89,0.90,0.91,0.92,0.93,0.94,0.95,0.96,
0.97,0.98,0.99,1.00}
do
   ((i=i%N))
   ((i++==0)) && wait
#  process does work, test_process to test loop
#  echo "occ = $o"
   process &
#  test_process &
done

wait

#If results.csv exists, then saves as results_new.csv, results_new_new.csv, etc
sheet="results"
resultfile="${sheet}.csv"
until [ ! -f "$resultfile" ]
do
   sheet="${sheet}_new"
   resultfile="${sheet}.csv"
done

#calls resultcsv function to copy data from refmac output to csv file
echo "Tabulating B-factor data to ${resultfile}:"
echo
"================================================================================
"
        echo
"occupancy,A_Fe_Bfactor,A_N_Bfactor,B_Fe_Bfactor,B_N_Bfactor,C_Fe_Bfactor,C_N_Bfa
ctor,D_Fe_Bfactor,D_N_Bfactor,E_Fe_Bfactor,E_N_Bfactor,F_Fe_Bfactor,F_N_Bfactor"
| tee $resultfile
for o in
{0.00,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.10,0.11,0.12,0.13,0.14,0.15,
0.16,0.17,0.18,0.19,0.20,0.21,0.22,0.23,0.24,0.25,0.26,0.27,0.28,0.29,0.30,0.31,0
.32,0.33,0.34,0.35,0.36,0.37,0.38,0.39,0.40,0.41,0.42,0.43,0.44,0.45,0.46,0.47,0.
48,0.49,0.50,0.51,0.52,0.53,0.54,0.55,0.56,0.57,0.58,0.59,0.60,0.61,0.62,0.63,0.6
4,0.65,0.66,0.67,0.68,0.69,0.70,0.71,0.72,0.73,0.74,0.75,0.76,0.77,0.78,0.79,0.80
,0.81,0.82,0.83,0.84,0.85,0.86,0.87,0.88,0.89,0.90,0.91,0.92,0.93,0.94,0.95,0.96,
0.97,0.98,0.99,1.00}
do
   resultcsv
done
echo
"================================================================================
"

#prints final completion message with time taken
echo "Refined $(ls occ*.*/*.log | wc -l) structures with NO occupancy 0.00-1.00
and tabulated B-factors."
echo "Saved REFMAC output to \"occX.XX/\" directories, and B-factor data to
${resultfile}."
end=$(date +%s)
echo "Took $(($end-$start)) seconds to run."
```

2. Occupancy refinement output log file

```
Input PDB file:      /dls/mx-scratch/pete/2023-
02_DtpB_power_titration_2/occ_refine/20percent/9_2023-
02_dtpb_20percent_2_xyzout_prosmart_refmac.pdb
Input MTZ file:      /dls/mx-scratch/pete/2023-
02_DtpB_power_titration_2/occ_refine/20percent/9_2023-
02_dtpb_20percent_2_fphiout_prosmart_refmac.mtz
Input restraint file:   /dls/mx-scratch/pete/2023-
02_DtpB_power_titration_2/occ_refine/20percent/9_2023-
02_dtpb_20percent_2_dict_prosmart_refmac.cif
Refining structure with NO occupancy=0.08 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.01 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.06 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.00 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.09 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.02 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.03 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.05 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.07 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.04 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.05
Refined structure with NO occupancy = 0.01
Refined structure with NO occupancy = 0.04
Refined structure with NO occupancy = 0.07
Refined structure with NO occupancy = 0.03
Refined structure with NO occupancy = 0.02
Refined structure with NO occupancy = 0.09
Refined structure with NO occupancy = 0.00
Refined structure with NO occupancy = 0.06
Refined structure with NO occupancy = 0.08
Refining structure with NO occupancy=0.10 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.12 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.13 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.14 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.17 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.16 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.11 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.15 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.18 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.19 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.13
Refined structure with NO occupancy = 0.19
Refined structure with NO occupancy = 0.17
Refined structure with NO occupancy = 0.18
Refined structure with NO occupancy = 0.10
Refined structure with NO occupancy = 0.14
Refined structure with NO occupancy = 0.12
Refined structure with NO occupancy = 0.15
Refined structure with NO occupancy = 0.11
Refined structure with NO occupancy = 0.16
Refining structure with NO occupancy=0.25 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.22 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.23 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.21 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.27 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.29 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.24 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.28 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.20 with 10 cycle(s) of refinement.
```

