Imperial College London

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

---

Time-Resolved Protein Dynamics Using X-ray Crystallography and Optical Spectroscopy

---

Author:

James Michael Baxter

Supervisor:

Prof. Jasper van Thor

Department of Life Sciences

Faculty of Natural Sciences

2021

Statement of Originality

I confirm that the work in this thesis is my own. Where information or data originates from other sources I confirm that this has been indicated with citation or explanation.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law

Abstract

Reversibly switchable fluorescent proteins (RSFPs) switch between a highly and weakly fluorescent state upon photoexcitation. This switching is known to be driven by reversible photoisomerisation and a change in protonation state of the chromophore. However, the exact sequence of events after optical excitation is unclear. In this thesis, time-resolved crystallography and vibrational spectroscopy are used to study the ultrafast dynamics of a new RSFP.

Time-resolved serial femtosecond crystallography (TR-SFX) is used to probe the trans/cis reaction and reveals ultrafast sub-Angstrom motions of the chromophore and surrounding hydrogen bonding network. The addition of a stokes pulse (shown to dump the excited state population) increases these displacements and allows assignment of the light-induced signals to a vibrationally excited electronic ground state structure instead of an excited state structure. Temperature analysis of the steady-state switching rates resolves an intermediate ground-state structure using conventional crystallography which supports the TR-SFX assignments. Ultrafast pump-probe vibrational spectroscopy confirms that excited state photoisomerisation occurs (∼70 ps), proceeded by rearrangement of a nearby arginine residue (∼7 ps) with deprotonation (trans-neutral/cis-anion) occurring between 2 ns – 1 ms. An additional dump pulse is then shown to repress product formation and reforms the ground state, further supporting the structural observations of the pump-dump-probe crystallography measurements. These results improve our understanding of photoswitching in RSFPs and have important implications for structural assignments in future TR-SFX experiments.

Time-resolved serial synchrotron crystallography (TR-SSX) is an accessible and emerging technique for studying protein dynamics on millisecond timescales. Here, the trans/cis isomerisation of an RSFP is probed at 10 ms with a range of pump intensities to demonstrate precise control of the cis-state formation yields. The population transfer is resolved crystallographically using quantitative analysis of the occupancy, showing excellent agreement with optical flash-photolysis measurements. The issue of X-ray induced heating, which leads to partial thermal recovery, is also addressed. A temperature increase of 82–112 K is extracted for the fixed-target sample delivery used, which should be accounted for in all future TR-SSX experiments.

Finally, the effect of pH on the protonation step is investigated by studying a RSFP mutant with a destabilised ground state conformation. Structural and spectroscopic evidence is presented for a new class of reaction at low pH, involving a cis-neutral/trans-cation isomerisation, in contrasts to the typical cis-anion/trans-neutral reaction at neutral pH. These steady-state observations expand our understanding of acid/base equilibria in the photoreactions of RSFPs.

Acknowledgements

A huge number of people have made my PhD possible. I would like to thank Jasper for supervision though my PhD and being a bottomless pit of knowledge and ideas. Chris, thank you for a wide array of help over the years, especially through all of the experimental situations we found ourselves in... I would potentially have burnt down the Sir Ernst Chain building if I had to purify my own protein so I owe Karim and Violeta a huge amount for the litres of fluorescent protein they have produced over the years. Additionally, I don't think I could have completed my PhD if it was not for the environment around me. So thank you to Olga, Alisia, Stefania, Giovanni, Rafal, Beth, Bill, Andrea and others for comradery and discussions of experiments and for providing a fun place to work. Also thank you to James for the many discussions on crystallographic analysis and for introducing me to tools from the 1980s buried within CCP4i and FORTRAN code. Also Marc, for teaching me how to fish crystals and organisation of countless beamtimes.

In addition, a number of collaborators were fantastic to work with. At Diamond, Pierre and Agata were very helpful in organising serial crystallography experiments and Ann was kind enough to share laser and interview advice. In Hamburg, at the EMBL, Sam, Michael, David and Arwen helped in so many ways, from sample preparation to experimental suggestions and more.

I would like to thank friends who without which I'm not sure I would have enjoyed three years in London (as much) without. Louise, thank you for snacks, positivity and kindness, Tom and Richard for living with me, Adam, John, Mary and Sam for keeping me sane during lockdown.

Finally, I have been very lucky to grow up with such a supportive family. To my Mum, Dad, Aunt, Uncle, Gran and Brother, thank you for the wooden swords, Quadro and chemistry set growing up. Also thank you for the encouragement to pursue whatever I became interested in.

Contents

1 Introduction **5**

1.1 Motivation and Aims . . . 5

1.2 Modern X-ray Sources for Crystallography . . . 6

1.3 Studying Dynamics with X-ray Crystallography . . . 11

1.4 Analysis and Interpretation of Time-Resolved Crystallography Data . . . 18

1.5 Ultrafast Spectroscopy . . . 24

1.6 Thesis Outline . . . 29

2 Development of Reversibly Switchable Fluorescent Proteins **30**

2.1 Introduction . . . 30

2.2 Methods . . . 35

2.3 Results . . . 37

2.4 Discussion . . . 44

2.5 Conclusions . . . 45

3 Cryo-Kinetics and Cryo-Crystallography of rsKiiro **47**

3.1 Introduction . . . 47

3.2 Methods . . . 50

3.3 Cryo-Spectroscopy Reveals Convex Arrhenius Behaviour . . . 52

3.4 Cryo-Trapping Steady States with Macromolecular Crystallography . . . 54

3.5 Discussion . . . 61

3.6 Conclusions . . . 62

4 Time-Resolved Serial Femtosecond Crystallography of rsKiiro **64**

4.1 Introduction . . . 65

4.2 Experiment Overview and Implementation . . . 67

4.3 Data Processing of Merged TR-SFX Data . . . 81

4.4 Experiment Results and Discussion . . . 87

4.5 Conclusions . . . 94

5 Serial Synchrotron Crystallography Feels the Heat **96**
5.1 Introduction . . . 96
5.2 Methods . . . 98
5.3 Results and Discussion . . . 103
5.4 Conclusions and Outlook . . . 114

6 Ultrafast IRTA with Optical Control Reveals rsKiiro Isomerisation Mechanism **115**
6.1 Introduction . . . 115
6.2 Methods . . . 117
6.3 Results . . . 121
6.4 Discussion . . . 126
6.5 Conclusions . . . 133

7 Observation of a Neutral to Cation Switching of a GFP-like Protein **134**
7.1 Introduction . . . 135
7.2 Methods . . . 137
7.3 Results . . . 140
7.4 Discussion . . . 153
7.5 Conclusions . . . 157

8 Conclusions **158**

A Appendix **202**
A.1 Development of RSFPs . . . 202
A.2 Cryo-Kinetics and Crystallography . . . 204
A.3 SFX of rsKiiro . . . 205
A.4 Quantitative Serial Synchrotron Crystallography of rsKiiro . . . 219
A.5 Ultrafast IRTA with Optical Control Reveals rsKiiro Isomerisation Mechanism . . . 222
A.6 Observation of a Neutral to Cation Switching of a GFP-like Protein . . . 224

1 | Introduction

This chapter provides a general introduction to modern X-ray sources and their use in optically pumped, time-resolved crystallographic studies. This is followed by a review of current data analysis practises for the determination of light-induced structural differences in protein crystallography. The details of ultrafast pump-probe transient absorption spectroscopy are then introduced along with the data analysis framework of global fitting.

1.1 Motivation and Aims

Observing the structure and dynamics of biological molecules is key to understanding the complexity and diversity that makes up living organisms. Proteins are large molecules assembled from peptide-linked amino acids. X-ray crystallography has long been the most successful method for protein-structure determination. With the increasing brightness of modern X-ray sources it is becoming possible to make time-resolved crystallography measurements with high spatial and temporal resolution. Here, work will focus on method and sample development for studying the dynamics of a reversibly switchable fluorescent protein (RSFP) using X-ray crystallography and spectroscopy.

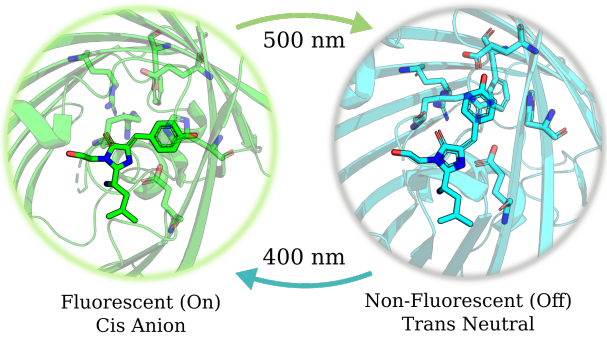


Figure 1.1: The on (5OQE green [1]) and off (5OQA cyan [1]) structures of the reversibly switchable fluorescent protein Skylan-NS [2]. Illumination with 500/400 nm switches from on-to-off/off-to-on.

RSFPs are molecules that can be switched between a highly-fluorescent 'on' and weakly-fluorescent 'off' state with photoexcitation (figure 1.1). In the class of RSFPs studied in this thesis (introduced in chapter 2) the on-state corresponds to an anionic-cis chromophore conformer, and the off-state to a

neutral-trans [3]. Switching is generally driven with excitation around 500 nm (on to off) or 400 nm (off to on) although there are many variants with different excitation schemes [4]. Currently, the exact sequence of events that leads to photoisomerisation and fluorescence switching in these proteins is not clear due to conflicting spectroscopic assignments [5–8] (chapter 6). In this thesis, these assignments will be addressed using ultrafast multi-pulse spectroscopy experiments and multi-pulse time-resolved crystallography experiments. The broad aim of this thesis is to improve our understanding of switching mechanisms in RSFPs and improve the methodology used to study these reactions and other proteins.

Improving our understanding of the switching mechanisms in RSFPs is deemed interesting and worthwhile for three main reasons. Firstly, photoisomerisation and subsequent macromolecular motions play key roles in photoreceptor biological function across plants [9–11], vision [12–14] and even in bioelectronics [15–17]. The photoswitching behaviour of RSFPs is therefore an opportunity to gain insight into these fundamental motions, which has implications for understanding a wide array of protein motions throughout biology [18]. Secondly, RSFPs are used across several super-resolution microscopy techniques [19–24]. Improved understanding of the switching mechanism in these proteins may lead to improved engineering of these proteins, which could improve the resolution of super-resolution microscopy measurements [25]. Thirdly, the development of X-ray free electron lasers and improvements in the brightness of synchrotron X-ray sources is driving a new era for studying the dynamics of proteins in their crystalline form [26–28]. It is important to properly understand the fundamental limitations of these measurement techniques, in particular, the effects of intense X-ray and optical pulses on protein systems during time-resolved X-ray crystallography experiments. Reactions in RSFPs are kinetically simple compared to more complex, multistep, enzyme reactions which means they are ideal model systems for these types of fundamental experiments.

1.2 Modern X-ray Sources for Crystallography

X-ray crystallography relies on the diffraction of hard X-rays (keV photon energies) by a periodic array of molecules to obtain structural information. The brightness, spot-size and stability of the X-ray source can dictate the range of possible experiments and the quality of the measured diffraction data. The first dedicated synchrotron X-ray sources were built in the 1980s [29] as replacements for previously used cathode-ray sources [30]. Synchrotron sources drove an explosion in the development of macromolecular crystallography (MX) [31], which is now responsible for around 130,000 protein structures deposited in the protein data bank (PDB) [32]. Access to these brighter X-ray beams also led to the evolution of time-resolved (or kinetic, or dynamic) crystallography [33] which was originally limited in time resolution to around 30 minutes [34]. As synchrotron technology has matured [28] and with the development of extremely bright X-ray free electron lasers (XFELs) [35] dynamic crystallography experiments with up to femtosecond time-resolution have become possible [36].

The two types of X-ray photon sources that are relevant for studying biomolecular dynamics are synchrotron storage rings and X-ray free electron lasers (XFELs). Both sources generate bright and spatially coherent X-rays by accelerating electron beams at GeV energies through undulators to induce synchrotron emission. Synchrotron is emitted spontaneously when charged particles are accelerated by an external force such as a magnetic field. To evaluate the differences between XFELs and synchrotron X-ray sources, several metrics which describe the characteristics of the source can be defined. The size of an X-ray source, $\Sigma_{x,y}$ is determined by the size of the creating electron beam, $\sigma_{x,y}$ and effects of the radiation generation process itself, σ_r.

$$\Sigma_{x,y} = \sqrt{\sigma_{x,y}^2 + \sigma_r^2}$$

The divergence of the X-rays, $\Sigma_{x',y'}$ depends on the divergence of creating electron beam, $\sigma_{x',y'}$ and radiation-generating process, $\sigma_{r'}$. Both these quantities can be expressed mathematically as:

$$\Sigma_{x',y'} = \sqrt{\sigma_{x',y'}^2 + \sigma_r'^2}$$

These equations show how the divergence and size of an X-ray beam is directly related by the size of the generating electron beam. If the electron beamsize and the divergence is small compared to the radiation generating effect ($\sigma_{x,y,x',y'} \ll \sigma_{r,r'}$) an X-ray source is said to be diffraction limited [37]. Typically, the divergence and size on each axis are expressed as the product: $\epsilon_{x,y} = \sigma_{x.y}\sigma_{x',y'}$ known as emittance. Spectral brightness is the most useful metric, especially for cross-comparison of sources. Brightness is defined as [38]:

$$B = \frac{\Phi}{4\pi^2 \Sigma_x \Sigma_y \Sigma_{x'} \Sigma_{y'}} \text{ at 0.1 \% energy bandwidth} \tag{1.1}$$

Where Φ is the spectra flux, the number of photons per second at the chosen bandwidth. Coherence is an important property of any X-ray light source. Temporal/longitudinal coherence describes the extent to which a wave maintains phase between two points in time. Often it is specified as a coherence-time after which temporal coherence is said to be 'lost' [37]. Transverse/spatial coherence is defined as the distance across the wavefront of a beam over which there is total dephasing between the two points [39].

1.2.1 Storage Ring/Synchrotron X-ray Sources

Over the past few decades, synchrotrons (also known as storage rings) have come to replace home sources as the primary way of producing X-ray photons to solve protein crystal structures. Currently, there are over fifty synchrotron facilities internationally that are used for macromolecular structure determination by protein crystallography [40]. Modern, 3$^{\text{rd}}$ generation, synchrotron facilities generate X-rays from long-lived (hours) electron bunches which are kept at relativistic energies (MeV–GeV) in a storage ring. Whilst in the storage ring, bending and focusing magnets are used to maintain the electron-bunch size and spread. The lifetime of each bunch is limited by collisions with residual gas and

so is topped up periodically to maintain the overall current. X-rays are generated by passing electrons through bending magnets, a wiggler or an undulator. Most modern MX beamlines use undulators which are a periodic array of dipole-dipole magnets and allow high-energy electron bunches to produce photon energies in the hard X-ray region which is ideal for crystallography.

In undulators, the magnetic field-strength and dipole separation is chosen to give in-phase oscillation of the electrons with respect to the emitted radiation. This amplifies the synchrotron radiation in the direction of the electron beam (z-axis) at the desired energy. The enhancement of harmonics means the energy spectrum of an undulator has enhanced spikes with a typical width (FWHM) of $\Delta E/E \sim 10^{-2}$ [41]. The exact details of X-ray generation and the angular divergence dependence is beyond the scope of this thesis. Overall, the total output power (P_{Total}) of an undulator integrated over the emission angle is given as [42]:

$$P_{Total} = \frac{2\pi ec}{6\lambda_u} N Z_0 I \gamma^2 K^2 \tag{1.2}$$

Where λ_u is the undulator period, N is the number of dipoles, I is the electron current (normally of the order ∼100 mA), γ is the Lorentz term and K is the so-called deflection parameter which is a dimensionless factor that depends on the magnetic field strength and undulator spacing. All other symbols have their usual meaning. Importantly, this equation shows the undulator output-power scales directly with the number of undulator dipoles. In most MX experiments, a double crystal monochromator is used to narrow the bandwidth of the generated radiation to around $\Delta E/E \sim 10^{-4}$ [43, 44]. The X-ray beam is then focused using Kirkpatrick-Baez (KB) mirrors [45] and compound refractive lenses (CRLs) which can achieve spot sizes on the order of 10 µm [46]. The spot size and therefore brightness is primarily limited by the emittance of the electron beam. For the generation of 12 keV X-rays using an undulator, an electron beam emittance of ∼ 8 pm rad is required to match the effects of the radiation generation process (σ_r) [47]. This is well below the current nm rad (horizontal) specification of most 3rd generation synchrotron sources.

Equation 1.1, shows the brightness of a source will increase with higher photon fluxes (achieved with more electron bunches/higher ring-current) or smaller beam sizes. Decreasing the emittance of the electron bunch leads to a smaller beam-size which thereby increases the brightness. The degree of spatial coherence is determined by emittance. Improved emittance leads to a higher level of coherence for emitted X-rays. The generation of radiation in an undulator can be considered as essentially a Gaussian spatial laser mode of uncorrelated emitters. This means there is a low level of spatial coherence in synchrotron X-ray beams [48]. Next-generation synchrotron sources aim to reduce the emittance of the electron beams [49]. This will be beneficial for techniques that require a high degree of coherence, such as coherent imaging, and also techniques such as serial crystallography which will benefit from the higher spectral brightness of smaller spot sizes.

In this thesis, synchrotron experiments used the P14-2 end station at Petra III (chapter 5), Germany and i24 end station at Diamond Light Source (chapter 3 and 7). At i24, the undulator period is 21 mm with 94 periods [50]. To tune the X-ray energy (between 6.4-20 keV), the gap between upper and lower magnet arrays can be adjusted. This alters the peak magnetic field and hence the photon energies. The distance between the undulator and the interaction region at the beamline is about 40 m [50]. The smallest spot size is 6×8 µm with a photon flux of around 10^{12} photons/s which gives a spectral brightness of around 10^{20} photons/0.1% b.w./mm^2/mrad at 10 keV.

The pulse structure of synchrotron X-rays is dictated by the number of electron bunches. The pulse duration is roughly equivalent to the electron bunch length. Different facilities can support different specific bunch modes. At Diamond Light Source, the standard operation mode has a filling pattern of 828 bunches 2 ns apart followed by a 72 bunch (144 ns) gap. The electron bunch pulse length is on the order of 10 mm (FWHM) long which leads to X-ray pulses around 30 ps in length [51]. Specialised bunch modes have been developed, along with fast shutter mechanisms for time-resolved measurements [52]. However, for the purpose of pump-probe experiments in this thesis, we will consider the synchrotron beam to be analogous to a CW laser as individual pulses are too weak to collect diffraction information from (unlike XFEL pulses).

1.2.2 X-ray Free Electron Lasers

Conventional lasers rely on the stimulated emission from atomic or molecular state to amplify light. This works well for visible light energies (∼1 eV) as it corresponds to electronic transitions with lifetimes which allow population inversion to be achieved with relative ease. To generate hard X-ray photons (∼10 keV) using the same methods would require excitation of inner-shell electron transitions which decay on serveral-femtosecond time scales [53]. This makes the accumulation of population inversion impossible at X-ray energies using atomic/molecular energy levels as a gain medium. Instead, free electron lasers (FELs) use electron bunches as a gain medium and can amplify radiation over a broad radiation spectrum, from infrared to hard X-rays [54]. The first hard X-ray free electron laser (XFEL) was the Linac Coherent Light Source (LCLS) which started operation in 2009 [35]. Most obviously, XFELs differ from synchrotrons as they accelerate electron bunches linearly and generate radiation in a single pass through a long undulator. This leads to enhanced amplification of X-rays and generates pulses with very high photon fluxes (∼10^{12} photons/pulse), low emittance (∼nm rad) with very short pulse lengths (∼10 fs) [55].

The most common mode of operation in XFELs is known as self-amplified spontaneous emission (SASE) mode [56]. In SASE mode, discreet electron bunches are accelerated to relativistic energies (14 GeV at LCLS [35, 57]) and then passed through a series of undulators (112 m of active undulator length at

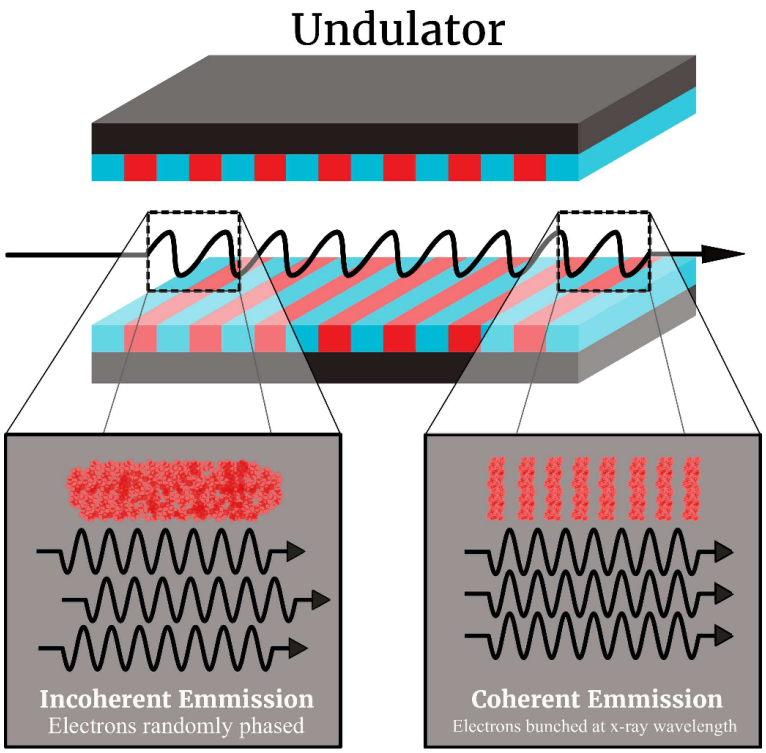


Figure 1.2: Schematic of the self-amplified spontaneous emission SASE operation mode at an XFEL. An electron bunch passes through an undulator and is transversely oscillated, initially emitting incoherent, synchrotron, radiation. As the intensity of this radiation grows it couples with the electrons transverse velocity causing a gradual split into microbunches which leads to coherent in-phase amplification of X-rays.

LCLS [35, 57]). Moving through the undulator causes the electrons to emit incoherent radiation (as in storage ring undulators). As the emission builds there is a gradual formation of in-phase-emitting micro-electron bunches (figure 1.2). This microbunching reduces the available phase-space for photon emission by each electron compared to the continuous incoherent emission in storage ring insertion devices. This leads to an exponential amplification in X-ray intensity through the undulator [55]:

$$P \propto \exp(z/l_g) \quad \text{where} \quad l_g = \frac{\lambda_u}{4\pi\rho} \tag{1.3}$$

where z is the distance through the undulator, l_g is the gain length, λ_u is the undulator period and ρ is the dimensionless FEL parameter. In SASE operation the amplified wavelength is seeded by spontaneous emission of electrons in the first section of the accelerator. This emission arises from the inherent shot-noise of the electron bunch and leads to stochastic behaviour of the X-ray spectrum [58]. The final X-ray pulse has a bandwidth of around $\Delta E/E \sim 10^{-3}$ but shows a spiked spectrum in the frequency domain [38]. During amplification of the pulse, a majority of the higher-order spatial modes become filtered out as the micro-bunching encourages the growth of the fundamental laser mode over others. This means XFELs have low emittance meaning the X-ray spot size can be focused very tightly [59]. Overall, the exponential amplification and high transverse coherence mean a spectral brightness on the order of 10^{32} photons/0.1% b.w./mm^2/mrad over a duration of $\sim$ 10 fs can be achieved in the hard

X-ray regime [35]. This is 12 orders-of-magnitude brighter than 3rd generation synchrotron sources and allows single XFEL pulses to be used to collect diffraction data from crystals in a technique known as serial femtosecond crystallography (SFX) [36, 60, 61].

The pulse structures and energies of XFELs varies between facilities but typically there are around 10^{11} photons/pulse at 10 keV, giving a pulse energy of between 0.5–1 mJ/pulse depending on the beamline [38]. The repetition rate at each machine also varies, 120 Hz at LCLS (California, USA) [57], 100 Hz at SwissFEL (Switzerland) [62] 30 Hz at SACLA (Japan) [63] and PAL-XFEL (Korea) [64] whilst the EuropeanXFEL (Germany) has 370-420 pulses (each separated by 2 ns) per a pulse train each of which is separated by 100 ms [65]. Experiments in this thesis used XFEL sources at LCLS, SACLA and PAL-XFEL. Of these three facilities LCLS offers the highest spectral brightness in the hard X-ray regime (1-3 mJ in a 1-2 µm focus size [35, 57]). The pulse energy densities at XFELs are so high that protein crystals are destroyed by a Coulombic explosion after the interaction and so must be constantly replenished for structure determination using SFX.

1.3 Studying Dynamics with X-ray Crystallography

Modern X-ray sources provide the opportunity to probe structure at high resolution, and on ultrafast time-scales. As mentioned, crystallography is the most successful structure determination technique in structural biology, making up around 90% of the 155,000 structures in the PDB [66]. Low packing densities, high solvent content and increased flexibility [67] mean that proteins can still be functionally active and undergo reactions whilst remaining in crystal form [68, 69]. Time-resolved Laue crystallography has been developed to measure changes in electron density from seconds to picoseconds in single crystals [70–73]. However, stroboscopic pump-probe experiments are limited due to the accumulation of laser and X-ray induced radiation damage [74]. Wide-angle X-ray scattering (WAXS) has shown that proteins can undergo conformational changes on the picosecond time scales [75, 76]. But, WAXS signals are very weak [77], meaning experiments rely on substantial averaging which cannot provide the same level of spatial resolution as crystallography, although it can be used to observe larger conformational changes [78].

Time-resolved crystallography measurements at XFELs [36] and synchrotron [79] are made using serial crystallography. In serial crystallography structures are obtained with data collection from multiple crystals. The idea was first implemented in the 1970s, when early crystallographers successfully merged intensities from multiple single-crystal rotation datasets to minimise the effects of damaging radiation by sharing the dosage [80]. In the late 1990s and early 2000s, third-generation synchrotron sources with lower emittance led to the development of micro-focus beamlines [81] and microcrystallography [82, 83]. Smaller X-ray spot sizes lead to enhanced diffraction signal in the small crystals as background

scattering from the solvent and mounting are reduced [84]. However, these smaller crystals also succumb more quickly to radiation damage. To solve this, in microcrystallography, rotation slices of a few degrees from multiple crystals can be collected in data-subsets which are then scaled and merged to determine the full structures.

With the development of XFELs it was clear this rotational method would not be suitable for structure determination. Not least because it would require the mounting of thousands of tiny crystals, as the highly intense XFEL pulses almost immediately obliterate samples. Instead, at XFELs, serial femtosecond crystallography (SFX) involves the collection of stationary (not rotated) diffraction patterns from the randomly oriented crystals before the onset of disorder due to radiation damage (diffraction before destruction) [60]. Recently, it has also become possible to make serial synchrotron measurements (SSX) using the same injection and data analysis methods of SFX [79]. Throughout this work SSX refers to the collection of stationary diffraction patterns, although partial-rotation datasets (similar to microcrystallography) techniques are sometimes referred to as SSX in the literature [85].

In general, reactions in crystals can be initiated through the addition of substrates [86, 87], buffer exchange [88] or through optical excitation [89]. Even in microcrystals, diffusion rates limit efficient reaction initiation to milliseconds [90] for all but light-activated reactions [91]. In photoactivated reactions, the generation of detectable photoproduct concentrations is limited by the primary quantum yield and photodamage thresholds with ultrafast excitation or a combination of the excited state lifetime, primary quantum yield as well as photodamage thresholds with laser pulses of longer duration (ns–ms). Given efficient reaction initiation there are three crystallographic techniques that can be used to resolve intermediate states: macromolecular cryo-crystallography (cryo-MX), time-resolved serial femtosecond crystallography (TR-SFX) and time-resolved serial synchrotron crystallography (TR-SSX).

1.3.1 Macromolecular Cryo-Crystallography

Whilst conventional cryogenic macromolecular crystallography (cryo-MX) is usually only used to collect static structures, ground-state intermediates can be captured by flash freezing crystals after reaction initiation [92]. The shortest lifetime of these intermediates is normally limited by the cooling rates for the particular crystal morphology and activation energy barriers involved in the initiated reaction [93]. Low temperatures can be used to increase lifetimes/decrease reaction rates and to accumulate or 'cryo-trap' an intermediate state (chapter 3). However, cryogenic structure determination methods can introduce artefacts caused by non-physiological temperatures or alter the reaction pathway [94]. While the vast majority of structures obtained are static/steady-state structures, cryo-MX remains an important tool for studying dynamics in proteins.

With modern synchrotron sources and detectors cryo-MX involves rotating flash-frozen crystals, mounted

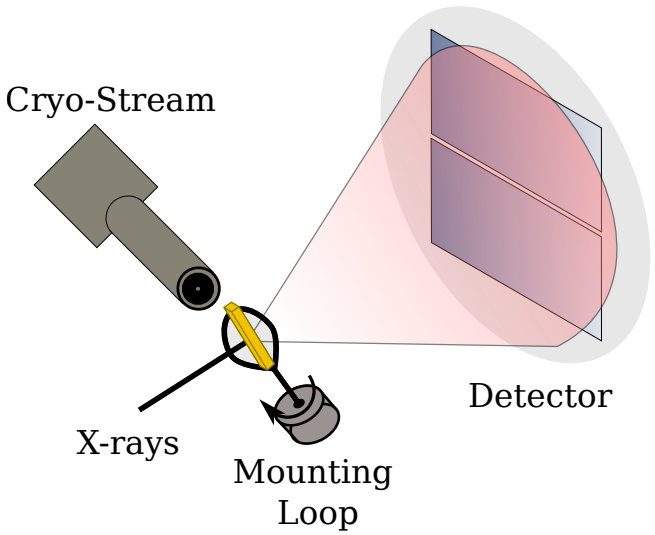


Figure 1.3: Cryo-macromolecular crystallography data collection. A crystal mounted in a loop is aligned and rotated. diffraction images are collected on the detector.

in nylon or polyimide loops, through a monochromatic ($\Delta E/E \sim 10^{-4}$ [43, 44]) X-ray beam whilst continuously acquiring images every 0.1-1° of oscillation with around 10 ms of exposure depending on the X-ray flux and crystal space group/unit cell size (figure 1.3). A cryostream is used to keep the crystal at 100 K during the measurement to prolong the crystal's lifetime and minimise the effects of radiation damage [95, 96]. Constructive interference of X-rays diffracting from the periodic crystal lattice leads to intense Bragg peaks also known as reflections. The intensity of these peaks is measured and indexed against possible space groups and unit cell dimensions. Symmetry equivalent intensities are then scaled and merged to determine the overall intensity of reflections from individual Bragg planes (hkl). These are converted to structure factors from which the electron density of the protein is determined from the phased Fourier transform. As the phases of the intensities are not measured they must be inferred with specific phasing experiments or, more commonly, with molecular replacement methods [97]. Overall, the time to collect a full dataset using a single crystal cryo-MX is on the order of a minute. For most reactions, this is far too slow to resolve useful rate dependant information even if the effects of radiation damage caused by multiple collections can be outrun.

1.3.2 Time-Resolved Serial Femtosecond Crystallography

With the development of hard X-ray FELs, the field of serial femtosecond crystallography (SFX) has emerged [36, 60]. In SFX, single XFEL pulses are used to obtain stationary (also known as partial) diffraction patterns from liquid samples containing high densities of randomly oriented micro/nanocrystals.

The still images acquired in serial crystallography means that for a given miller index (of indices h,k,l) the full reflection intensity is not measured. This is known as the partiality problem and occurs because

a single X-ray exposure of a stationary crystal can only access a thin slice of the Ewald sphere through the Bragg condition [98, 99]. The inherent crystal mosaicity contributes to each reflection having a finite 'volume' in reciprocal space [100]. This is dissected when using monochromatic radiation which means the reflection intensity may be only partially measured. In cryo-MX measurements, rotation of the crystal avoids this problem and allows for the measurement of full reflection intensities. In serial crystallography this problem is solved by making highly redundant measurements across large numbers of crystals. Sample injection randomly orients crystals (following a Poisson distribution) meaning adequate sampling of each crystal orientation and therefore reflection intensities can be achieved with the collection of a large enough number of diffraction patterns. A Monte-Carlo approach can then be used to integrate and merge the final hkl intensities as well as estimating associated errors (section 1.4.1) [101, 102].

In SFX, each X-ray pulse is well beyond the tolerable dose of the sample so each diffraction snapshot originates from a new crystal. This means the effects of radiation damage are limited to processes occurring on the ∼10 fs time-scale of the XFEL pulses [60]. In conventional crystallography, radiation damage from X-rays involves the disintegration of crystal structure and specific chemical changes caused by radicals [103]. As these processes occur on time-scales longer than the XFEL pulse duration, SFX measurements can essentially outrun most radiation damage effects [104]. However, in proteins with heavy atoms [105, 106], or when using excessively high X-ray fluxes [107], a significant population of atoms can be photoisomerised which can lead to specific or global damage on the timescale of the XFEL pulse [108]. To suppress these effects, either shorter XFEL pulses can be used (ideally shorter than the Auger decay lifetime of carbon, around 10 fs [109]), and/or the X-ray energy density limited (ideally to below 10^{19} W/cm^2 for sub micron focus sizes [108]). Overall, for the study of light-atom proteins with moderate XFEL beam intensities, SFX can be used to collect essentially damage free structures at room temperature.

As SFX can generally outrun radiation damage, it is well-suited to determining structures from crystals much smaller than is possible with cryo-MX (even with the development of microcrystallography [82]). The intensity of a Bragg peak from a diffracted X-ray beam (I_{Spot}) is approximately proportional to the product of the X-ray beam intensity (I_{Beam}), crystal volume (V_{XTAL}), magnitude of the structure factor ($|F_{hkl}|$) and unit cell volume (V_{Cell}) [110]:

$$I_{Spot} \propto I_{Beam} \frac{V_{XTAL}}{V_{Cell}^2} |F_{hkl}|$$

This means that for small crystals, more intense X-ray beams are necessary to make high signal-to-noise measurements, especially if the unit cell size is large. However, in cryo-MX, smaller protein crystals are damaged within fractions of seconds which prevents the collection of usable data [84, 111]. This problem is especially exacerbated for membrane proteins which often crystallise in multimeric unit

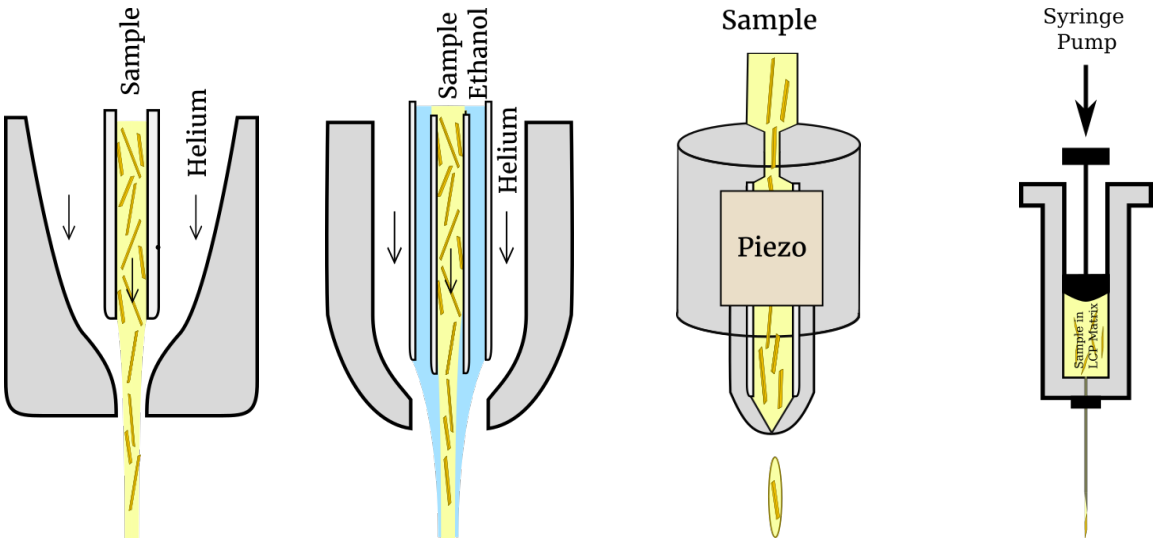
cells with very large sizes [112]. SFX is therefore one of the only viable methods for determining the structures of these types of crystals.

Dynamics can be studied using time-resolved SFX (TR-SFX) where measurements can be made analogously to flash photolysis [113], using optical excitation as a 'pump' followed the hard X-ray diffraction as a probe. By varying the delay between the pump laser and the XFEL, structural snapshots at different time-points can be collected to build up a 'molecular movie' of the reaction being studied. A range of time-delays can be collected at XFELs. For sub 1 ps measurements, there is an inherent jitter between the synchronised optical laser and X-rays which limits the temporal resolution. This is due to the stochastic nature of the SASE operation mode. Depending on the facility, this jitter can have a bandwidth between 0.1 - 2 ps [57, 64, 114], much longer than the ∼100 fs optical pump pulses [115] and ∼30 fs XFEL pulse [114]. Delays are corrected using a timing tool which measures the relative arrival time for each pulse and allows binning of diffraction patterns into sub-datasets (summarised in section 4.2.4). Longer delays from 1 ps - 1 ns can be set using optical delay stages but do not require binning using the timing tool. Delays from ns - µs are not practical to achieve using a delay stages so are normally obtained by offsetting the phase of the laser synchronisation signal [115]. Delays beyond the repetition rate of the XFEL (8.3 µs at LCLS) require highly controlled injection rates so samples can be illuminated before they have transited into the interaction region [116].

Sample Delivery Methods in Serial Femtosecond Crystallography

Sample delivery remains the largest bottle-neck in SFX experiments. Crystals need to be delivered at a rapid exchange rate with high stability, to maximise hit-rate whilst minimising sample wastage/consumption. A majority of early SFX experiments used liquid jets, in part because this was one of the only methods which is compatible for injection into the vacuum chambers which is necessary at the CXI end station as the detector only operates in vacuum [117, 118]. Rayleigh liquid jet injection involves the flow of a dense crystal suspension (∼10^7 crystals/ml) through a thin (50-200 µm inner diameter) glass capillary at high flow rates (10-200 µl min^{-1}). This forms thin focus with diameter 20-50 µm [119]. The gas dynamic virtual nozzle (GDVN) improved on the Rayleigh jet by using a flow of Helium around the internal capillary to focus the sample to around 10 µm [120]. This better matches the XFEL beam size, lowering background scatter and sample consumption rates. Further developments of liquid jets include double flow-focusing nozzles (DFFNs) which adds a second external sheaf of ethanol between the sample and helium. This further focuses the sample jet, improves stability and can reduce sample consumption by up to an order of magnitude compared to GDVN [121]. Jet methods typically cannot be used for delays beyond 100 µs as the flow rate has to be constantly adjusted to maintain stability meaning the pump laser must be aligned at the interaction region.

Droplet injectors are similar to jet methods in terms of sample densities but they have much lower



(a) GDVN **(b)** DFFN **(c)** Droplet Injector **(d)** Syringe Injection

Figure 1.4: Droplet and jet injector methods typically used at SACLA and LCLS for time resolved serial femtosecond crystallography experiments.

sample consumption rates. Instead of a continuous stream of the sample, droplet injectors use acoustic actuators to eject single droplets in time with the arrival of the XFEL pulses [122, 123]. The properties of the droplets can be varied by changing the amplitude of the acoustic wave which propagates through the sample that forms the droplet. This can be used to change the droplet shape/thickness for optimised hit rates/minimised background scatter. The lower sample consumption rates mean that samples with lower expression yields become viable targets.

More recently, fixed target sample injection systems have been developed for use in SFX measurements. These delivery systems use a 'fixed' medium which is loaded with crystals and translated through the XFEL beam using stages. There are several variants [124–126], but here we will focus on the one developed by Owen *et al.* [127]. In this fixed target system, a dense ($\sim 10^6$ crystals/ml) crystal suspension is spread over a silicon chip with lithographed openings around 10×10 µm in dimension. Removal of crystal mother liquor traps the microcrystal in the holes while a stepper motor is used to probe each opening. This method has a much lower sample consumption rate than liquid jet injection and can achieve up to 100 % hit rates with spectral mapping of successfully loaded wells prior to injection [128]. Additionally, a hit-and-return method has been developed which allows excitation in a pump 'run' and followed by return to probe at set delays [129, 130].

Viscous jet methods involve re-suspension (or growth) of microcrystals in a lipidic cubic phase (LCP) such as monoolein [119]. For membrane proteins, LCP can be the only stable medium for crystallisation/injection [131]. The high viscosity of the LCP suspension can be loaded into a syringe for sample delivery. This allows much slower extrusion speeds than similar liquid-jetting methods which

significantly reduces sample consumption. Overall, sample consumption per structure can be very high in SFX experiments. Per a structure, liquid jet methods typically require tens of mg of protein samples, whilst fixed target and viscous injection require sub-mg [132]. Droplet injectors lie in between - requiring close to 1 mg of protein per structure [122]. However, currently, only liquid jet methods are suitable for high vacuum environments (such as those at the LCLS CXI end-station) and use at MHz repetition rates such as those at the EuropeanXFEL.

1.3.3 Time-Resolved Serial Synchrotron Crystallography

Improving synchrotron technology with higher spectral-brightness [51] and faster-frame rate detectors [133] have meant it is possible to make time-resolved serial synchrotron crystallography (TR-SSX) measurements using broadly the same methodology as TR-SFX: the collection of X-ray diffraction snapshots before the rapid onset of disorder and radiation damage [36, 104]. Serial crystallography experiments at synchrotrons can use the injection methods established at XFELs for sample delivery [44, 79, 129, 134, 135], although generally liquid jets are not used as the transit times through the X-ray focus is too quick and sample consumption rates are too high to justify. Although TR-SSX requires much larger quantities of crystals compared to cryo-MX/conventional crystallography it allows room temperature and therefore dynamic measurements in addition to lowering the radiation dosage per crystal which is useful for damage sensitive samples [136].

Single synchrotron pulses are not intense enough to give measurable diffraction patterns when using undulator/monochromator combination installed at most MX beamlines. Instead, detectors are normally electronically gated over nanosecond time-scales at millisecond repetition periods for data collection and the synchrotron beam is considered essentially a CW X-ray source. This limits the time-resolution of TR-SSX measurements to the millisecond domain. Initially, it was thought that much higher multiplicity would be required compared to SFX measurements as the energy bandwidth is lower at synchrotron beamlines ($\Delta E/E \sim 10^{-4}$) compared to XFEL ($\Delta E/E \sim 10^{-3}$) [79]. It was thought this may increase the partiality problem meaning SSX would require much higher multiplicity compared to SFX [137]. However, the first SSX experiments have shown this is not the case, potentially as crystals can translate/rotate over the longer exposure times [138]. This means statistics typically converge with around 10,000 indexed images, as with SFX. Laue beamlines that produce polychromatic X-ray beams can achieve higher time resolutions (up to 100 ps [44, 139, 140]) and require fewer indexed patterns for convergence [44], but are not as widely available as monochromatic beamlines. One of the attractions of TR-SSX is the ease at which traditional crystallography beamlines could be retrofitted with fixed target systems for TR-SSX experiments.