```
Refining structure with NO occupancy=0.26 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.29
Refined structure with NO occupancy = 0.27
Refined structure with NO occupancy = 0.20
Refined structure with NO occupancy = 0.23
Refined structure with NO occupancy = 0.25
Refined structure with NO occupancy = 0.24
Refined structure with NO occupancy = 0.28
Refined structure with NO occupancy = 0.26
Refined structure with NO occupancy = 0.22
Refined structure with NO occupancy = 0.21
Refining structure with NO occupancy=0.36 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.34 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.39 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.33 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.37 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.38 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.30 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.32 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.31 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.35 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.38
Refined structure with NO occupancy = 0.36
Refined structure with NO occupancy = 0.37
Refined structure with NO occupancy = 0.39
Refined structure with NO occupancy = 0.33
Refined structure with NO occupancy = 0.31
Refined structure with NO occupancy = 0.34
Refined structure with NO occupancy = 0.30
Refined structure with NO occupancy = 0.35
Refined structure with NO occupancy = 0.32
Refining structure with NO occupancy=0.49 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.42 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.45 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.43 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.48 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.47 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.41 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.40 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.46 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.44 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.46
Refined structure with NO occupancy = 0.44
Refined structure with NO occupancy = 0.43
Refined structure with NO occupancy = 0.45
Refined structure with NO occupancy = 0.41
Refined structure with NO occupancy = 0.47
Refined structure with NO occupancy = 0.48
Refined structure with NO occupancy = 0.40
Refined structure with NO occupancy = 0.42
Refined structure with NO occupancy = 0.49
Refining structure with NO occupancy=0.56 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.51 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.54 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.53 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.50 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.59 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.57 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.55 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.58 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.52 with 10 cycle(s) of refinement.
```

```
Refined structure with NO occupancy = 0.50
Refined structure with NO occupancy = 0.53
Refined structure with NO occupancy = 0.54
Refined structure with NO occupancy = 0.51
Refined structure with NO occupancy = 0.59
Refined structure with NO occupancy = 0.57
Refined structure with NO occupancy = 0.56
Refined structure with NO occupancy = 0.58
Refined structure with NO occupancy = 0.52
Refined structure with NO occupancy = 0.55
Refining structure with NO occupancy=0.65 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.68 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.62 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.69 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.61 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.67 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.66 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.63 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.64 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.60 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.61
Refined structure with NO occupancy = 0.66
Refined structure with NO occupancy = 0.68
Refined structure with NO occupancy = 0.64
Refined structure with NO occupancy = 0.60
Refined structure with NO occupancy = 0.67
Refined structure with NO occupancy = 0.65
Refined structure with NO occupancy = 0.63
Refined structure with NO occupancy = 0.69
Refined structure with NO occupancy = 0.62
Refining structure with NO occupancy=0.76 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.74 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.71 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.78 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.72 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.77 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.70 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.73 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.75 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.79 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.74
Refined structure with NO occupancy = 0.76
Refined structure with NO occupancy = 0.77
Refined structure with NO occupancy = 0.78
Refined structure with NO occupancy = 0.70
Refined structure with NO occupancy = 0.71
Refined structure with NO occupancy = 0.73
Refined structure with NO occupancy = 0.72
Refined structure with NO occupancy = 0.79
Refined structure with NO occupancy = 0.75
Refining structure with NO occupancy=0.85 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.84 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.88 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.86 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.89 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.81 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.80 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.83 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.87 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.82 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.86
```