Typically, a protein crystal will absorb around 2% of X-rays photons primarily through the photoelectric

effect (1.5%) and Compton scattering (0.15 %) [141]. This absorbed energy leads to stronger thermal vibrations as well as chemical reactions caused by the generation of radical species which propagate further through the crystal lattice at higher temperatures [142, 143]. In TR-SSX measurements, crystals are singly exposed to the X-ray beam for around 10-20 ms. This short exposure reduces the overall radiation dosage required to collect a structure by around an order of magnitude compared to cryo-MX or multi-crystal averaging techniques (the combination of partial rotation datasets) [85]. However, at room temperature, crystals have significantly lower tolerable dose limits (around two orders of magnitude lower) than crystal at 100 K [141, 142]. As TR-SSX is a relatively new technique, understanding the net effects of this trade off is not clear and is an ongoing research topic. D. Mora *et al.* [144] recently suggested that room temperature crystals in SSX measurements should not exceed an average radiation dosage of 0.2 Gy for damage free data collection. This corresponds to X-ray exposure time of 20-30 ms for a 3 µm cubic crystal using microfocus beamlines with no X-ray attenuation [43, 51]. As data of sufficient signal to noise can typically be collected using dosages of around 0.1 Gy, TR-SSX seems like a viable technique for low-dosage crystallography measurements. The thermal effects of X-ray induced heating in TR-SSX are currently less well understood. The thermal coupling of crystals to their sample environment is not as clear and varies for different injection techniques. In chapter 5, the heating effects of TR-SSX are explicitly investigated and derived from the kinetic behaviour of protein crystals exposed to X-rays.

1.4 Analysis and Interpretation of Time-Resolved Crystallography Data

The merging of measured intensities and conversion to structure factors has become essentially automated for simple cryo-MX experiments. By comparison, the processing of serial crystallography data, is much more demanding, partially, as the software tools are less mature but also as large volumes of data are collected (typically ∼240 Gb of raw data per a structure). Furthermore, in time-resolved experiments, analysis is often complicated, due to the heterogeneous populations of intermediate/ground-state structures that need to be properly analysed. In this section, an overview of the data processing and analysis techniques used in time-resolved crystallography experiments is presented.

1.4.1 Data Analysis in Serial Crystallography

As only a limited number of frames collected in serial crystallography experiments contain Bragg peaks, one of the first steps in processing is to perform hit-finding. Ideally, this is done in real-time to provide feedback for sample optimisation during the experiment. Data-rates in SFX experiments can be as high as 720 Mb/s (LCLS at 120 Hz) which would make indexing all images very slow/computationally expensive. As a result, hit-finding can also be used to filter images for indexing which is also useful for

reducing disk space and required computation time for processing of diffraction patterns. A number of programs have been designed for hit-finding/pre-processing images (section 4.2.5): Cheetah [145], OnDa [146], psocake [147] and CASS [148]. As well as hit-finding, these programs can correct for detector and experiment artefacts with detector masking and background subtraction.

For indexing and integration of the images, there are currently two main software suites: CrystFEL, developed mainly by T. White and co-workers [102], has a number of options for indexing algorithms: MOSFLM [149, 150], DIRAX [151], XGANDALF [152], XDS [153], FELIX [154], TakeTwo [155] and pinkindexer [156]. DIALS, developed mainly by N. Sauter and co-workers [157], is built upon the computational crystallography toolbox (cctbx) [158] and uses cctbx.spotfinder [159] and specific still indexing algorithms [157, 160]. Both programs have options for merging with or without post-refinement of partialities (section 4.2.5).

After pre-processing, the first step in data analysis is indexing. Indexing assigns the crystal orientation and unit cell size to each image. Ideally, prior knowledge of the space group and approximate unit cell size is known from cryo-MX measurements which can be used as a starting point. To achieve reliable indexing, the location of spots on the detectors must be known accurately in 3D space. As SFX beamlines are highly configurable they often use custom detectors which can change in configuration between experiments. For example, at the CXI end station at LCLS, the CSPAD detector is made of 32 individual detector panels which are replaced periodically as they are damaged [117]. As a result, geometry optimisation of the detector distance, and relative location of each detector panel must be refined for accurate indexing. Both CrystFEL and DIALS have built-in functions for this purpose [161–163].

The final step of processing is to integrate/merge the indexed Bragg peaks. As mentioned above, a Monte Carlo method is used to approximate the solid angle integration. In both CrystFEL and DIALS, there are also a number of options for post-refinement. Post-refinement involves the re-refinement of the geometry/unit cell for each image within the context of the final distributions of these indexing parameters. This can be used to overall improve the final merged data merging statistics [164]. During this step, data quality metrics are calculated for the overall dataset (detailed in the next section). If the data quality is sufficient then the structure factors can be calculated and phased. The best method for analysis light-induced structural differences in time-resolved crystallography data is using Fourier difference maps (section 1.4.3). In chapter 4, a pipeline for systematic data analysis of TR-SFX data is presented with exact details of the data analysis steps.

1.4.2 Metrics for Data Quality Assessment

Macromolecular Cryo-Crystallography

In cryo-MX data collection one of the main indicators of the quality of data is assessment of the residual, R_{merge} parameter (sometimes denoted R_{sym}) which measures the spread of n independent measurements over each hkl reflection for each reflection intensity, $I_i(hkl)$ around the mean/merged intensity, $\langle I_i(hkl)\rangle$ [30]:

$$R_{merge} = \frac{\sum_{hkl}\sum_{i=1}^{n}|I_i(hkl) - \langle I_i(hkl)\rangle|}{\sum_{hkl}\sum_{i=1}^{n} I_i(hkl)}$$

More consistent data will have a lower R_{merge}. However, R_{merge} may increase with higher redundancy as each individual reflection influences the overall average intensity. Hence, Diederichs & Karplus [165] introduced the metric R_{meas}, which adjusted R_{merge} by a factor of $\sqrt{n/(n-1)}$ to normalise against multiplicity:

$$R_{meas} = R_{merge}\sqrt{\frac{n}{n-1}}$$

Typically, overall values below 0.1 are considered good although it is typical to calculate R_{merge}/R_{meas} by resolution bin and discard bins above 0.6-0.8 [166]. Another important metric for interpreting data quality is the signal to noise ratio (SNR). This is calculated as the overall average of the merged intensity for each reflection divided by the merged error ($\langle\sigma_I\rangle$):

$$SNR = \left\langle \frac{\langle I_i\rangle}{\langle\sigma_{I_i}\rangle} \right\rangle \tag{1.4}$$

Generally, a signal to noise cut off of around 2.0 is desirable. However there are variations in how values of σI are estimated which can make comparison of SNR unreliable in cryo-MX [166]. More recently, a measure known as CC* has been introduced which calculates a Pearson correlation coefficient between two halves of the data set ($CC_{1/2}$) [166]:

$$CC* = \sqrt{\frac{2CC_{1/2}}{1 + CC_{1/2}}}$$

For evaluating the accuracy of refinement, the fit of model coordinates can be assessed with the residual R factor between the observed (F_{obs}) and calculated (F_{calc}) structure factors:

$$R_{Work/Free} = \frac{\sum_{hkl}|F_{obs}(hkl) - F_{calc}(hkl)|}{\sum_{hkl} F_{obs}(hkl)}$$

Where work/free refer to sets of structure factors known as the 'working' - those used in reciprocal space refinement (normally 90-95% of the entire set) or the 'free' - those not used for refinement (normally 5-10% of the set). The use of free structure factors helps reduce model bias. The appropriate values for $R_{Work/Free}$ depend on the resolution of the data, but typically below values below 0.25 are considered good.

Serial Crystallography

As serial crystallography data is collected and analysed in a very different way to conventional crystallography data, some of the metrics used to evaluate data quality are no longer appropriate. For example, R_{Merge} depends on the variance of the measured reflection intensities which is naturally large in serial methods, even for high-quality data. Instead the statistic, R_{Split} was introduced by White *et al.* [102] as a figure of merit for assessing the merged quality of serial crystallography data. The consistency of the data is measured by splitting the intensities in two halves with 'odd' and 'even' reflection indices. The R_{Split} residual is defined as:

$$R_{Split} = \sum_{hkl} \sqrt{2} \frac{\sum_i^n |I_{even} - I_{odd}|}{\sum_i^n (I_{even} + I_{odd})}$$

Which will ideally be close 0. The SNR is also calculated differently in serial data. As only partial reflections are measured, the SNR for each reflection cannot be calculated based on the standard deviation of all the measured reflections (as expressed in equation 1.4). Instead the error of each reflection measurement is calculated from the Monte Carlo integration as: [36]:

$$\sigma_{I_i} = \frac{\sqrt{\sum_{hkl}(I_{spot}(hkl) - \langle I(hkl) \rangle)}}{N_{hkl}}$$

Based on the number of reflection observations (N_{hkl}). This is then used to calculate overall SNR as usual. Given adequate spot intensities (I_{Spot}), but a low number of measurements, the SNR will strongly depend on the number of reflection observed. Hence, the SNR should be interpreted cautiously if too few reflections have been measured for a given resolution shell [102].

1.4.3 Fourier Difference Maps

In time-resolved crystallography, the observation of light-induced structural changes are limited by the population transfer size (the photolysis yields) and the size of the real-space atomic displacement. This means that differences are small and not easily analysed by directly comparing superposed refined-structure-coordinates. Isomorphous Fourier differences allow a more sensitive comparison of structures, especially in the case of partial occupancies. To compare two structures, ideally, it would be possible to directly calculate the difference electron density (DED) ($\Delta\rho$) of dark (ρ_{Dark}) and light-excited (ρ_{Light}) species. This is equivalent to taking the differences in the Fourier transform of the complex structure factors ($\tilde{F}$):

$$\Delta\rho = FT\{\tilde{F}_{Light}\} - FT\{\tilde{F}_{Dark}\}$$

Where the complex structure factors ($\tilde{F}_i$) can be expressed in terms of their phases (ϕ_i) and amplitudes(F_i):

$$\tilde{F}_i = \mid F_i \mid \exp(i\phi_i)$$

However, as the excited-state structure is normally unknown it is difficult to provide accurate phases for the light-state. If the light-induced structural changes are small in atomic displacement, then the

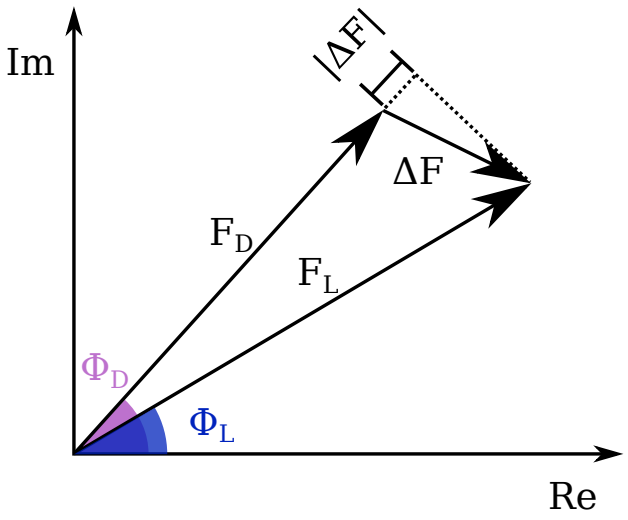


Figure 1.5: Complex plane of the structure factor triangle for light induced structural difference calculation.

phases for the dark and the light structures will very close and so are approximated as: $\phi_{Dark} \approx \phi_{Light}$. This gives the equation for finding isomorphous differences as:

$$\Delta\rho = FT\{(\mid F_{obs,Light} \mid - \mid F_{obs,Dark} \mid) \exp(i\phi_{calc,Dark})\} = FT\{\Delta F \exp(i\phi_{calc,Dark})\} \tag{1.5}$$

Where subscript 'obs' indicates observed structure factors and 'calc' indicates values calculated from the model coordinates. Given uncorrelated phases between two conditions, the difference structure factor magnitude ($\Delta F = F_{obs,Light} - F_{obs,Light}$) will account for only half the true value [167] (figure 1.5) which means this approximation can underestimate the magnitude of real-space structural changes. However, as the light and dark structure phases are generally highly correlated, the difference structure factor magnitude using dark phases should be close to the full magnitude using the difference phases. There are cases where, in single crystals measurements (cryo-MX), this approximation does not hold. For example when comparing crystals with large unit cell variations or in proteins which undergo vey larger conformational changes where the new state may have a significantly different structure and therefore phases (or unit cell size) [168]. In serial crystallography measurements the merged unit cell size is calculated from the thousands of crystals so is normally very consistent between conditions. Additionally, in this thesis, the light-induced structural changes only effect a small number of atoms, meaning this phase approximation hold (although this is discussed in more detail in section 4.3).

Scaling

Structure factor differences occur due to real atomic displacements but also due to systematic (local), random and overall (global) discrepancies. In cryo-MX this can be due to different quality/sizes of crystals as well as differences in experimental setup such as X-ray intensity or detector defects. In serial crystallography analysis (using CrystFEL), different data conditions are merged and scaled separately which can lead to different relative scales for reflection intensities, particularly if the redundancy differs

across the datasets being compared. Therefore, to calculate accurate structural differences, scaling must be applied to account for these variations [169, 170]. Global scaling ensures that the intensities of both datasets are being projected onto the same overall intensity scale and involves fitting a linear scale factor using the method of least squares [168, 169] or Wilson scaling [171, 172]. In general global scaling follows normal probability analysis [169] to calculate a scaling factor, K, which minimises the sum over reflections i and j:

$$\sum_{i,j} \left| \frac{(F_{i,Obs,Dark} - KF_{j,Obs,Light})}{\sqrt{\sigma^2(F_{i,Obs,Dark}) + K^2\sigma^2(F_{j,Obs,Light})}} \right|^2 \tag{1.6}$$

Where σ^2 represents the variance over the individual (i and j) structure factors.

Local scaling can account for different sensitivities in detector regions and variation in experimental set-up between beamlines. Generally, local scaling calculates a scale factor, for individual subsets of reflections, again minimising the sum in equation 1.6. As local scaling is based on smaller subsets of reflections it is statically more prone to error. The decision to apply local scaling to a particular reflection and area Fourier space is made by comparing the average systematic variation in local scale-factors ($\langle \Delta K \rangle$) with the overall average random error for a dataset ($\delta(K)$). Specifically, when $\langle \Delta K \rangle > \delta(K)$ local scaling is applied [168].

Q-Weighting of Difference Structure Factors

Sources of noise in electron density are predominately due to poor quality phases as X-ray exposure times/intensities are chosen to give accurate structure factor measurements [167]. However, there are cases where statistical/random noise becomes significant and degrades the accuracy of structure factor measurements. This can be due to bad segments on a detector (which are hopefully already masked out) or from the reduction in structure factor intensities in the highest resolution shells. To account for these errors, in conventional difference density maps, the largest ΔF values, beyond a set threshold are discarded. In Q-weighted maps, given the determination of reliable σ_I (and corresponding σ_F) values, prior knowledge of the structure factor distribution [173] and error distribution [174] can be used to improve the reliability of calculated difference structure factors.

This was first implemented by Ursby *et. al* [174, 175] who used the interpretation of Baye's theorem by French and Wilson [172] to improve the estimation of difference structure factor amplitudes between two data sets. They assumed a Gaussian distribution of errors for each structure factor and calculated a weight, w which should be applied to each individual (i) structure factor difference ($\Delta F_i = F_{i,Light} - F_{i,Dark}$):

$$w_i = \frac{1}{1 + \frac{\Delta F_i^2}{\langle |\Delta F| \rangle^2} \frac{\sigma^2_{\Delta F_i}}{\langle |\Delta \sigma_{\Delta F}| \rangle^2}} \tag{1.7}$$

Where $\sigma_{\Delta F_i}$ is the variance of the structure factors difference and $\langle X \rangle$ represent the mean values of

a variable. The variance of a structure factor difference is calculated as mean of the light and dark structure factor variance: $\sigma^2_{\Delta F_i} = \sigma^2_{F_{i,Light}} + \sigma^2_{F_{i,Dark}}$.

This weight is applied as $w\Delta F$ for each structure factor difference. This type of weighting is known as Q-weighting. To produce the final DED maps the Fourier transform is taken over all calculated structure factors. Q-weighting does not reduce the error due to poor quality phases, however, it is particularly useful for the interpretation of small structural displacements at high resolution. It is therefore the most effective was of resolving light-induced structural differences in crystallographic data.

Extrapolated Structure Factors

To relate difference density signals back to structural/chemical changes in the protein, extrapolated structure factors and corresponding real space maps can be generated. A model can then be refined to the real space map to give the best representation of the light-induced structure. To generate this extrapolated map the difference structure factors, ΔF are added in a linear combination in multiples of N_{EXT} to the calculated dark structure ,F_{Dark}, to generate extrapolated structure factors, F_{EXT}. The negative electron density around areas of interest in the protein is then integrated and extrapolated to find the characteristic N_{EXT} value. The population transfer (PT) can then be approximated from the N_{EXT} value as $PT = 2/N_{EXT}$.

1.4.4 Summary

Throughout this thesis, light-induced structural differences from a reversibly switchable fluorescent protein are studied. CrystFEL is used to process serial crystallography data after and pipeline resulting in the calculation of Q-weighted DED is developed for the analysis of TR-SFX data (section 4.3). Q-weighted maps are used throughout for analysis of crystallographic data with custom-written scripts invoking PHENIX [176] and CCP4i [177] modules for the calculation of structure factors, scaling and the final DED generation. For the final interpretation of DED extrapolated structure factors and density is generated to allow the modelling of light-state coordinates (section 4.3).

1.5 Ultrafast Spectroscopy

Some of the earliest ultrafast spectroscopy studies focused on biological processes. In particular, the cis/trans isomerisation of the retinal chromophore was one of the first ultrafast spectroscopic measurements made in the 1970s [14]. A picosecond component of the isolated chromophore was measured and later related to the mechanism which drives vision in the protein bacteriorhodopsin in the late 1980s [12]. Through the 1990s, commercially available lasers capable of generating ultrashort pulses drove the development of ultrafast spectroscopy [113]. In particular, the technique of femtosecond transient absorption (TA) is one of the most commonly used methods to investigate photolysis reactions

on ultrafast timescales [178].

1.5.1 Infrared Transient Absorption Spectroscopy

Vibrational spectroscopy of proteins, in particular the interpretation of difference spectra can be used to assign localised changes in chemical structure (conformation, redox state, bond angles, properties of surrounding amino acids, hydrogen bonding environment) in proteins [179]. Steady-state vibrational spectra are normally acquired using Fourier transform infrared (FTIR) spectroscopy while transient spectra can be studied infrared transient absorption spectroscopy (IRTA).

In IRTA measurements, a small volume of sample is excited using an optical pump pulse to a higher energy optical/vibronic excited state. A weaker, probe pulse, overlapped with this volume is then used to measure changes in absorbance (ΔA) as a function of delay (T) and wavelength (λ). Usually, the pump pulse is chopped to half the frequency of the probe pulse and the absorbance changes are recorded as the difference:

$$\Delta A(\lambda, T) = A(\lambda, T) - A(\lambda)$$

TA measurements therefore yield an absorbance difference measurement as a function of wavelength and delay which contains spectral information of the molecular processes occurring. Normally, the data are fitted using global or target analysis to reduce the wealth of information into assignable spectral components (section 1.5.3).

Signals measured in TA experiments can be characterised as a ground state bleach (GSB), stimulated emission (SE) or excited state absorption (ESA) [180]. GSB signals are negative changes in absorption relative to the ground state spectrum. Physically, they represent a reduction in the population of the ground state species after it has been excited. SE are signals negative in ΔA that generally follow the fluorescence spectrum and do not correspond to the ground state absorption (although the GSB and SE bands can overlap). The origin of SE signals can be understood from Einstein's treatment of a two-level electronic system. The coefficients for spontaneous absorption/emission (B_{12}/B_{21}) are derived to be equally probable. This means an excited system can emit a photon given the presence of a resonant pulse (probe). This is observed as an increase in the probe intensity and therefore a negative ΔA signal. ESA is observed from optically allowed transitions of the excited state species which were created by the pump pulse. In cases where the GSB overlaps with the ESA, the excited state spectrum will not resemble components of the steady-state spectrum.

After the excitation of a photochemical system, reactions can occur which lead to the formation of photoproduct species. This will be normally be observed as a long-lived spectral component that may be assigned to a corresponding ground/steady-state difference spectrum. In biological systems, conformational changes are often driven by ultrafast dynamics [18], but the final product state may

evolve more slowly (potentially on ms timescales). It is therefore important to characterise the evolution of spectral components over the broadest delays possible to observe the product sate formation. The time-bandwidth product ($\Delta t \Delta \nu = 2ln(2)/\pi \sim 0.44$) for a Gaussian laser pulse limits the spectral and temporal selectivity of pump/probe pulses in TA experiments. In the case of a 400 nm pump pulse of 80 fs duration, the transform-limited bandwidth would be around ∼3 nm. This is more than reasonable for investigating most ultrafast reactions and means that pulse lengths/bandwidths are not typically the limiting factor in spectroscopy experiments.

1.5.2 Implementation of Ultrafast Infrared Transient Absorption

Modern-day TA setups can be broadly divided into two main categories depending on the repetition rate of the fundamental laser pulse. High repetition rate systems generate low energy pulses (0.5 - 10 nJ at the sample) at rates in the range of 40-250 kHz [178]. Alternatively, low repetition rate systems produce higher pulse energies (∼100 nJ at the sample) as rates between 1-5 kHz [178]. In this thesis (chapter 6), a 1 kHz Ti:Sapphire laser system was used to generate infrared probe light and visible pump pulses. As a result, low repetition rate systems and vibrational spectroscopy will be the focus of the following section.

In chapter 6, a Spectra Physics Spitfire Pro laser system is used to generate 70 fs 800 nm pulses. The system generates the final 1 kHz, 4 mJ pulses through regenerative amplification of the pulses. Firstly, a Maitai broadband oscillator (Spectra Physics) is used to generate a ∼30 fs duration, 800 nm, 80 MHz pulse-train one of which is gated into the main cavity by a Pockel cell at 1 kHz as a seed pulse. The seed pulses are too weak for performing spectroscopic TA measurements so needs to be amplified. Efficient amplification is only possible using a technique known as chirped pulse amplification (CPA) due to the extremely high instantaneous powers reached with femtosecond pulses [181]. The seed pulse is stretched to ∼100 ps duration before being focused into the Ti:Sapphire gain medium. Population inversion is created in the Ti:Sapphire crystal by pumping with a 532 nm nanosecond Q-switched Nd:YLF (yttrium lithium fluoride) Empower (Spectra Physics) laser. As the stretched seed pulse travels through the rod it induces stimulated emission which reaches saturation after 10 passes. The pulse is then re-compressed using a grating compressor to the final pulse length (70 fs) at a 1 kHz repetition rate with 4 mJ of energy at 800 nm. Visible pump pulses can be generated from the harmonics of the fundamental (400/200 nm) and/or from the output of an OPA (240-2600 nm). The mid-IR probe beam (2-10 µm) in this system is generated from difference frequency generation (DFG) of the signal and idler beams from an optical parametric amplifier (OPA) which are driven by the fundamental beam. The exact details of the experimental setup are detailed in section 6.2.3.

The pump and probe pulses are focused and overlapped temporally and spatially in the sample. The

pump energy is chosen to excite between 2-20 % of molecules to ensure a linear response to the pump. The single measured pump-on/pump-off change in absorption (ΔA) is generally not of high-enough signal to noise quality in IRTA. This is due to fluctuations in the pump and probe intestines and caused by imperfections in the sample as it is translated during the measurement [182]. To increase the signal to noise of the overall measurement, typically, a few thousand absorption difference measurements are made for each time-point and averaged. Normally, around 100 pump-probe delays are collected between from negative pump-probe delays to nanosecond delays. The final measurement output of a TA experiment is a 3-dimensional vector of delay, wavelength and averaged change in absorption. As the first step, a background subtraction is made using the earliest negative time point. This corrects for signals caused by accumulative averaging [183] and long-lived photoproducts caused by pump scatter/overlap with the previous pulse. A dispersion correction (usually a polynomial of 2nd degree) is applied to correct for temporal chirp the probe. In IRTA measurements the probe pulse normally requires very little chirp correction (first-order correction around 0.1 $\mathrm{ps\,nm^{-1}}$). The temporal resolution of an experiment is limited by the convolution of the pump and probe pulses which is known as the instrument response function (IRF). The IRF is modelled as a Gaussian function and is fitted for each wavelength, as described in section 1.5.3

1.5.3 Data Analysis

Global fitting of spectroscopy data provides a convenient way to reduce the dimensionality of spectroscopy data into time-dependent components as a function of wavelength [184, 185]. Van Wilderen *et al.* developed the ultrafast fitting toolbox [186] for analysing spectroscopy data. The toolbox has options to perform chirped pulse (dispersion) corrections, fitting of the IRF, global/target analysis and singular value decomposition of spectroscopy data.

Chirp correction for the central arrival time (μ) is fitted using a polynomial to the n^{th} order with coefficients (μ_i), around time-zero (t_0) and a central wavelength (λ_c)[186]:

$$\mu = t_0 + \sum_{i}^{n} \mu_i \Big(\frac{\lambda - \lambda_c}{100}\Big)^i$$

After chirp correction, the IRF can be fitted around the central pump arrival time. This is fitted assuming the Gaussian pump pulse induces a response of exponential decay[186]:

$$IRF = \frac{1}{2}\exp\Big(k\Big(\mu + \frac{k\delta_t}{2} - t\Big)\Big)\left(1 + erf\Big(\frac{t - (\mu + k\delta_t^2)}{\sqrt{2}\delta_t}\Big)\right)$$

Where δ_t is the $1/e^2$ of the IRF, k is the relaxation rate. Normally this is fitted around the signal at time-zero which is referred to as the coherent artefact, coherent spike or cross-phase modulation [187].

In chapter 6, a sequential non-branched model is invoked to globally fit the ESA and GSB evolution of visible-pump IR probe TA. This meant data was fitted with a single connectivity assumption where

one species evolved into another (i.e. A→ B →C → ...) which was appropriate given the unimolecular reaction of chromophore isomerisation being studied.

A sequential model involves the fitting of n components to make an idealised description of the entire dataset as a function of wavelength and delay ($\Psi(\lambda, T)$) [186]:

$$\Psi(\lambda, T) = \sum_{i=0}^{n} A_i(\lambda) \exp\left(\frac{-(t - \tau_{i-1})}{\tau_i}\right) \tag{1.8}$$

Where the A_i is the amplitude and τ_i is the lifetime (reciprocal rate constant) for each evolution-associated difference spectra (EADS). The number of components required to describe a dataset is best chosen by evaluation of singular value decomposition (SVD).

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\dagger}$$

Where $\mathbf{M}$ is the n×m matrix of time dependent data, $\mathbf{\Sigma}$ is diagonalised matrix orthogonal to $\mathbf{U}$ and $\mathbf{V}$, which are the left and right singular vectors on the time and spectral domains of dimensions n×m and m×n respectively. Generally, the number of SVD components is increased until the left and right singular vectors start to show noise and/or loose physically relevant structure.

1.5.4 Multi-Pulse Experiments

In some cases, pump-probe experiments are not definitive in determining the dynamics of a system. This can be due to complicated connectivity schemes which occur when a system has multiple decay pathways. In these cases, additional pulses can be included in the pump-probe scheme to help assign the pathway between states [186, 188]. In this thesis (chapter 6), a pump-dump-probe (PDP) scheme, sometimes referred to as stimulated emission pumping, is used to obtain a kinetic trace which assigns the excited state behaviour of a reversibly switchable fluorescent protein. In PDP TA, an additional dump pulse is introduced at a chosen delay (τ) relative to the pump pulse pulse. The dump wavelength is chosen to induce stimulated emission from the excited state (i.e. dump the population) whilst, ideally, being non-resonant with the ground-state [189]. In this case a double-difference spectrum can be calculated:

$$\Delta\Delta A(\lambda, T, \tau) = \Delta A_{PDP}(\lambda, T, \tau) - \Delta A_{PP}(\lambda, T) \tag{1.9}$$

From the measured changes in absorption of the pump-probe (ΔA_{PP}) and pump-dump-probe (ΔA_{PDP}) respectively. The double-difference absorption represents the response of the system to the dump pulse. It can be globally-fitted equivalently to pump-probe TA and the resulting spectral components interpreted.

1.6 Thesis Outline

In this thesis, reversible switchable fluorescent proteins are studied using TR-SFX, TR-SSX, cryo-MX, UV/Vis spectroscopy, FTIR spectroscopy and IRTA spectroscopy. In chapter 2 existing RSFPs are engineered to improve their steady-state visible spectral behaviour and crystal stability to create a favourable sample for study with ultrafast spectroscopic/crystallographic techniques. Two proteins named rsKiiro and JvT75 are brought forward. Chapter 3, investigates the steady-state structures of the rsKiiro using cryo-MX methods which directly supports the work of chapter 4 where ultrafast TR-SFX measurements using a multi-pulse excitation scheme (pump-probe and pump-dump-probe) investigate the ultrafast response and structural assignment of excited state proteins in ultrafast crystallography. In chapter 5, a 10 ms pump-probe delay is collected using TR-SSX. The effects of X-ray induced heating are then investigated by comparing population transfer measured in the crystal structures with flash-photolysis spectroscopy yields. In chapter 6, ultrafast vibrational spectroscopy is used to determine the early steps in isomerisation events of rsKiiro. Finally, chapter 7 investigates ground state conformations of the protein JvT75 using steady-state FTIR spectroscopy and TR-SSX.

2 | Development of Reversibly Switchable Fluorescent Proteins for Ultrafast Spectroscopy and Crystallography Experiments

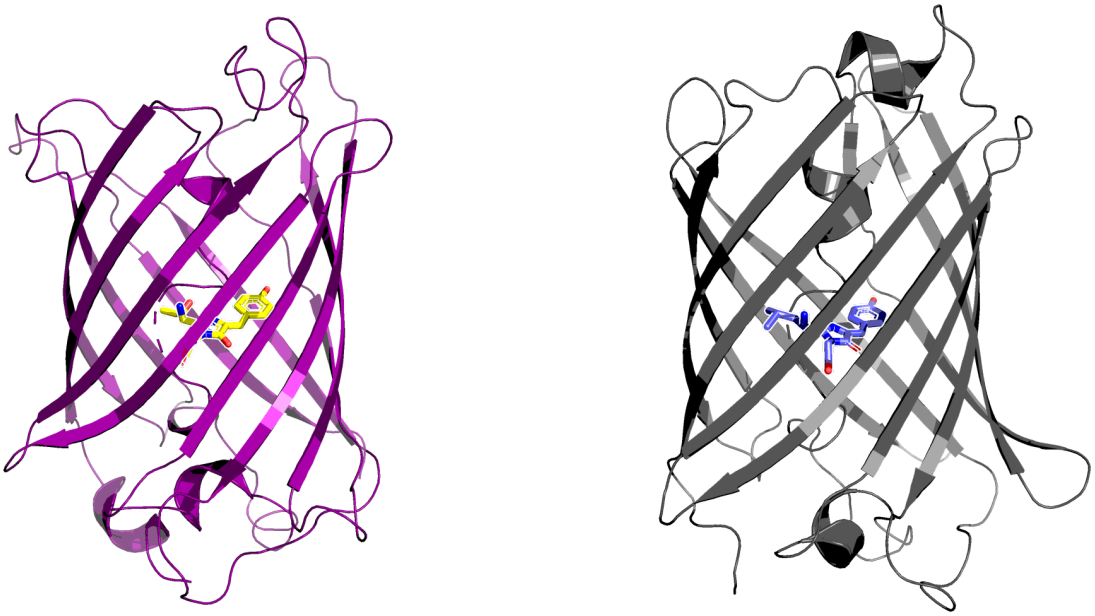
Reversibly switchable fluorescent proteins are a subset of fluorescent proteins which have been engineered to allow reversible photoconversion between a strong and weakly fluorescent state. Targeted mutations to wild-type fluorescent proteins can alter photochemical behaviour and affect a range of characteristics from folding efficiency and crystalline stability to expression yields in E. coli. Mutagenesis is, therefore, a tool that can be used to control the attributes of fluorescent proteins to permit ultrafast crystallography (and spectroscopy) experiments. The aim of this chapter is to introduce the reversibly switchable fluorescent proteins studied in this thesis and show the development of new proteins with suitable characteristics for time-resolved crystallography experiments. A range of mutations was introduced to the fluorescent proteins Skylan-NS [2], Thermal Green Protein [190], rsTagRFP [191] and rsEGFP2 [192] to improve their photoconversion rates, crystallographic resolution and expression yields. Two mutants of Skylan-NS were found to show the highest switching rates in combination with excellent crystallography and were selected as ideal systems for further experiments.

Note to the reader: In this chapter mutants were designed by J. van Thor. Mutagenesis, transformation, expression and purification was done by V. Cordon-Preciado and K. Maghlaoui.

2.1 Introduction

Fluorescent proteins (FPs) were first discovered in the form of the green fluorescent protein (avGFP), from the jellyfish Aequorea Victoria in 1962 [193]. In 1994, avGFP was used as a marker for gene expression [194] and has since been developed as a non-invasive, *in vivo*, reporter and biosensor for

measurements of numerous properties of biological systems [195]. Since its discovery, site-directed mutagenesis of avGFP has led to an array of variants with different absorption and emission spectra [196] as well as enhancement of the fluorescence quantum yields for easier detection in bio-functional expression assays and fluorescent microscopy measurements [197]. Following the isolation of GFP, many other FPs have been discovered. A majority of these can be split into two families: those originating from the taxonomic classes of Hydrozoa (jellyfish) or those emanating from Anthozoa (corals and sea anemone). FPs originating from both families share similar secondary-structure properties, seen in figure 2.1: an 11-stranded β-barrel with a central α-helix. The chromophore is bonded to the alpha-helix and is formed during folding from the oxidation of three amino acids, in a process known as maturation [195]. Most Hydrozoa and Anthozoa proteins are 220-240 residues in length with the chromophore tripeptide motif occurring close to residue numbers 62-65. Figure 2.2 shows the relative distance of residue number to the chromophore for two different proteins, Skylan-NS [2] and avGFP [190]. For the Anthozoa-type proteins studied in this section residues: 58, 61, 65, 66, 72, 89, 91, 105, 142, 173 and 212 all sit within 3 Å of the chromophore. Therefore, mutations to these residues are likely to affect the photochemical switching behaviour.

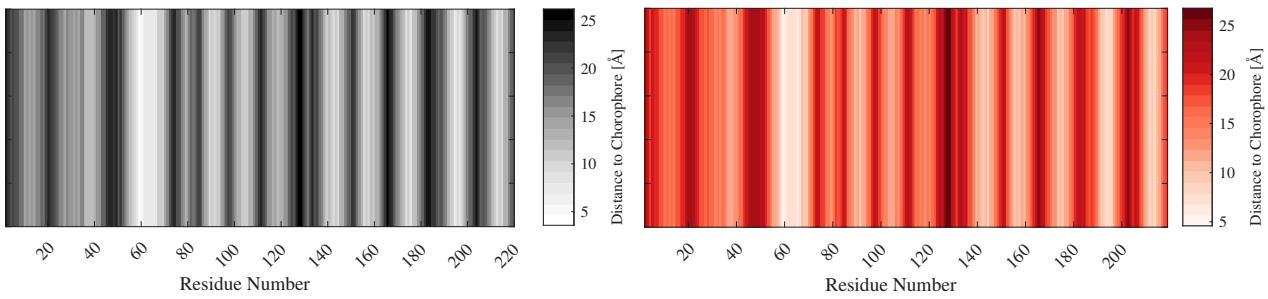


(a) Green fluorescent protein (1EMA [198])

(b) Skylan-NS (5OQE [1])

Figure 2.1: Chromophore (yellow and blue) and secondary structure (purple and grey) of two fluorescent proteins from **(a)** Hydrozoa (GFP) and **(b)** Anthozoa (Skylan-NS) origin. Both chromophores are in the cis conformation. The 11 stranded β barrel structure of GFP is noticeably wider than the Skylan-NS variant but otherwise, the structures are very similar with an α helix holding the chromophore in the centre.

FPs that undergo photochemical reactions are broadly known as photo-transformable fluorescent proteins (PTFPs). PTFPs can be split into classes describing their photo-induced behaviour. Following the definitions set out by Adam [199] we define three main categories of PTFPs: photoactivatable



(a) Skylan-NS (5OQE [1])

(b) TGP (4TZA [190])

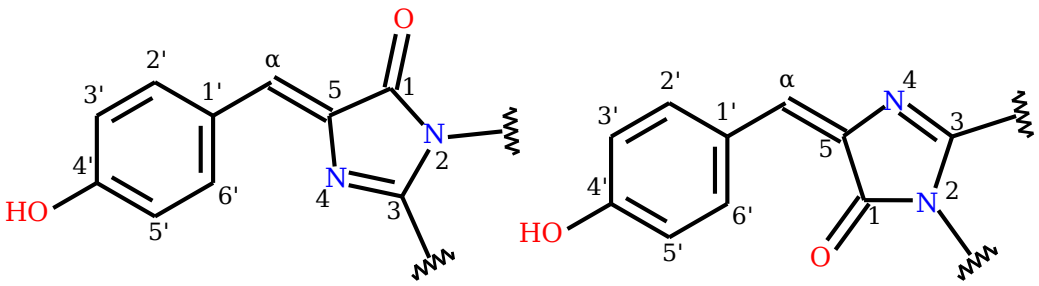
Figure 2.2: Contact map showing distance (shown by the colour bar) of the C_α in each residue (residue number) to the Cα of the chromophore for Skylan-NS and TGP.

fluorescent proteins (PAFPs), photoconvertible fluorescent proteins (PCFP) and reversibly switchable fluorescent proteins (RSFP).

Photoactivatable fluorescent proteins (PAFPs) are those which undergo an irreversible transition from a dark (non-fluorescent/off) state to a fluorescent state. This transition is normally associated with the cleavage of a covalent bond which leads to a change in the protonation state of the chromophore [200]. One of the first examples of a PAFP was photoactivatable-GFP which undergoes a decarboxylate reaction upon excitation which shifts the equilibrium of the chromophore's protonation state to an anionic and highly fluorescent state [201]. Techniques such as photoactivatable localisation microscopy (PALM) [19] rely on the irreversible activation of PAFPs to obtain super-resolution images. In PALM, a small concentration of PAFPs are activated with a weak pump source. Fluorescence imaging of the activated proteins is then carried out until substantial bleaching has occurred. This process is then repeated with numerous switching/bleaching cycles to localise single molecules at a resolution beyond the classical airy disk diffraction limit.

Photoconvertible fluorescent proteins (PCFPs), irreversibly transition from a highly-fluorescent state to a shifted absorption/emission wavelength. The transition normally involves chemical changes to the chromophore. For example in EosFP [202], a green-red PCFP, 400 nm excitation leads to autocatalytic dehydration of the chromophore backbone. This extends the chromophore's conjugated π-electron system which leads to an irreversible redshift of around 75 nm of the absorption/emission peaks [202, 203]. PCFPs have found use in tracking applications. Photoconversion can be localised to one area of a sample *in vivo* and images acquired to track the motion/movement through a sample over time [204].

Reversibly switchable fluorescent proteins (RSFPs) are a broad class of PTFPs where the fluorescence can reversibly be switched on or off through a photochromatic reaction often involving the cis/trans isomerisation and protonation state change of the chromophore. In both Anthozoa and Hydrozoa type



(a) Cis Neutral **(b)** Trans Neutral

Figure 2.3: Skeletal Structure of the HBDI chromophore in the cis and trans neutral states.

RSFPs the chromophore is formed from an autocatalytic condensation reaction during folding which creates the 4-hydroxybenzylidene-1,2 dimethylimidazolinone (HBDI) chromophore [195, 205], seen in figure 2.3. The HBDI chromophore undergoes isomerisation around the carbonyl methyl bridge (α carbon) when switching between the cis and trans isomers. There are two subdivisions of RSFP: most common are those considered negative-switchers which are most stable in the fluorescent/on-state. Positive-switchers show inverted behaviour and are most stable in their off/dark-state. Dronpa [206] was the first discovered RSFP. It was engineered from a wild-type protein found in the *Echinophyllia* sp. SC22 (Stony Coral). Dronpa is a negative-switcher meaning it is most stable in the fluorescent state and switches off under illumination (with 500 nm light). The off-state strongly absorbs at 400 nm, 100 nm blue shift compared to the maximum of the on-state. To date, Dronpa has been utilised in numerous cell-biology and microscopy experiments due to its high fluorescence yields and monomeric nature. RSFPs are not limited to emission in the green. Since the discovery of Dronpa, a colour palette of RSFPs have been identified with emission wavelengths from 500-610 nm with excitation between 460-580 nm [207].

Of the 811 entries in the fluorescent protein database (fpbase [207]) only 32 exhibit RSFP behaviour. The main use of RSFPs is in super-resolution microscopy techniques. The ability to switch fluorescent states on and off has led to the development of techniques such as reversible saturable optical fluorescence transition (RESOLFT) [208] and non-linear structured illumination microscopy (NL-SIM) [209] which allow localisation of features beyond the classical diffraction limit. The mass-adoption of these microscopy techniques has accelerated the development of new and improved RSFPs. A majority of super-resolution microscopy techniques rely on stable photoswitching behaviour in terms of predictable bleach rates, high contrast-ratios (ratio on to off fluorescence emission intensities) and high fluorescent yields to maximise the accuracy of localisation. Most RSFPs have therefore been developed with these criteria in mind [210].

To study RSFP switching mechanisms with TR-SFX on ultrafast timescales, photoproduct yields need to be large enough for single femtosecond laser pulses to generate observable light-induced structure

factor differences in photoactive crystals. As mutagenesis of amino acids can drastically alter the behaviour of a protein it can be used as a tool to generate a target with the specific required properties for TR-SFX measurements. In particular, existing sequence information from existing FP/RSFPs can be used to optimise crystallography, photoproduct yields an expression yields. This section follows the development of new proteins from several known FPs. The aim is to engineer a sample with the following properties:

- A high primary quantum yield of switching: previous TR-SFX experiments on the protein Skylan-NS [1] have suggested that a primary quantum yield of at least 10% is necessary to produce observable changes [211].
- High expression efficiency and stability: sample delivery methods employed in SFX experiments (section 1.3.2) can use around 0.3 grams of protein per a structure [27]. This can be extremely resource-intensive to produce. Hence, high expression yields in *E. coli* are necessary to produce enough material without a prohibitive amount of purification time for SFX measurements. In addition, the protein must be stable enough to be stored for long periods and survive shipping literally across the globe (in our case from London to Japan, California, Korea, Germany or Oxfordshire).
- High quality and stable crystals with high symmetry: crystals must stable and of suitable dimensions for injection (µm sized), diffract to high-resolution (ideally better than 2 Å) and be durable enough to withstand the forces exerted during injection. Furthermore, higher symmetry space groups allow data of higher multiplicity to be collected with less indexed frames which essentially allows structures to be collected to high completeness using less beamtime.
- High state homogeneity: some RSFPs such as rsEGFP2 [192] have been shown to not fully convert with illumination [212]. This structural heterogeneity can make crystallographic and spectroscopic analysis much more complex so is best avoided.