```
Refined structure with NO occupancy = 0.83
Refined structure with NO occupancy = 0.87
Refined structure with NO occupancy = 0.81
Refined structure with NO occupancy = 0.85
Refined structure with NO occupancy = 0.82
Refined structure with NO occupancy = 0.84
Refined structure with NO occupancy = 0.89
Refined structure with NO occupancy = 0.88
Refined structure with NO occupancy = 0.80
Refining structure with NO occupancy=0.90 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.93 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.92 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.99 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.96 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.97 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.94 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.91 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.95 with 10 cycle(s) of refinement.
Refining structure with NO occupancy=0.98 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 0.93
Refined structure with NO occupancy = 0.92
Refined structure with NO occupancy = 0.96
Refined structure with NO occupancy = 0.94
Refined structure with NO occupancy = 0.97
Refined structure with NO occupancy = 0.95
Refined structure with NO occupancy = 0.99
Refined structure with NO occupancy = 0.90
Refined structure with NO occupancy = 0.98
Refined structure with NO occupancy = 0.91
Refining structure with NO occupancy=1.00 with 10 cycle(s) of refinement.
Refined structure with NO occupancy = 1.00
Tabulating B-factor data to results.csv:
================================================================================
occupancy,A_Fe_Bfactor,A_N_Bfactor,B_Fe_Bfactor,B_N_Bfactor,C_Fe_Bfactor,C_N_Bfac
tor,D_Fe_Bfactor,D_N_Bfactor,E_Fe_Bfactor,E_N_Bfactor,F_Fe_Bfactor,F_N_Bfactor
0.00,28.19,45.26,33.88,47.60,30.27,44.40,28.86,40.00,29.15,48.41,32.31,54.93
0.01,28.12,27.55,33.81,33.48,30.17,29.83,28.76,29.07,29.09,28.81,32.22,32.20
0.02,28.14,26.92,33.75,32.75,30.17,29.03,28.78,28.14,29.09,28.22,32.24,31.82
0.03,28.04,26.16,33.69,31.76,30.08,28.58,28.67,27.17,29.03,27.74,32.25,31.44
0.04,27.97,25.79,33.57,30.72,30.00,27.92,28.63,26.69,28.95,27.11,32.04,30.69
0.05,27.98,25.24,33.59,29.89,29.99,27.21,28.61,26.18,28.94,26.62,32.08,30.33
0.06,27.88,24.60,33.47,28.97,29.91,26.57,28.50,25.33,28.83,25.98,31.99,29.99
0.07,27.89,24.02,33.46,28.25,29.96,26.03,28.50,24.91,28.90,25.66,32.04,29.53
0.08,27.85,23.55,33.40,27.41,29.91,25.49,28.47,24.51,28.84,25.15,32.02,29.20
0.09,27.77,22.90,33.28,26.47,29.78,24.92,28.38,23.74,28.74,24.66,31.94,28.80
0.10,27.69,22.63,33.19,25.98,29.75,24.37,28.29,23.30,28.65,24.25,31.89,28.50
0.11,27.70,22.32,33.17,25.74,29.74,23.98,28.29,23.13,28.65,23.96,31.89,28.27
0.12,27.69,21.98,33.18,24.87,29.80,23.65,28.33,22.79,28.70,23.78,31.98,28.06
0.13,27.63,21.66,33.13,24.38,29.72,23.21,28.20,22.46,28.67,23.46,31.89,27.74
0.14,27.59,21.43,33.09,24.39,29.68,22.86,28.18,22.21,28.64,23.26,31.88,27.66
0.15,27.58,21.21,33.00,23.64,29.64,22.49,28.11,21.92,28.64,23.10,31.88,27.54
0.16,27.54,21.01,32.97,23.34,29.62,22.24,28.07,21.77,28.54,22.94,31.88,27.48
0.17,27.49,20.86,32.88,23.01,29.57,21.97,28.05,21.44,28.58,22.78,31.83,27.36
0.18,27.50,20.91,32.87,22.89,29.63,21.81,28.00,21.44,28.55,22.82,31.90,27.38
0.19,27.45,20.68,32.81,22.58,29.57,21.63,27.99,21.19,28.56,22.65,31.85,27.31
0.20,27.40,20.60,32.78,22.38,29.54,21.36,27.96,21.02,28.54,22.59,31.83,27.41
0.21,27.34,20.51,32.73,22.30,29.49,21.20,27.87,20.89,28.47,22.48,31.79,27.35
0.22,27.33,20.45,32.68,22.07,29.49,21.08,27.86,20.85,28.44,22.47,31.80,27.34
0.23,27.31,20.43,32.62,21.98,29.50,21.06,27.76,20.85,28.39,22.56,31.79,27.47
0.24,27.23,20.29,32.56,21.98,29.43,20.98,27.75,20.68,28.38,22.55,31.73,27.56
0.25,27.25,20.31,32.55,21.91,29.45,20.92,27.70,20.67,28.42,22.51,31.77,27.73
```

0.26,27.23,20.34,32.52,21.84,29.43,20.86,27.68,20.65,28.39,22.58,31.73,27.80
0.27,27.22,20.43,32.49,21.77,29.42,20.94,27.66,20.66,28.42,22.67,31.77,28.08
0.28,27.15,20.49,32.44,21.80,29.40,20.84,27.54,20.80,28.35,22.87,31.74,28.34
0.29,27.12,20.42,32.37,21.73,29.42,20.93,27.57,20.83,28.37,22.81,31.70,28.40
0.30,27.14,20.51,32.36,21.81,29.42,20.93,27.55,20.81,28.39,22.92,31.74,28.75
0.31,27.12,20.68,32.34,21.83,29.37,21.01,27.51,20.87,28.37,23.04,31.70,28.95
0.32,27.12,20.76,32.30,21.83,29.37,21.06,27.53,20.99,28.38,23.19,31.72,29.25
0.33,27.11,20.92,32.31,21.89,29.39,21.17,27.48,21.12,28.37,23.31,31.73,29.59
0.34,27.11,21.04,32.28,22.11,29.37,21.21,27.47,21.32,28.34,23.69,31.75,29.80
0.35,27.05,21.15,32.24,22.12,29.37,21.36,27.39,21.49,28.32,23.89,31.73,29.97
0.36,27.07,21.27,32.22,22.18,29.34,21.40,27.43,21.47,28.36,23.94,31.73,30.39
0.37,27.04,21.54,32.18,22.31,29.35,21.57,27.37,21.73,28.34,24.40,31.73,30.66
0.38,27.03,21.63,32.19,22.42,29.34,21.69,27.35,21.91,28.31,24.63,31.76,30.91
0.39,26.98,21.82,32.16,22.55,29.30,21.72,27.31,21.83,28.30,24.62,31.71,31.22
0.40,26.98,22.02,32.14,22.57,29.33,21.87,27.30,22.00,28.29,24.96,31.76,31.54
0.41,27.02,22.31,32.17,22.84,29.38,22.07,27.36,22.22,28.36,25.28,31.80,32.11
0.42,27.00,22.41,32.16,23.03,29.36,22.27,27.33,22.47,28.33,25.56,31.78,32.30
0.43,26.99,22.71,32.11,23.20,29.34,22.43,27.31,22.60,28.34,25.81,31.76,32.82
0.44,27.02,22.96,32.16,23.44,29.40,22.68,27.34,22.91,28.35,26.16,31.85,33.17
0.45,26.99,23.23,32.13,23.63,29.37,22.85,27.31,23.05,28.37,26.46,31.83,33.59
0.46,26.99,23.47,32.13,23.87,29.36,23.06,27.30,23.26,28.35,26.69,31.83,34.05
0.47,27.00,23.71,32.09,24.14,29.40,23.30,27.31,23.67,28.37,27.08,31.85,34.36
0.48,26.98,24.04,32.08,24.37,29.41,23.46,27.29,23.90,28.38,27.45,31.82,34.84
0.49,27.00,24.31,32.07,24.68,29.37,23.67,27.25,24.15,28.37,27.81,31.82,35.22
0.50,27.00,24.56,32.07,24.97,29.41,23.95,27.26,24.39,28.39,28.03,31.86,35.67
0.51,27.02,24.83,32.09,25.19,29.40,24.24,27.31,24.70,28.39,28.37,31.87,36.12
0.52,27.01,25.09,32.08,25.48,29.38,24.44,27.28,24.92,28.41,28.72,31.87,36.57
0.53,26.98,25.41,32.04,25.77,29.39,24.68,27.27,25.17,28.39,29.07,31.86,36.84
0.54,26.97,25.76,32.04,26.21,29.36,24.94,27.23,25.38,28.38,29.41,31.84,37.20
0.55,26.99,26.06,32.05,26.45,29.40,25.20,27.22,25.76,28.41,29.65,31.85,37.68
0.56,27.04,26.44,32.10,26.86,29.44,25.60,27.28,26.13,28.46,30.24,31.92,38.23
0.57,26.98,26.69,32.02,26.98,29.37,25.88,27.22,26.32,28.43,30.48,31.90,38.51
0.58,27.03,26.92,32.08,27.48,29.45,26.08,27.27,26.62,28.48,30.81,31.92,39.03
0.59,27.00,27.28,32.03,27.69,29.40,26.37,27.23,26.91,28.44,31.10,31.90,39.37
0.60,27.03,27.67,32.07,28.12,29.41,26.67,27.25,27.19,28.47,31.52,31.91,40.03
0.61,27.03,27.98,32.05,28.38,29.42,26.94,27.26,27.51,28.49,31.91,31.95,40.18
0.62,27.07,28.34,32.09,28.77,29.43,27.27,27.26,27.84,28.50,32.26,31.98,40.65
0.63,27.03,28.69,32.06,29.16,29.40,27.55,27.24,28.09,28.48,32.67,31.93,41.11
0.64,27.02,28.94,32.04,29.43,29.43,27.93,27.27,28.38,28.53,32.98,31.95,41.48
0.65,27.01,29.28,32.03,29.69,29.35,28.23,27.19,28.69,28.45,33.36,31.90,41.90
0.66,27.04,29.66,32.03,30.19,29.43,28.43,27.22,29.06,28.46,33.70,31.93,42.27
0.67,27.05,30.06,32.07,30.60,29.42,28.82,27.25,29.41,28.49,34.03,31.96,42.82
0.68,27.06,30.27,32.10,30.84,29.42,29.11,27.27,29.71,28.50,34.45,31.99,43.27
0.69,27.10,30.78,32.09,31.35,29.43,29.43,27.27,29.99,28.51,34.90,32.02,43.77
0.70,27.07,30.98,32.05,31.59,29.40,29.73,27.24,30.32,28.47,35.21,31.98,44.03
0.71,27.08,31.37,32.09,32.00,29.44,30.07,27.26,30.69,28.50,35.64,32.01,44.61
0.72,27.08,31.63,32.04,32.35,29.43,30.29,27.24,30.81,28.53,35.96,32.00,45.00
0.73,27.13,31.98,32.12,32.62,29.44,30.60,27.28,31.29,28.53,36.36,32.03,45.38
0.74,27.12,32.36,32.12,33.12,29.44,30.94,27.22,31.71,28.53,36.96,32.07,45.71
0.75,27.14,32.66,32.13,33.55,29.42,31.27,27.28,31.75,28.57,37.04,32.03,46.24
0.76,27.12,33.05,32.12,33.85,29.46,31.45,27.27,32.14,28.54,37.41,32.04,46.50
0.77,27.16,33.35,32.15,34.32,29.47,31.91,27.29,32.47,28.59,37.92,32.05,47.01
0.78,27.14,33.78,32.17,34.61,29.46,32.13,27.26,33.04,28.52,38.42,32.09,47.34
0.79,27.13,33.96,32.11,34.96,29.45,32.43,27.27,33.06,28.56,38.60,32.04,47.77
0.80,27.17,34.36,32.17,35.31,29.49,32.82,27.32,33.41,28.59,39.07,32.07,48.22
0.81,27.17,34.63,32.15,35.73,29.49,33.13,27.32,33.68,28.61,39.43,32.08,48.64
0.82,27.17,35.13,32.19,36.04,29.52,33.37,27.28,34.24,28.57,39.86,32.13,49.05
0.83,27.20,35.32,32.21,36.41,29.49,33.73,27.34,34.40,28.62,40.12,32.12,49.45
0.84,27.19,35.62,32.20,36.76,29.46,33.99,27.33,34.60,28.61,40.41,32.09,49.82
0.85,27.25,36.13,32.24,37.17,29.56,34.35,27.37,35.18,28.66,40.96,32.16,50.22
0.86,27.20,36.27,32.18,37.46,29.49,34.66,27.36,35.29,28.61,41.13,32.14,50.58