There are many potential target proteins for TR-SFX measurements, the most widely studied is Dronpa. Extensive spectroscopic and theoretical studies have confirmed the spectroscopic contrast is caused by the isomerisation of the chromophore from a deprotonated cis conformation to a neutral trans [213–215]. Recent ultrafast spectroscopic measurements have suggested the deprotonation occurs as a ground state process [5, 7] as opposed to the previously reported excited-state proton transfer [216] which is observed in GFP [217]. Crystals of the fast-switching mutant of Dronpa [206] have been found to lose diffraction quality under illumination meaning they are not suitable for TR-SFX experiments [218, 219]. In contrast, mutants of EosFP [202] another Anthozoa FP isolated from *Lobophyllia hemprichii* have

been shown to photoswitch in crystal form. Mutants of EosFP, in particular, Skylan-NS [2], show good potential for the development of a sample for TR-SFX experiments.

Similarly to Dronpa, in its unilluminated state, on-state, Skylan-NS shows bright fluorescence and strong absorption (133,770 /M/cm) at 501 nm assigned to an anionic cis chromophore conformation [2]. Illumination with 500 nm light drives photoconversion to a low-fluorescence state which most likely corresponds to a meta-stable neutral-trans state with an absorption peak at 387 nm and a lifetime of around 4 hours. The structure of both the cis and trans states of Skylan-NS have been obtained using SFX to a resolution of 1.6 Å [211]. However, preliminary experiments suggested the photoproduct yield of Skylan-NS is too low to detect light induced differences in ultrafast TR-SFX experiments. This has led to the development of an array of new mutants as detailed in this chapter. Comprehensive mutagenesis analysis of GFP and GFP-like variants has been done by several authors [199, 220]. It is not the aim of this chapter to comprehensively investigate the effect of each mutation made, the aim and results of this chapter are to improve the photo-product yields and crystallography of Skylan-NS and other similar proteins for use in ultrafast crystallography and spectroscopy experiments.

2.2 Methods

Table 2.1 shows the mutants developed since October 2017. Gene synthesis of these constructs and cloning into the corresponding expression vector was carried out by GeneCust, transformed and expressed in *E.coli* strain BL21(DE3) using the pET28a(+) plasmid (unless stated otherwise in table 2.1). Mutants were derived from Skylan-NS [221], thermal green protein (TGP) [190], enhanced green fluorescent protein 2 (rsGFP2) [192], and a red fluorescent protein (rsTagRFP) [222]. During Nickel-affinity purification spectra of fractions were collected to determine if expression was successful.

To establish the suitability of a mutant for TR-SFX experiments the samples were characterised in three ways: steady-state kinetics, fs-flash photolysis (performed by C. Hutchison) and crystallography trials.

UV/Vis steady-state spectra were collected as explained in detail in section 2.2.1, 1-2 µl of concentrated protein solution was mixed to 100 µl of buffer to give appropriate optical density (OD) between 0.1-0.3. Data was collected in three buffers at pH 10 (0.1 M HEPES, 0.05 M NaCl), pH 8.4 (0.1 M Tris-Cl, 0.05 M NaCl) and pH 5 (0.1 M MES, 0.05 M NaCl). A defocused 405 nm laser and a 450 nm LED were attenuated to a power density of around 1 mW/cm^2. As exact power densities were difficult to determine Skylan-NS was used as an actinometrical calibration for each set of measurements to calculate the relative rate of switching.

The pK_a was measured by preparing 'low salt' buffers as mentioned above (HEPES, Tris-Cl and MES)

in pH steps of 0.5 between pH 4-10. Around 1 µl of concentrated protein was mixed with 100 µl of buffer at the selected pH and stationary spectra, as described in section 2.2.1, were acquired. The Henderson Hasslebalch equation was then used to calculate the pK_a:

$$pH = pK_a + \log_{10}\left(\frac{[\text{Off}]}{[\text{On}]}\right) \tag{2.1}$$

Where [Off] and [On] are the normalised concentrations of the off and on-states. In this case, the concentrations were normalised with the aromatic peak at 280 nm, and then the maximum peak height of the on/off state.

The thermal recovery rate was measured at room temperature over 12 hours. A protein sample was prepared at an optical density between 0.2-0.6 and converted to the off state using 500 nm light. Spectra were then taken every minute for 12 hours to track the recovery back to the on-state. A spectrum was taken before conversion (with the protein in the unilluminated state) to normalise the re-conversion too. Parafilm was used to seal the cuvette to minimise sample evaporation.

Crystallisation conditions were screened using a PEG-ion screen with a TTPLabTech Mosquito crystallisation robot, as explained in section 2.2.2. Sitting drop trials (section 2.2.3) aimed to produce large (mm) single-crystals for cryo-MX structure determination whilst batch (section 2.2.4) conditions were trialled to make crystal slurries at high enough crystal densities for serial crystallography and spectroscopic 'pancake' measurements.

2.2.1 Ultraviolet / Visible Spectroscopy

An Agilent 8453 UV/Vis spectrometer was used to acquire spectra in the range of 200-800 nm at a resolution of 1 nm. Background signal was subtracted from spectra taken with the buffer only in a 100 µl, 1 cm path length cuvette (Hellma Quartz Suprasil Micro Absorption Cell 1.000). When collecting kinetic data the sample was optimised to have a maximum absorption between 0.2-0.8. The spectrometer acquired at a rate of 2 Hz with 0.5 for each exposure. A continuous-wave laser or LED was used for illumination. Lasers were defocused and a mirror used to direct the at the sample. The LED was mounted directly above the cuvette. First-order rate equations fitted the change in optical density over time using MATLAB.

2.2.2 Mosquito Trials

Commercially available 96-well PEG/Ion screens (HR2-139 Hampton) were set up using a TTPLabTech Mosquito crystallisation robot. Each well contained two droplets of different protein concentration. For both positions, droplet volume remained the same but two different protein concentrations were used. The robot was programmed to dispense a 2 nl drop of protein which was mixed with a 2 nl of buffer solution. The PEG/ion screen varied in buffer composition through the different wells to find a

stabilising condition to grow crystals. Once prepared the 96 well plates were stored in a 25°C incubator and checked using a bifocal microscope every 3 days for crystal growth.

2.2.3 Sitting Drop crystallisation: 24-Well Plate Trails

24-well plates (Hampton Research) were prepared with micro-bridges placed at the centre of each well. A miniature coverslip was adhered using vacuum grease to a miniature cover slide for sitting drop crystallisation. A 2 µl drop of protein solution was placed on the coverslip and mixed with an equal volume of crystallisation buffer. A further 1 ml of crystallisation buffer was placed in the well below. The buffer composition was valued linearly over one axis of the plate, and protein over the other axis. The plates were then incubated for several weeks at 21°C and checked periodically for crystal growth.

2.2.4 Batch Crystallisation

1 ml round-bottomed Eppendorf tubes were prepared with a final volume of 100 µl of crystallisation solution. A 40 % solution of poly-ethylene glycol (PEG) in 0.2 M lithium sulfate, 0.1 M Tris-HCl at pH 8.5 was mixed with a low salt buffer (0.2 M lithium sulfate, 0.1 M Tris-HCl) and a small volume of concentrated protein solution (typically above 50 mg/ml) in 0.1 M Tris-HCl. Once added, the protein solution was well mixed with the crystallisation buffer using a pipette. The ratio of protein solution to crystallisation buffer was varied and a low salt buffer used to maintain a constant final volume. If available from other crystallisation trials, 1 µl volume of crystal seed solution was added.

2.3 Results

A summary of mutations made and their names is shown in table 2.1. Although many of the mutants were successfully expressed, JvT67 JvT70, JvT73 and JvT76 showed either very low levels FP concentration (below 10 mg/litre expression yields) or failed to produce any products with the expected FP spectra or colour. This is possibly due to misfolding during expression or inhibited chromophore maturation [223]. Figure 2.2 shows the distance of each residue position to the chromophore in Skylan-NS and GFP. A majority of mutations made in table 2.1 are correlated with positions that are close to the chromophore. Mutants JvT71, JvT74, JvT75, JvT77, JvT81 and JvT83 were all successfully expressed and purified. Stationary spectra of the off and on-states were collected for each construct at three difference pHs, shown in figure 2.5. A summary of peak maximum and absorption coefficients for each construct is shown in Table 2.2.

Overall, 96-well crystallisation trials initially appeared to be successful for the TGP based mutants which yield crystals in numerous wells. However, these crystals did not diffract to high quality, as summarised in table 2.3 despite the visual appearance, as seen in figure 2.4a. 24-well trials were tested JvT71, JvT74 and JvT75, but no crystals were successfully grown. For Skylan-NS mutants, JvT75 and

JvT81 batch crystallisation conditions gave needle-shaped crystals (seen in figure 2.4b), the average size of which could be controlled from 10-100 µm in length and around 1-20 µm in width by with changes in PEG and protein concentration buffer composition. As a general rule, it was found higher PEG concentrations gave smaller crystals, whilst higher protein concentrations gave larger crystals (within limits).

Mutant Name	Parent	Mutations	Notes
JVT67	TGP	P59T A60V Q62A T69A F173S	Low Expression
JVT68	rsEGFP2	V151A	
JVT69	rsTagRFP	M63A	
JVT70	Skylan-NS	L62A R66L H194A	No Expression
JVT71	Skylan-NS	L62A F173S	
JVT72	rsTagRFP	M63A L194V	
JVT73	rsTagRFP	R67L	No Expression
JVT74	Skylan-NS	A60V L62A L93M N116Y V118I M146L	
JVT75	Skylan-NS	L62A S142A M159T	
JVT76	rsEGFP2	V151A (in pQE31 plasmid)	No Expression
JVT77	TGP	T69A (in pQE31 plasmid)	
JVT81 (rsKiiro)	Skylan-NS	L62A M159T	
JVT83	TGP	T69A (with new C-terminus)	

Table 2.1: Specific point mutations of RSFPs engineered by the van Thor group since October 2018.

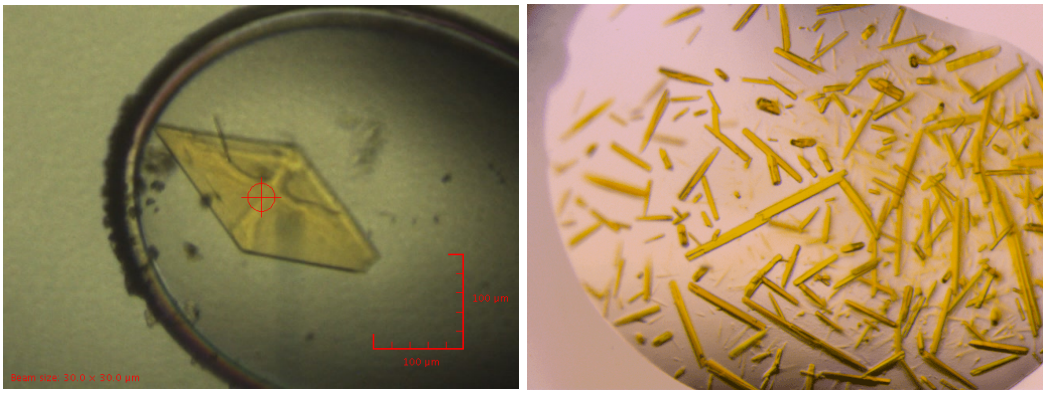
JvT71 shows a blue-shifted on-state with a decreased molar absorption coefficient. Whilst the off state is red-shifted and show the lowest molar absorption coefficient of the proteins successfully purified. The F173S mutation is a major change, from a sterically large and hydrophobic amino acid to a small polar one. Looking at table 2.2 ensemble switching rates of JvT71 are notably higher than related protein, Skylan-NS. Looking at figure 2.5, JvT71 is also the only construct to show no spectral differences at low pH. The peak position blue shifts at higher pH's which is perhaps a sign of a second protonation site around the chromophore. Crystal trials of JvT71 were unsuccessful with no diffracting crystals found.

JVT74 shows the highest absorption coefficient compared to other mutants at pH 8.4. This can be seen in figure 2.5 where the on-state absorption peak is around 2.25 times stronger than the aromatic peak at 280 nm. Ensemble kinetics of JvT74 were much slower than those for Skylan-NS, and, as seen in figure 2.5. It seems the on-state does not fully convert to the off state under 500 nm illumination.

Construct	On-state		Off State		Normalised Rate* (pH 8.4)	
	Max. [nm]	Molar Absorptivity [$M^{-1}cm^{-1}$]	Max. [nm]	Molar Absorptivity [$M^{-1}cm^{-1}$]	On to Off	Off to On
JvT71	487	79290	394	24787	1.55	1.16
JvT74	503	232476	392	48002	0.46	0.52
JvT75	503	76971	394	60500	2.61	1.45
JvT77	504	174246	399	45215	0.54	1.22
JvT81	489	117333	391	45280	6.66	1.49
JvT83	504	181415	399	42059	0.45	0.67
Skylan-NS[2]	501	133770	387	34595	1.00	1.00
TGP [190]	493	64000	-	-	-	-

Table 2.2: Absorption maximum and molar absorption coefficients of the off and on states of mutants of TGP and Skylan-NS (for reference) at pH 8.4. [2]. * The normalised rate is the relative rate of switching as compared to Skylan-NS, normalised for the absorption coefficient of each state: $\text{rate}_{\text{rel. norm.}} = \text{rate}_{\text{rel.}} \frac{\epsilon_{\text{Construct}}}{\epsilon_{\text{Skylan-NS}}}$

Crystal trials of this mutant were also unsuccessful with no diffracting crystals produced.



(a) TGP-based JVT83 crystals **(b)** JvT81 (rsKiiro) crystals

Figure 2.4: (a): JvT83 crystals in cryoloop for MX data collection. **(b):** Batch grown crystals of JvT81 (rsKiiro) around 100 µm long (right).

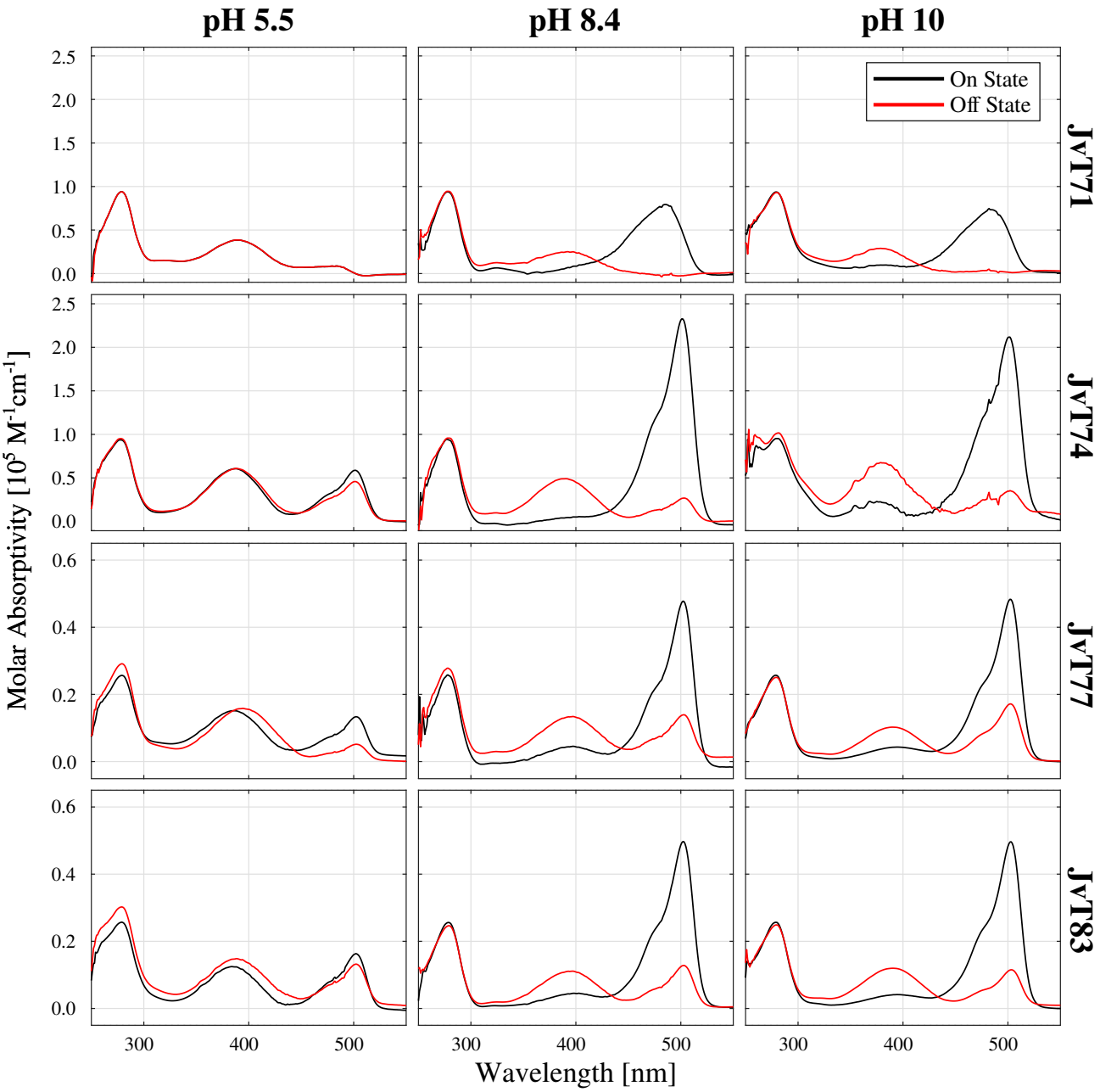


Figure 2.5: Molar absorption coefficient between 250-550 nm at pH 5.5 (left column), 8.4 (middle column) and 10 (right column) for mutants JvT71 (top row), 74 (second from top row), 77 (second from bottom row), and 83 (bottom row), as labelled on the right side and with mutations defined in table 2.1. Shown in black is the unilluminated/dark/ground state of the protein which corresponds to the fluorescent, on-state for all constructs at pH 8.4. The red line, labelled off state shows protein spectra after illumination with 500 nm light.

JvT75 shows the second-quickest rate of photoconversion in the trans to cis (off to on direction). It also shows the highest absorption coefficient in the off state of any mutant. The low pH switching behaviour, seen in 2.7 is interesting to note. At pH 8.4, it was also the only protein not to fully photo-reconvert after switching to the off state and had the largest off state peak whilst unilluminated. Instead, the reaction of the off to on (driven with 400 nm light) only returned to 40% of the original peak height. The measurement of thermal re-conversion (seen in figure 2.6) followed first-order rate equations only

at pH 10 (black squares) but not at pH 8.4 (blue diamonds). The rate constant at pH 10 was fitted to be 2.5×10^{-4} s^{-1}. These features are all indications of more complex photophysical behaviour which is investigated later in this thesis (chapter 7). Fitting the pK_a gave a value of 8.2 (figure 2.7). The fluorescence spectra, shown in 2.6, has a maximum of 512 nm, showing a Stokes shift of 9 nm. Overall, the crystallography of JvT75 was very successful with many high-resolution crystals produced via the batch-crystallisation method. In particular, the condition of 7-10 mg/ml protein concentrations at 18.75 % PEG 3350 concentration consistently diffracted to better than 2 Å resolution. Photoswitching was observed in crystals from colour change under 400/500 nm illumination and was clear in the phased electron density and refined coordinates of refined structures. Table A.1 shows the crystallographic statistics of the collected cis and trans states. Both are limited to a resolution of 1.24 Å and refinement and merging statistics are appropriate for data of this quality.

JvT77 showed efficient switching in the off to on but slow switching in the on to off direction. The on-state peak shows the largest redshift of any mutant studied here. Full conversion was not seen under 500 nm illumination at pH 8.4, but (similar to JvT75 at pH 5) there are some spectral shifts seen after illumination in the low pH measurements. Attempts to crystallise JvT77 gave perfect-looking octagon shaped crystals which unfortunately showed very poor diffraction quality, despite attempts to optimise the growth conditions.

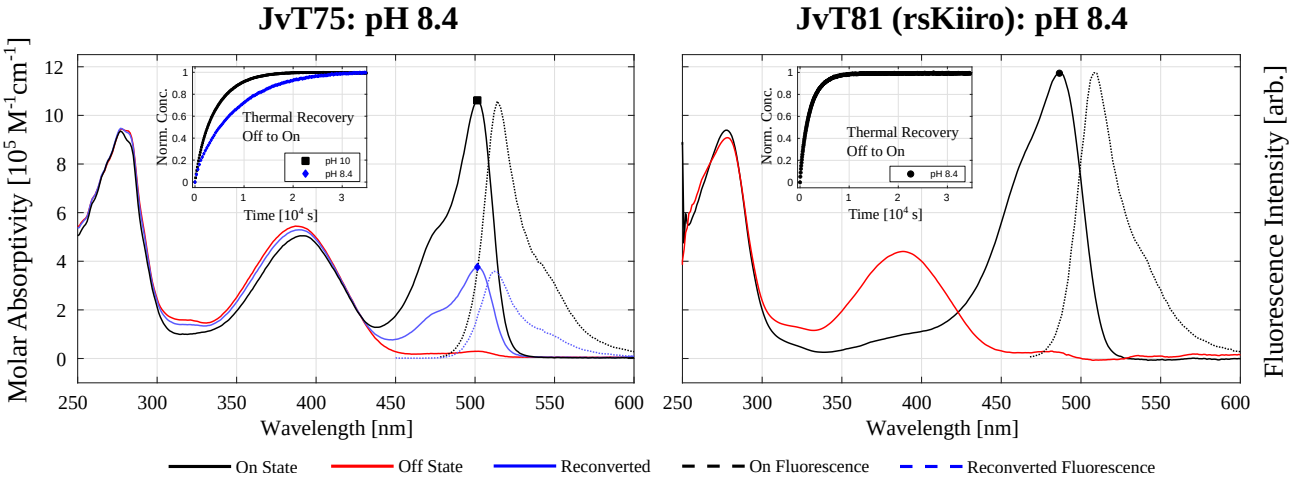


Figure 2.6: Molar absorption coefficient between 250-600 nm of JvT75 (left) and JvT81 (right) at pH 8.4 of the fluorescent on-state (black) and photoconverted (with 500 nm light) off state (red). The on-state fluorescence is plotted in dashed black line. Shown in blue is the only partial re-conversion seen after illumination with 500 nm followed by 400 nm. The inset plots show the thermal recovery of the on-state peak from the off state for pH 10 & 8.4 (JvT75) and pH 8.4 for JvT81.

JvT81 showed the highest rate of switching from the on to the off state. The absorption coefficients of the red and blue-shifted on and off-state peaks absorb -10% and +30% less/more than Skylan-NS. This mutant (along with JvT71) was one of the only two samples to switch fully between the off and on

states with very little bleaching observed after a switching cycle. The pK_a was 5.4 as seen in figure 2.7, which is a large jump compared to the JvT75 mutant which only differs by the S142A mutation. The fluorescence emission spectra of the on-state peaked at 508 nm with excitation from 420-500 nm, slightly blue-shifted compared to JvT75 but with a much larger Stokes shift, of 19 nm. The thermal recovery rate of the off to on-state was measured and found to follow a first-order rate equation with a rate constant of 5.0×10^{-4} s^{-1} (see figure 2.6). The crystallisation of the JvT81 mutant was highly successful with crystals formed in batch and sitting drop trials. Batch crystals were formed at protein concentrations of 8-17 mg/ml and PEG 3350 concentrations of 15-28 %. Sitting drop trials were successful in two conditions at 10 mg/ml protein concentration and 22-23 % PEG 3350 concentration using a 2:2 µl mother-liquor to protein mix for the sitting drop. After optimisation, the highest resolution structure of JvT81 was captured to 0.94 Å using MX, placing it the 8^{th} highest resolution structure of a fluorescent protein as of early 2021 (based on all FP PDB submissions). Crystals clearly changed colour from yellow to transparent when illuminated with 500 nm light, indicating photoswitching had occurred. MX data collection of both the on and off states confirmed this with 100% conversion seen in fitted model coordinates as seen in figure 2.8. The structures and statistics for JvT81 (rsKiiro) are presented in Chapter 3 of this thesis.

JvT83 showed an enhanced absorption coefficient almost 3 times higher than its parent construct, TGP, but showed low switching rates in both directions. The switching did not continue to completion, with around 25% of the on-state peak remaining after continued illumination with 500 nm. Interesting, at low pH this mutant seemed to show a denaturing effect. Illumination with 400/500 nm light increased the background scatter in the measurement significantly. Perhaps the T69A, E45A E73A mutations allowed too much solvent interaction with the secondary structure elements which denatured the protein at low pHs.

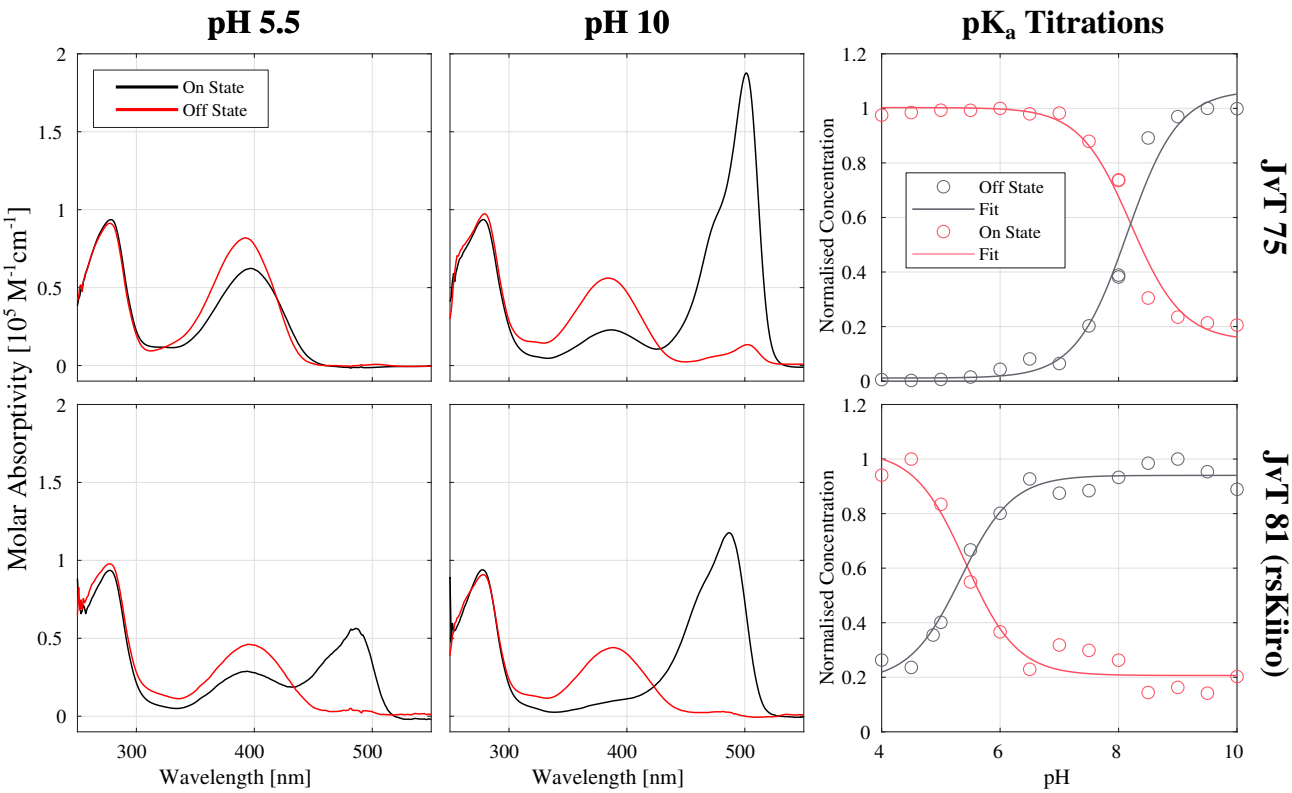
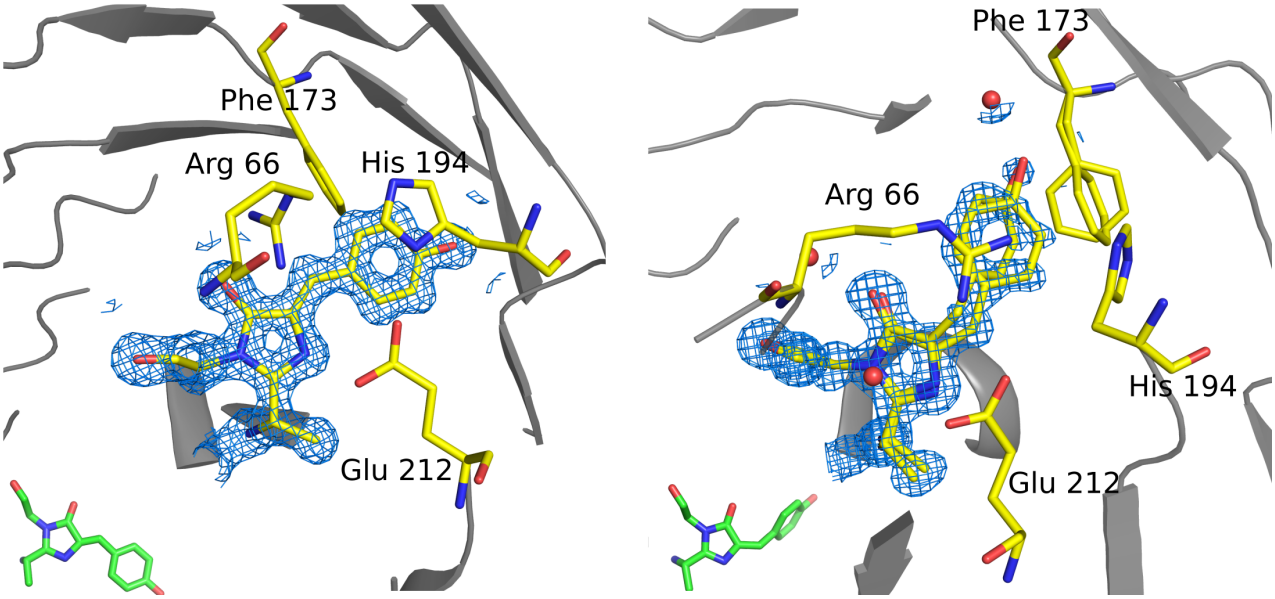


Figure 2.7: Molar absorption coefficient between 250-550 nm at pH 5.4 (left column) and 10 (middle) of mutants JvT75 (top) and JvT81 (bottom). On the right is the normalised concentration of the off and on peaks in red and black circles respectively plotted as a function of pH. The solid lines follow the Henderson–Hasselbalch equation for fitting the pK_a for both mutants. Fits converged to a value of 8.17 and 5.52 for JvT75 and JvT81 respectively.

		Number of Full Datasets at:	
Construct	No. Fished Crystals	Low Res. (2-3 Å)	High Res. (<2 Å)
JVT75	93	8	16
JVT77	12	1	0
JVT81 (rsKiiro)	28	0	13
JVT83	16	6	3
Total:	153	15	32

Table 2.3: Summary of collected crystals and diffraction quality.



(a) On-state (cis)

(b) Off-state (trans)

Figure 2.8: $2F_o$ - F_c electron density contoured at 2 σ (blue) of JvT81 for the fluorescent on-state in the cis conformation (left) and off-state, trans conformation (right). In yellow are the coordinates of the chromophore and surrounding amino acids which rearrange with photoisomerisation. Grey shows the secondary structure. In the bottom corner left corner an isolated sketch of the chromophore shown for clarity.

2.4 Discussion

Overall, eight mutants were successfully expressed and spectral properties determined. In general, substitution of larger, hydrophobic residues for smaller polar amino acids in the chromophore region encouraged quicker switching rates. However, this had to be balanced with ensuring close enough contact was maintained between stabilising residues and the chromophore. Spectral shifts are seen across the range of mutants studied here. The largest blue shift of the on-state is seen in the JvT71 mutant to 487 nm and the largest redshift to 504 nm. Alongside these spectral shifts, there are large changes in the oscillator strength of the visible absorption as evidenced by the large variation in absorption coefficient. Given the nature of the mutations, it is likely that these shifts are caused by changes to the electrostatic environment of the chromophore leading to a twisted/non-co-planar of the phenol and imidazolinone rings [224, 225].

The first residue in the triple-peptide which forms the chromophore on maturation varies between different FPs. In Skylan-NS a Lys-Tyr-Gly amino acid sequence forms the chromophore [2] whilst in TGP the chromophore is formed from Gln-Tyr-Gly residues [226]. Studies of reversibly switchable GFP (rsGFP) [192] and mEos2 [226] have found that mutation of the first residue in the chromophore tripeptide sequence to alanine leads to an increased switching rate. The equivalent mutation in Skylan-NS is L62A which was found in JvT71, JvT75 and JvT81. These mutants all showed higher

switching rates than Skylan-NS, in agreement with the studies on rsEGFP and mEos2. The equivalent substitution in TGP, T69A, however, did not produce the same effects as the rate of switching did not increase.

The F173S mutation in mutant JvT71 has been postulated to increase switching rates as it may reduce the steric pressure on the chromophore during isomerisation [227]. Although this mutant showed high switching rates they were not as high as other Skylan-NS mutants. Possibly, the introduction of the polar serine disrupts hydrogen bonding around the alpha helix. Therefore, this evens out the reduction in steric pressure, which prevents the switching rates from increasing by a large amount.

The M159T mutation was found in the most efficient switchers, JvT81 and JvT75, which both originated from the parent protein Skylan-NS. In another Anthozoa based protein, Dronpa, the M159T mutation was found to increase the on to off switching rates by a factor of 1000 [228, 229]. While the effect is not so dramatic here, the same mutation certainly increases the rate of switching. The crystal structure of JvT81 shows that the S142 residue donates a hydrogen bond to the cis chromophore (seen in figure 2.9) which is removed in the JvT75 mutation (S142A). This single mutation changes the behaviour of the chromophore greatly with a shift of 3 in pK_a. It also leads to a pH dependent equilibrium seen of the reconversion kinetics; at pH 10 the protein fully reconverts after 400 nm illumination and follows first-order rate kinetics in the thermal recovery, as JvT81 does. At pH 8.4 however, only 30% of the original on-state peak is recovered under illumination although this peak does thermally recover fully after around 4 hours at room temperature. Furthermore, at pH 8.4 the recovery rate is characteristic of a two-step process indicating pH plays a role in shifting the equilibrium of the off to on reaction. This behaviour is investigated thoroughly in Chapter 7.

2.5 Conclusions

Overall, this chapter has presented several mutants of TGP and Skylan-NS which all undergo similar photochemical reactions as evidenced by spectroscopic and crystallographic behaviour. The fastest switching mutant, with the best crystallographic characteristics and high expression yields, was mutant JvT81, an L62A, M159T mutant of Skylan-NS. In addition to the high rates measured with steady-state spectroscopy, flash photolysis yield measurements by C. Hutchison using a 100 fs 400 nm pulse showed that JvT81 had the highest photoproduct yield (whilst maintaining a low bleach rate) of all the tested mutants. High-resolution-diffracting and stable microcrystals could be produced in a range of sizes which could support several sample delivery methods for TR-SFX and TR-SSX measurements. As a result, the JvT81 mutant was taken forward for ultrafast crystallographic experiments. In the following chapters, it is referred to as rsKiiro. Another mutant, JvT75, an L62A, M159T, S142A mutant of Skylan-NS also showed desired crystallography and behaviour although the photoswitching behaviour

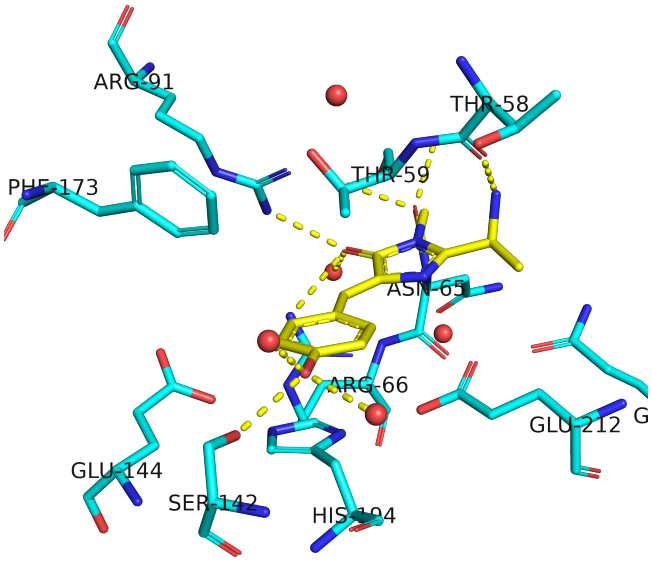


Figure 2.9: Interactions of the cis chromophore (shown in yellow) with surrounding residues (blue) and waters (red spheres) in JvT81 (rsKiiro).

was atypical with an unexplained equilibrium shift upon photo-reconversion. As a result, both of these proteins were taken forward as systems to study the dynamics of photoisomerisation with the crystallographic and spectroscopic experiments presented in this thesis.

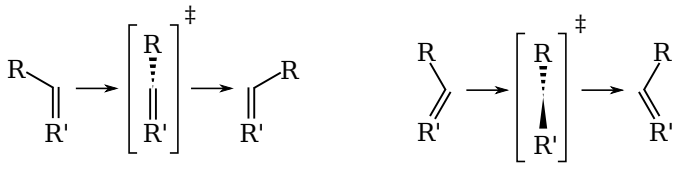
3 | Cryo-Kinetics and Cryo-Crystallography of rsKiiro

The reversibly switchable fluorescent protein rsKiiro was developed in chapter 2 for ultrafast spectroscopy and crystallography experiments. In this chapter cryo-spectroscopy studies of rsKiiro are undertaken to explore the energetic landscape of the off/on and on/off switching mechanism. Spectroscopy reveals convex Arrhenius behaviour with 'breaks' at 229, 208 and 236 K for the on-to-off, off-to-on (preilluminated at low temperature) and the off-to-on (preilluminated at room temperature) conditions respectively. Using cryo-macromolecular crystallography the off and on structures are characterised to atomic-resolution and electron density differences between the conditions is calculated. Cryo-trapped intermediate crystal structures were accumulated with illumination at 150 K with 400 or 500 nm light. Q-weighted difference density maps were then used to calculate extrapolated electron density which resolved a partially occupied (and cryo-trapped) intermediate conformation of the protein. Analysis of this partially occupied intermediate reveals a distinct planar-cis conformation which is shown to relax to the cis state after annealing for 3 seconds to room temperature.

Note to the reader: Cryospectroscopy measurements of protein samples were carried about by S. Bodet Lefèvre under the supervision of this author and C. Hutchison. Analysis of the thermodynamic behaviour was done by J. van Thor and C. Hutchison with input from this author.

3.1 Introduction

Switching of reversibly switchable fluorescent proteins (RSFPs) involves the rotation around the double bond which can proceed via a one-bond-flip (OBF) or hula-twist (HT) mechanism (figure 3.1). The HT mechanism is considered volume-conserving as the rotation sweeps out a smaller space in comparison to the OBF. The OBF mechanism is therefore more likely to occur only when there is no steric hindrance to constrain the larger motion - i.e. in a low-viscosity medium. In RSFPs the chromophore is situated in a tightly-packed environment which is believed to promote the HT over the OBF mechanism [3, 230]. In two recent studies on rsEGFP2 (Chang *et al.* [231]) and Dronpa (Romei *et al.* [224]),



(a) One bond flip **(b)** Hula twist

Figure 3.1: The two isomerisation pathways possible around a carbene double bond: the one bond flip and hula twist mechanisms [231].

a mono-halogen group was attached to the chromophore phenol ring. This broke the symmetry of the phenol-ring motif in the chromophore meaning it was possible to tell if the reaction proceeded via the OBF or the HT mechanism. Depending on the mechanism the halogen group would change side in the photoswitched crystal structure. Remarkably, evidence for both pathways was observed, though the yields were dependant on the crystal packing arrangement [231]. However, it is not clear if the presence of the additional electrophilic group could alter the isomerisation pathway. Furthermore, these studies gave no insight into the exact intermediate structures and sequence of steps involving the protonation/deprotonation of the chromophore *and* the rearrangement of residues Arg66, Phe173 and His194 and Glu 212.

The FP Dronpa has been the subject of a number of electronic and vibrational spectroscopy studies to resolve the switching mechanism. Generally, it is agreed the protonation change for both the cis/trans and trans/cis occurs as a ground state process [6–8, 212, 232, 233]. This is primarily supported by IRTA measurements, due to the observation of a downshifted C=O stretch frequency occurring between 2 ns - 1 ms [6–8, 233]. The C=O stretch originates from the imidazoline ring and acts as a reporter group for the protonation state of the chromophore (chapters 6 and 7). Visible transient absorption measurements also show evidence of the ground state protonation changes with bleaching/formation of the 500 nm anionic absorption band on µs timescales for both the forward and reverse reactions [233]. While the cis/trans reaction is relatively poorly studied due to low photoswitching yields, the trans/cis mechanism has been the subject of a number of studies and the mechanism is highly debated. In particular, there is disagreement about the order in which the structural rearrangements happen. Warren *et al.* suggested the isomerisation occurs as an excited state process in the first 9 ps based on shifts in the reporter C=O and methyl bridge C=C stretching frequencies of the chromophore [5]. This assignment relied on a product peak at 1627 cm^{-1} being slightly up-shifted from the ground state C=C stretch at 1623 cm^{-1} [5]. Kaucikas et al. [8] extended the study by Warren *et al.* to a broader spectral window and made density functional theory (DFT) calculations of a twisted chromophore geometry. The calculations showed that a twisted geometry would show an up-shifted C=C frequency and so supported the excited state isomerisation mechanism proposed by Warren *et al.*

Meech and co-workers have suggested an alternative mechanism where isomerisation is a ground state process [6, 7]. This was initially suggested in Lukacs *et al.* [6] where excited-state absorption features (also seen by Warren *et al.* and Kaucikas *et al.*) were assigned to the reorganisation of the hydrogen bonding network instead of isomerisation. They put forward that this primes the chromophore for full isomerisation at a later point. Subsequent work by Laptenok *et al.* [7] extended the probe delays of Lukacs *et al.* to 200 µs and concluded the chromophore conformation is ill-defined within the first two decay components (16 and 596 ps) assign a component with a 91 ns rise-time to the formation of the cis neutral state which is that protonated on µs timescales.

Overall, further visible spectroscopy studies by Yadav *et al.* [233], theoretical studies [234] and recently TR-SFX experiments [212, 235, 236] have supported the conclusions of Warren *et al.* [5] and Kaucikas *et al.* [8]. Isomerisation in the excited-state is, therefore, the most widely accepted mechanism, but is still disputed.