```
0.87,27.23,36.63,32.21,37.86,29.50,34.98,27.35,35.62,28.68,41.64,32.13,50.98
0.88,27.25,36.99,32.25,38.11,29.53,35.17,27.34,36.14,28.62,42.19,32.19,51.42
0.89,27.27,37.27,32.20,38.51,29.53,35.58,27.38,36.23,28.65,42.38,32.17,51.88
0.90,27.24,37.62,32.26,38.87,29.55,35.83,27.33,36.73,28.59,42.96,32.20,52.03
0.91,27.23,37.86,32.19,39.16,29.46,36.15,27.33,36.84,28.62,42.98,32.11,52.50
0.92,27.30,38.35,32.27,39.53,29.53,36.46,27.37,37.40,28.65,43.53,32.22,52.81
0.93,27.27,38.55,32.24,39.74,29.52,36.76,27.41,37.51,28.63,43.72,32.18,53.29
0.94,27.28,38.83,32.26,40.12,29.55,37.11,27.40,37.75,28.67,44.05,32.20,53.56
0.95,27.22,39.04,32.27,40.50,29.48,37.39,27.40,38.08,28.62,44.43,32.14,53.91
0.96,27.33,39.58,32.33,40.89,29.58,37.64,27.41,38.60,28.72,45.11,32.27,54.40
0.97,27.27,39.70,32.27,41.20,29.54,38.00,27.42,38.63,28.68,45.27,32.23,54.71
0.98,27.28,39.97,32.30,41.65,29.55,38.28,27.41,38.95,28.69,45.58,32.19,55.11
0.99,27.37,40.47,32.36,41.79,29.59,38.48,27.42,39.44,28.68,45.98,32.32,55.42
1.00,27.33,40.59,32.32,42.24,29.56,38.94,27.44,39.57,28.74,46.32,32.23,55.98
================================================================================
Refined 101 structures with NO occupancy 0.00-1.00 and tabulated B-factors.
Saved REFMAC output to "occX.XX/" directories, and B-factor data to results.csv.
Took 1953 seconds to run.
```

3. Difference density retrieval Bash script