The switching mechanism and characteristics of rsKiiro are investigated throughout this thesis. In this chapter, we aim to resolve structures in ground-state minima on the ground state surface and can be isolated with illumination at low temperatures and measured using cryo-MX. Characterisation of the ground state is essential for assignment of ultrafast crystallography (chapter 4) and spectroscopy measurements (chapter 6). For now, we will assume rsKiiro follows similar behaviour to the FP Dronpa, although this will be explicitly investigated in later chapters. In Dronpa, evidence suggests the trans-neutral state is formed after 488 nm excitation of the cis-anion. Isomerisation occurs within the first 9 ps which is proceeded with relaxation to a hot-ground state and followed by the deprotonation [5, 8], a ground-state process occurring in 12.5 µs [233]. The cis-anion is formed after excitation of the trans-neutral state at 400 nm. Isomerisation appears to occur within the first 100 ps which is followed by ground state deprotonation on µs time scale [8]. So far, IRTA spectroscopy has not been able to definitively assign the order of rearrangement of the surrounding amino acids in the chromophore pocket.

In this chapter, the ground state behaviour of rsKiiro is studied using cryo-kinetic spectroscopy measurements which identify a so-called 'convex' Arrhenius plot where the activation barrier sharply changes below a transition temperature. We attribute this to a shift in equilibrium between two states at low temperatures which could be structurally distinguishable from the cis/trans global/local ground-state minima. To isolate these states, crystals were illuminated below the Arrhenius break temperatures to accumulate the cryo-trapped states. The structures of these crystals were then resolved using Cryo-MX and compared to thermally annealed and room temperature illuminated structures using Fourier difference calculations and extrapolated structure coordinates.

3.2 Methods

3.2.1 Cryo-Spectroscopy

The rates of photoconversion in rsKiiro were monitored between 180-260 K to extract the activation energy barrier for photoisomerisation using the Arrhenius fitting. A protein-glycerol mix: 15-16 µl glycerol and 4-5 µl protein at a concentration of 14 mg/ml, was pipetted onto a 1 mm sized crystallography loop to give an optical density of around 0.2 when placed at the focus of an inverted microscope (seen in figure 3.2). Glycerol acts as a cryoprotectant and minimises ice formation at low temperatures. The formation of ice could cause a glass transition in the sample being studied. White-light (425-800 nm) was focused through the sample and spectra were collected using a fibre-coupled spectrometer. A 473 nm, 25 mW laser was attenuated using neutral density filters and directed, unfocused, (beam diameter of 1.5 mm) onto the sample at a power density of 0.5 mW/mm^2. This drove the reaction from the cis to the trans-state. A 400 nm LED, was focused using a 2-inch lens, focal length 50 mm to a 5 mm spot size onto the sample. It was attenuated using neutral density filters and a 280-700 nm optical filter to around 1 mW/mm^2. The 280-700 nm filter was used to remove unwanted ultraviolet light produced by the LED source. The 400 nm LED drove the reaction back, from trans towards the thermally-stable cis state. A sample of known yield, Skylan-NS, was used as an actinometrical control between measurements.

The sample was cooled using an Oxford Instruments Cryostream 700 which was positioned around 5 mm from the loop. A thermometer checked the systematic offset from the cryo-stream's set point to the actual sample-temperature. The temperature set point was varied from 180-260 K and room temperature (294 K) in steps of 5 K. Spectra were collected at 2 Hz and the interconversion rates were extracted by fitting a 1st order rate constant to the optical density at 500 nm.

3.2.2 Cryo-Trapping Crystallography

Crystals were preilluminated at 150 K to 'trap' states which only stable only at low temperatures. Batch grown crystals around 10x100 µm in size were grown in conditions of 25% Poly-ethylene glycol 3350 in 0.2 M lithium sulfate, 0.1 M Tris-Cl pH 8.5 with a final protein concentration of 10 mg/ml (similar to [1]) and as described in section 2.2.4. A single crystal was 'fished' in an appropriately sized loop and illuminated with either a 400 nm LED or 473 nm laser at power densities of around 0.5 mW/mm^2. Using the same power densities as found in the spectroscopy ensured that optical heating effects of illumination did not differ. A crystal was assumed to have an optical density of 1. Using the measured rate constants at 200 K it was calculated that 95% conversion would occur after 30 minutes of 400 nm and 69 minutes of 473 nm illumination respectively. As soon as illumination time was complete the crystals were flash-frozen into liquid nitrogen. Macromolecular crystallography measurements were then

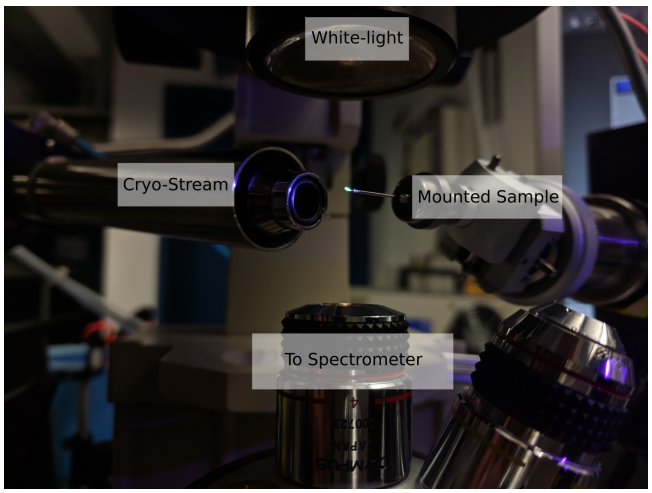


Figure 3.2: A drop of rsKiiro mounted in a loop, the white light source comes from above and passes through an microscope objective to focus it on the sample. Spectra are collected from below using a condenser lens which is coupled into a spectrometer. On the left the positioning of the cryo-stream can be seen.

undertaken at Diamond Light Source, UK. Where the mounting orientation and crystal size allowed, after the first diffraction collection, crystals were annealed for 3 seconds to room temperature and a second structure was obtained from the same crystal (at a different position).

Crystallographic data was highly anisotropic and so was merged using the STARANISO pipeline (Global Phasing Limited, Cambridge, UK) which applies an anisotropic (elliptical on the crystal axes) resolution limit which increases the overall signal to noise for anisotropic crystals. Datasets were then phased with a published structure of Skylan-NS [1] with appropriate mutations (L62A M159T) applied. REFMAC5 [237] was used to refine the data and evaluate the final model.

3.2.3 Q-Weighted DED and Extrapolated Maps

Difference electron density (DED) maps between light and dark conditions were calculated using the program FFT [238] to take the Fourier transform of weighted difference structure amplitudes based on the Bayesian statistics analysis by Ursby *et. al* [174, 175]. As implemented in [239–241] and explained in section 1.4.3, difference structure factors weighted according to:

$$w_i = \frac{1}{1 + \frac{\Delta F_i^2}{\langle|\Delta F|\rangle^2}\frac{\sigma^2_{\Delta F_i}}{\langle|\Delta\sigma_{\Delta F}|\rangle^2}}$$

Published FORTRAN code [241] and custom written Bash scripts were used to calculate the maps. The dark state structure was used for phases in the DED. Structure factors were scaled using the `anisotropic` option in the CCP4i module SCALE IT [238]. This scaled data onto a normalised absolute scale as compared to calculated structure factors which were generated using the CCP4i program SFALL [238]. An upper and lower resolution limit was applied during the scaling step. The high-resolution cutoff was chosen by looking at the trend of calculated Fourier shell correlation (FSC) and choosing

a value close to 0.5 whilst also looking at trends in merging statistics such as $CC_{1/2}$ as a function of resolution shell.

Extrapolated structure factors (F_{Ext}) were calculated using the Q-weighted difference structure factors. A multiple (N_{EXT}) of difference structure factors were added to the dark-state calculated structure factors (F_C) to calculate the extrapolated structure factors:

$$F_{EXT} = F_C + N_{EXT}\Delta F$$

The program FFT [238] was used to calculate the real-space electron density using dark state phases. Negative electron density was then integrated in a 6 Å sphere centred on the chromophore methyl bridge using published FORTRAN code [241]. The point of inflexion, where the magnitude of negative density suddenly increases, was used to determine the characteristic extrapolated multiple, N_{EXT}. This multiple should 're-generate' the structure factors of a model with full occupancy which *caused* the difference density. Model coordinates can then be fitted and refined to the extrapolated density to extract low-occupancy coordinates which would otherwise be very difficult to identify. In this chapter, extrapolated maps are used to refine light-induced structure coordinates for the cryo-trapped structures.

3.3 Cryo-Spectroscopy Reveals Convex Arrhenius Behaviour

This section summaries the cryospectroscopy measurements made by S. Bodet Lefèvre under the supervision of this author and C. Hutchison. Analysis of the thermodynamic behaviour was done by J. van Thor and C. Hutchison with input from this author.

The kinetic rates of photo-conversion close to room temperature are relatively stable with temperature and follow Arrhenius behaviour with an activation energy barrier of 2 and 3.5 $kJ\,mol^{-1}$ for the off-to-on and on-to-off reactions respectively (figure 3.3 and table 3.1). At cryogenic temperatures, the slope of figure 3.3 changes dramatically below the transition temperatures of 229, 208 and 236 K for the on-to-off (500 nm illumination at low temp), off-to-on low temperature (500 nm illumination followed by 400 nm illumination *both* at low temperatures) and the off-to-on annealed (room temperature conversion to the off state followed by low-temperature conversion). Below these transition temperatures the gradient of the Arrhenius plot increases, indicating a larger activation energy barrier is present at low temperature compared to high. This is known as a convex Arrhenius behaviour due to the convex shape of the graph. Thermodynamically, this behaviour can understood as the increased ability of a molecule to explore larger regions of phase-space at higher temperatures [242]. This can increase the probability of being in a non-reactive conformation which will not proceed to the product state [243]. Cryogenic cooling of the molecule lowers this probability and therefore reduces occupancy of this unreactive state which can be implied by a change in activation energy at low temperature [243]. In a protein this could be interpreted as the existence of two competing forms/conformations which dominate in different temperature ranges

[242, 244]. If the lower energy form of the protein is the most reactive then increased rate constants (convex behaviour) will be measured [242]. It is important to note that similar Arrhenius breaks can be observed when a temperature change occurs over a glass transition point of a material [245]. Whilst we cannot completely rule out this effect in the cryo-spectroscopy measurements we find the transition temperature ($T_{Transition}$) is path dependant which therefore supports the thermodynamic interpretation already stated.

In rsKiiro we interpret the changes in activation energy barrier as a shift in equilibrium of population between of two wells on the potential energy surfaces: a 'relaxed' and a 'strained' geometry/state. The transition between Arrhenius barriers (figure 3.3) is modelled as a difference in Gibbs Free Energy (ΔG) between the relaxed and strained states. The overall rate constant, k_{all} can be written as a sigmoidal function on the two observed rate constants:

$$k_{all} = \frac{k_1}{1 + \exp(+\frac{\Delta G}{RT})} + \frac{k_2}{1 + \exp(-\frac{\Delta G}{RT})} \tag{3.1}$$

Where k_1 and k_2 are the rate constants for relaxation of the first and second wells - equivalent to the first and second segments of the Arrhenius plot (figure 3.3). The phenomenological Arrhenius activation energies (E_1 and E_2) and prefactors (A_1 and A_2) can be extracted by taking:

$$k_1 = A_1 \exp(-\frac{E_1}{RT})$$

$$k_2 = A_2 \exp(-\frac{E_2}{RT})$$

To fit equation 3.1. The Gibbs free energy difference of the two states can be defined on the changes of enthalpy (ΔH) and entropy (ΔS) between the two states as:

$$\Delta G = \Delta H - T\Delta S$$

and fitted to the Arrhenius data. The fitted values are summarised in table 3.1 using equation 3.1. As illumination power densities could not be accurately characterised the fitted values are unscaled.

In summary, the cryo-kinetic spectroscopy indicates there is an equilibrium between two states, on either the S_1 or ground-state surfaces. The equilibrium between these states shifts with temperature and causes a change in activation energy barrier height. At cryogenic temperatures, the increased barrier could be evidence of a structural changes due to the adoption of a new ground state structure. It may be possible to cryo-trap this intermediate structure and resolve it using cryo-MX, as will be shown in the next section.

Reaction	E_{A1} [kJ mol^{-1}]	E_{A2} [kJ mol^{-1}]	A_1 [s^{-1}]	A_2 [s^{-1}]	H [kJ mol^{-1}]	S [K kJ mol^{-1}]	$T_{Transition}$ [K]
On-Off	2±0.4	32±3	1.5±0.7	32±5	66±4	0.29±0.02	229±2
Off-On (Low temp. recon.)	3.5±1.2	40±4	0.9±1.2	22±3	51±20	0.24±0.1	208±4
Off-On (Room temp. recon.)	-	55±20	-	26±10	58±25	0.25±0.13	236±10

Table 3.1: Fitted Arrhenius plots values following equation 3.1

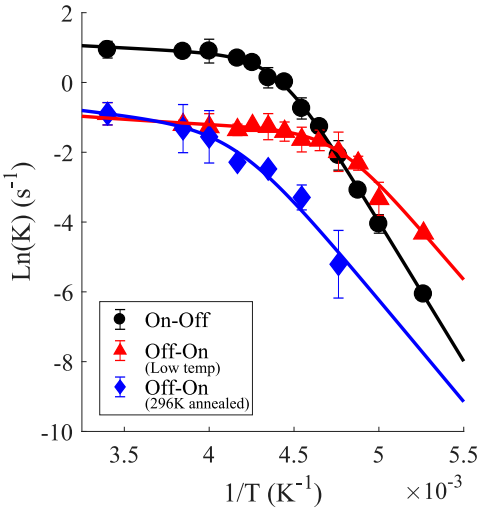
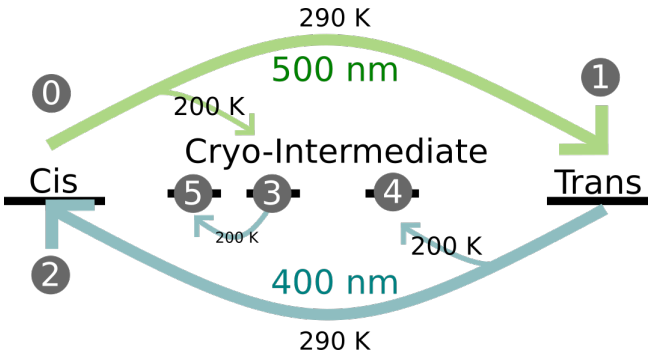


Figure 3.3: Convex Arrhenius behaviour of rsKiiro. Figure prepared by C. Hutchison.

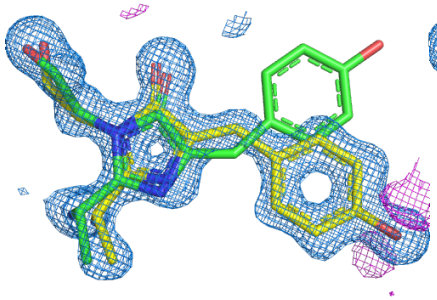
3.4 Cryo-Trapping Steady States with Macromolecular Crystallography

In total, 35 crystal structures of rsKiiro with different illumination schemes shown in figure 3.4 (a). Data was processed using the autoPROC + STARANISO pipeline [177, 246–249] and refined in REFMAC [237] using either anisotropic or isotropic B-factors depending on the resolution limits (table A.2 – A.3). Preliminary selection of the crystallography data was done to find similar unit cell sizes in data, comparable overall Wilson B factors and similar levels of overall signal to noise. Datasets were selected to be as similar as possible to minimise the scaling required for the calculation of Fourier difference maps. Throughout this results section states will be referred to according to the preparation shown in figure 3.4 (a).

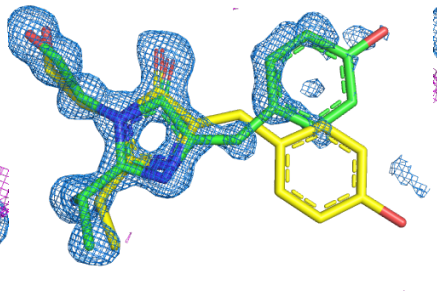
High-resolution crystal structures were collected for all conditions except for the low-temperature 500 nm followed by low-temperature 400 nm illumination (**5** in figure 3.4 (a)). As crystals were mounted in a loop and illuminated for around 1.5 hours in this condition it seems the crystals became too dehydrated/stressed which led to a loss of order in the crystal for state **5**. As seen in tables A.2 and A.3 data was collected to high-resolution, between 0.94-1.45 Å, for all conditions. The highest resolution



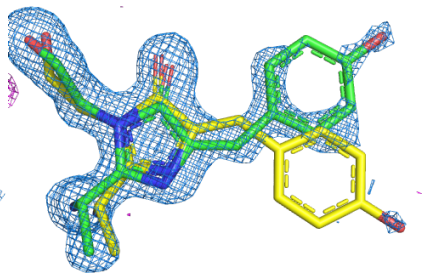
(a) Illumination schemes



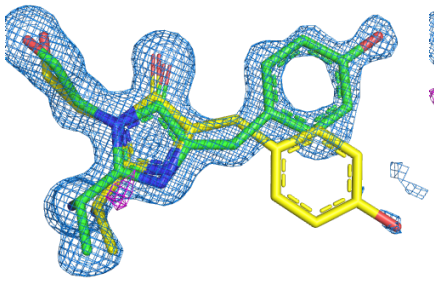
(b) 0



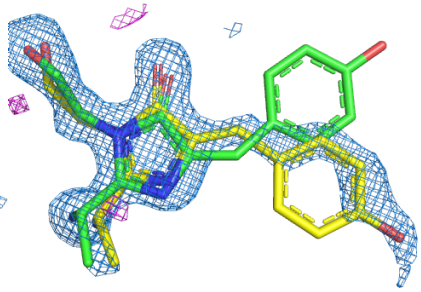
(c) 1



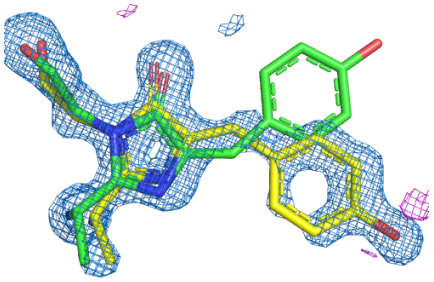
(d) 3



(e) 3a



(f) 4



(g) 4a

Figure 3.4: Showing calculated chromophore-omitted maps for the different structures/illumination conditions labelled in (a). Marine and purple show negative and positive features at 4.00 RMS in the $2mF_o - DF_c$ maps respectively.

structures were obtained for the unilluminated and 500 nm illuminated room temperature condition. Both these structures were highly anisotropic and so were refined with anisotropic B-factors. Other structures were less anisotropic and of lower resolution so were refined using isotopic B-factors instead. Overall, the merged data show excellent statistics and have very similar unit cell sizes with a difference of less than 1% across all unit cell volumes. Initially, to determine the chromophore conformation omit maps (figure 3.4) were prepared with a single round of rigid-body refinement using a dark state

structure with a chromophore occupancy of 0. Difference electron density maps were calculated to determine the extent of structural differences between illumination conditions. Extrapolated structure factors were then calculated and extrapolated coordinates (section 1.4.3) were refined to determine the cryo-trapped structures.

Cis and Trans Structures from Dark and Room Temperature Illuminated Crystals

The dark (**0**) and 500 nm illumination at room temperature (**1**) structures clearly show cis and trans states (figure 3.4) respectively. Both structures were obtained at very high resolution (1.16 and 0.94 Å) and strong refinement statistics (table A.2). Figure 2.8 shows the phased electron density of the chromophore and associated coordinates movement of residues Arg66, Phe173, His194 and Glu212. These structural changes are reflected in the Q-weighed (method defined in section 1.4.3) difference electron density between the two states. This can be seen plotted in 1-dimension above 3σ level and within 2 Å for each atom in the protein chain (figure 3.5). Strong signals are seen at residues closest to the chromophore which is clearly reflected in the coordinates (figure 2.8) and in some areas of the protein backbone around residues which move significantly. This 1-D plot can also be used to gain insight into the structural changes of the cryo-trapped states which will be discussed below.

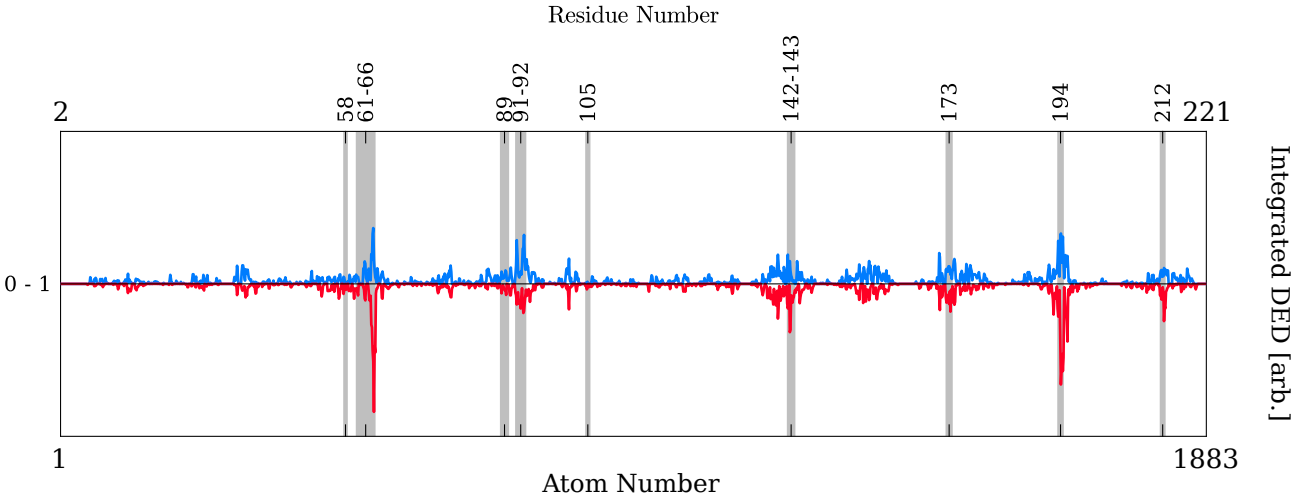


Figure 3.5: 1-Dimensional representation of difference electron density between the dark and 500 nm illuminated state. Density is integrated within 2 Å at 3σ level along each atom of the protein chain, on a normalised scale. Highlighted on the top axis are residue numbers with 3 Å of the chromophore. Code to produced this figure is based on the published tool by Wickstrand *et al.* [250].