```
module load ccp4
echo "Analysing difference density at ligand nitrogen position for refined
structures."
For struc in {40,20,10,05,0p5}
do
   echo "structure: ${struc}percent" >> density_output.csv
   for occ in
{0.00,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.10,0.11,0.12,0.13,0.14,0.15,
0.16,0.17,0.18,0.19,0.20,0.21,0.22,0.23,0.24,0.25,0.26,0.27,0.28,0.29,0.30,0.31,0
.32,0.33,0.34,0.35,0.36,0.37,0.38,0.39,0.40,0.41,0.42,0.43,0.44,0.45,0.46,0.47,0.
48,0.49,0.50,0.51,0.52,0.53,0.54,0.55,0.56,0.57,0.58,0.59,0.60,0.61,0.62,0.63,0.6
4,0.65,0.66,0.67,0.68,0.69,0.70,0.71,0.72,0.73,0.74,0.75,0.76,0.77,0.78,0.79,0.80
,0.81,0.82,0.83,0.84,0.85,0.86,0.87,0.88,0.89,0.90,0.91,0.92,0.93,0.94,0.95,0.96,
0.97,0.98,0.99,1.00}
   do
      coot –pdb ../${struc}percent/occ$occ/occ*.pdb –auto
../${struc}percent/occ$occ/REFMAC*.mtz –no-graphics -s density_${struc}.py
      echo $occ >> density_output.csv
   done
done
echo "#### Done ####"
```

4. Difference density retrieval example Python script for *Coot*, density_40.py

```
density_a = density_at_point(1,-42.98,-15.11,7.78)
density_b = density_at_point(1,2.06,8.56,38.08)
density_c = density_at_point(1,-39.87,-2.73,48.10)
density_d = density_at_point(1,-14.10,-30.91,24.15)
density_e = density_at_point(1,-34.17,25.05,24.47)
density_f = density_at_point(1,-3.20,2.88,1.97)

f = open("density_output.csv", "a")
f.write(str(density_a) + "," + str(density_b) + "," + str(density_c) + "," +
str(density_d) + "," + str(density_e) + "," + str(density_f) + ",")
f.close()

exit()
```

5. Difference density retrieval script output (abridged)

```
structure: 40percent
0.56919926405,0.607405662537,0.478522062302,0.493275702,0.593108892441,0.761418402195,0.00
0.559231162071,0.599429428577,0.472777962685,0.575083851814,0.582068502903,0.749957442284,0.01
0.552556812763,0.591965734959,0.464419841766,0.568236470222,0.574745416641,0.740307629108,0.02
...
-0.0656774044037,-0.0935266986489,-0.136973619461,-0.140445187688,-0.07212677598,-0.041556764394,0.99
-0.0757924392819,-0.103534892201,-0.140208020806,-0.14966148138,-0.0766198337078,-0.0441017821431,1.00
structure: 20percent
0.583626151085,0.53008055687,0.645498037338,0.797028005123,0.550472915173,0.356058716774,0.00
...
structure: 10percent
0.39210948348,0.535253882408,0.508890867233,0.420545101166,0.369890153408,0.48885846138,0.00
...
structure: 05percent
0.316481560469,0.360270768404,0.261154681444,0.316856682301,0.289140999317,0.194566071033,0.00
...
structure: 0p5percent
0.162947103381,0.284459292889,0.172173559666,0.214752569795,0.170785069466,0.379865109921,0.00
...
structure: static
-0.155969426036,-0.112239286304,0.0320361964405,0.130466103554,0.0942721292377,-0.0129051264375,0.00
...
-0.518309831619,-0.491184681654,-0.324650585651,-0.318280816078,-0.338334977627,-0.350260823965,1.00
```