Trans-State Intermediate Trapping

The omit map of the 200 K, 400 nm illuminated structure, (**3** in figure 3.4), shows a predominantly trans conformation with a small residual cis population indicated by the density in the position of the cis phenolic oxygen. This confirms that the decrease in absorption at 500 nm at low temperatures (section 3.3) is indeed due to the formation of a trans-state product. An alternative explanation for

the decrease in absorption could be due to the creation of a cis-neutral state which would also likely show a blue shift of the 500 nm peak (Chapter 7). However, the crystal structure clearly excludes this possibility and switching is confirmed, even at cryogenic temperatures.

To directly compare the trapped state pure-trans ground state (**1**) an 'annealed' dataset (**3a**) was collected from the same crystal. Annealing was carried out by blocking the cryo-stream for 3 seconds during which the crystal should briefly return to room temperature for 1-2 seconds. Given the quickest rate constants are on the microsecond time-scale in spectroscopic studies of similar FPs [5, 233], warming to room temperature for this period should be sufficient for any relaxation processes of rsKiiro to occur. Annealing is expected to relax the secondary structure and the chromophore pocket which will allow the trapped structure to move out of the local minima accumulated at low temperatures and led to the formation of the final ground-state product, **0**.

The annealed omit map (**3a** in figure 3.4) shows slightly stronger density on the trans-state, especially around the top of the phenol ring. This is indicative of increased order of the trans chromophore after annealing. This could imply a partially occupied intermediate was present with an activation barrier that could not be surpassed at cryogenic temperatures. An argument could be made that this change in omit density could be due to differences in data quality between **3** and **3a**. However, both datasets were taken from the same crystal and show very similar merging statistics (table A.2 and A.3). **3a** structure has only a slightly higher resolution (1.21 Å) compared to **3** (1.29 Å) and refined coordinates are almost identical in fitting metrics (both with R_{Free}/R_{Work} 0.19/0.16). It therefore seems likely that the changes in omit density are caused by the presence of an intermediate state.

Difference electron density is extremely sensitive to small changes in structure and can be used to extract intermediate-state structural differences which are not clear in the modelled coordinates. Shown in figure 3.6 is the 1-D plot of difference electron density comparing the pure trans (**1**), the trapped (**3**) and the annealed (**3a**) states. At first, the signals look very different from those in the purely cis and trans difference density (figure 3.5). However, on closer inspection both **3-1** and the **3a-1** maps show largest signals at the same positions with a similar distributions as the signals the trans-cis signals (figure 3.5). For example the DED signals around residues 91 and 61-66 indicate the **3** and **3a** structures have residual cis-state population. However, there are also new DED signals which cannot be assigned from only cis/trans isomerisation signals. We, therefore, suggest the **3** state is a mixed occupancy of at least three conformational states: the cis, trans and an intermediate. This mixture of states makes analysis of an intermediate structure difficult but not impossible so, extrapolated maps were calculated and cryo-trapped coordinates were fitted.

The extrapolated map calculation gave a characteristic N_{EXT} value of 8.8 (figure 3.7 (a)). The density

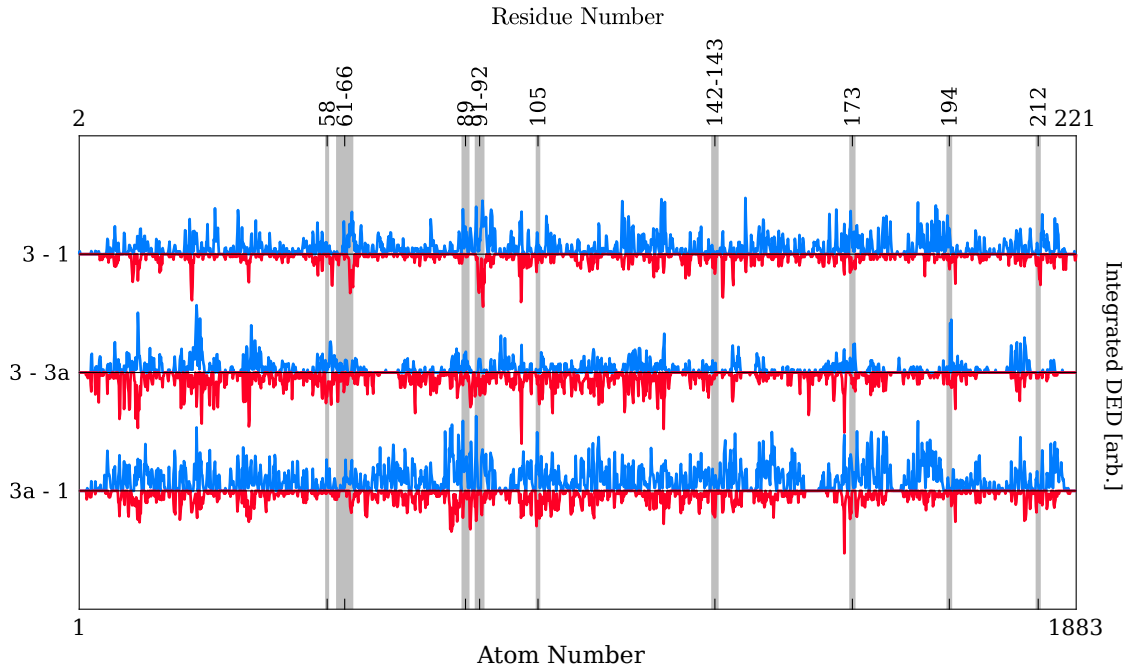
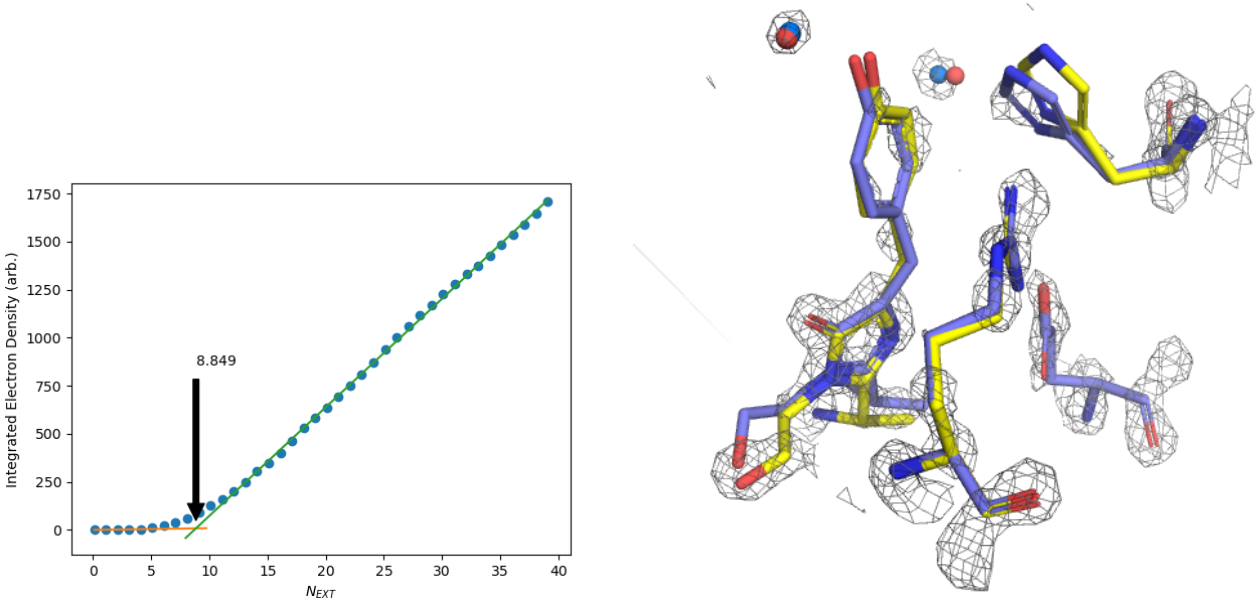


Figure 3.6: 1-Dimensional representation of difference electron density between the 500 nm illuminated (1) and the 500 nm illuminated at 200 K states. Density is integrated within 2 Å at 3σ level along each atom of the protein chain, on a normalised scale. Highlighted on the top axis are residue numbers with 3 Å of the chromophore. Code to produced this figure is based on the published tool by Wickstrand *et al.* [250].



(a) Integrated electron density as a function of N_{Ext} **(b)** Extrapolated coordinates (blue) and density at 1.5 σ

Figure 3.7: (a): Plots of integrated electron density in arbitrary units against multiples of added difference electron density (N_{Ext}) for different illumination powers. **(b):** Extrapolated electron density (grey) plotted at 1.5σ with the extrapolated-refined coordinates in blue and the dominant trans conformations shown in yellow.

and refined coordinates are shown in figure 3.7 (b) and show a slightly tilted trans conformation and twisted in residue histidine 194. Although these changes in atomic positions could be interpreted as an intermediate state the extrapolated electron density is not well defined around the chromophore phenol ring *or* the histidine. It therefore stands to reason that there are potentially multiple intermediate states isolated in structure **3** and the exact interpretation of the electron density is unclear. Perhaps with higher structural homogeneity, it may be possible to extract more accurate extrapolated coordinates which could then be related to an intermediate structure.

Cis-State Intermediate Trapping

Structure **4** was preconverted with 500 nm light at room temperature and then illuminated at 200 K with 400 nm light. The structure clearly shows a cis conformation (figure 3.4 (f)) but the density around the bottom of the phenol ring is poorly defined. The annealed structure (**4a**) shows increased omit-density strength which suggests the chromophore cis-state is more highly populated after annealing (figure 3.4 (g)). This change indicates some form of structural change caused by the thermal relaxation process allowed at room temperature which is frozen-out at cryogenic temperatures.

The 1-D DED plots (figure 3.8) do not show signals which are suggestive of large differences in cis/trans population when compared to the pure cis state (**0**). This is seen in the **4-0** plot which shows mostly non-site-specific signals which is surprising given the main changes were expected in only the chromophore region. While the **4-4a** DED shows very strong difference signals, the **4a-0** shows much lower integrated DED amplitudes. This directly suggests that the **4** state originates from a different molecular conformation/state than either **0** or **4a** *and* implies that **0** and **4a** are structurally similar - i.e. that thermal annealing relaxes the trapped intermediate.

Initially, the refined coordinates of **4** (figure 3.9) suggest the phenol ring is flattened into the chromophore ring plane when compared to the ground state/relaxed cis structure (purple coordinates figure 3.9). The calculation of Q-weighted difference density (figure 3.10 (a)) clearly shows complementary positive/negative signals on the chromophore phenol ring which support the observations of the shifted coordinates. Furthermore strong signals are seen in the hydrogen-bonding network of water molecules which suggests the waters move after annealing of the crystal.

The calculation of extrapolated structure factors allows regeneration of the electron density which caused the initial difference density. This technique allows for partially occupied states to be modelled on the extrapolated electron density obtained from the characteristic N_{EXT} value (figure 3.10 (c)). In the case of condition **4**, interpretation of the DED suggests there is a partially occupied intermediate state. The refined coordinates show an ensemble average of the relaxed and cryo-trapped states which are too close in geometry to model an alternative conformation. The extrapolated density (figure 3.10

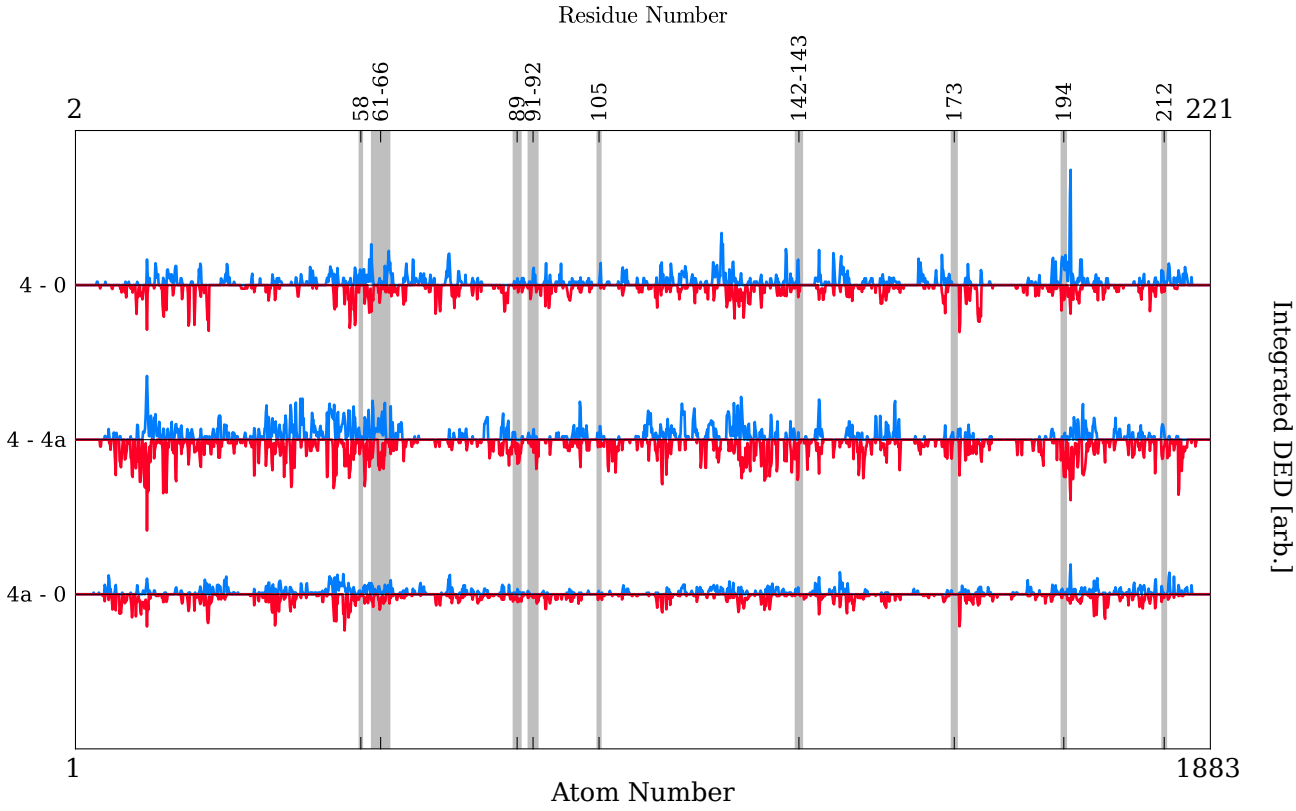
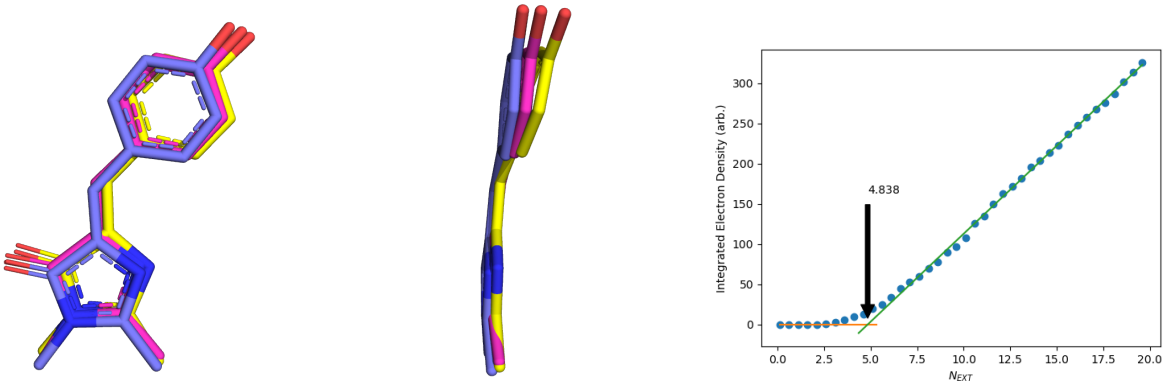


Figure 3.8: 1-Dimensional representation of difference electron density between the 400 nm illuminated (0) and the 500 nm preconverted followed by 400 nm illuminated at 200 K and annealed (4a) states. Density is integrated within 2 Å at 3σ level along each atom of the protein chain, on a normalised scale. Highlighted on the top axis are residue numbers with 3 Å of the chromophore. Code to produced this figure is based on the published tool by Wickstrand *et al.* [250].



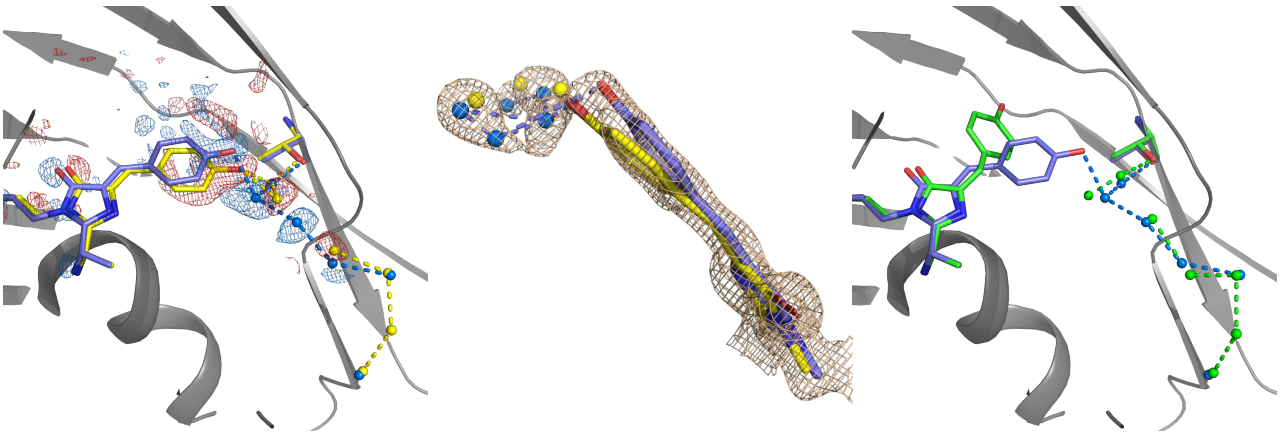
(a) **(b)** **(c)** Extrapolated Density Fitting and Chromophore Geometry

Figure 3.9: (a) and (b): Chromophore coordinates for the purely cis state, **4a** in yellow, the refined coordinates of a single conformation fitted to the cryo-trapped intermediate structure, **4** in purple and the extrapolated coordinates of **4** shown in blue.

(b)) and subsequent refinement of extrapolated coordinates in real space clearly show a new intermediate state. Changes in the chromophore coordinates, with a planar phenol/imidazolinone ring geometry with the rearrangement of a series of hydrogen-bonded waters, are strong evidence of a cryo-trapped intermediate state in **4** which relaxes on returning to room temperature. Further comparison to the trans-state coordinates (figure 3.10 (c)) shows that the new arrangement of water molecules is unique to the cryo-trapped intermediate state.

3.5 Discussion

Overall, we show that switching of rsKiiro proceeds at cryogenic temperatures in crystals and show that two intermediate states exist on the ground state surface. The intermediate closest to the cis product appears to be well-defined and is well resolved in the crystallography measurements. The intermediate closest to the trans coordinate was less well defined, although this is possibly due to incomplete conversion during preillumination. A convex Arrhenius plot behaviour has also been seen in the positive switching RSFP, Padron where crystallographic cryo-trapping also showed a similar different geometry [251]. This is opposed to Dronpa where switching is not seen in crystals or solution below the glass transition temperature (around 180 K) [251]. An explanation for this is that a 'breathing' mode may be required for switching in Dronpa [252] which is prevented by low temperatures. Given rsKiiro closely follows the behaviour of Dronpa we would expect to also see no isomerisation at low temperatures. However, as both the low temperature on and off-state conditions show isomerisation we can conclude that the ground state activation barriers are lower in rsKiiro and show a local minima of a flattened neutral-cis-state must not be accessible at low temperatures in Dronpa. It could be that the breathing motion of the beta-barrel is not required for switching in rsKiiro or that the crystal packing



(a) Q-Weighted DED **(b)** Extrapolated Density **(c)** Extrapolated and Trans coordinates

Figure 3.10: (a): Q-weighted difference electron density plotted at $\pm 2\sigma$ in blue and red respectively. Blue coordinates show the extrapolated model and yellow shows the pure-cis/annealed coordinates. **(b):** Extrapolated and relaxed coordinates shown with the extrapolated density plotted at 2σ **(c):** The Extrapolated coordinates plotted with the purely-trans coordinates (green).

arrangement is less-constrained [231] and so allows for greater flexibility, even at lower temperatures.

Given the small barriers that were extracted in table 3.1 it is important to consider the global heating effects of illumination. The maximum possible heating rate of the solution and crystals caused photon absorption of the CW laser was calculated to be less than 2 K/s based on an adiabatic calculation using $\Delta T = \frac{E_{abs}}{mc_p}$ where c_p is the heat capacity of water, E_{abs} is the absorbed photon energy and m was taken as the mass of a crystal. Given the crystal/solution are well thermally grounded by the cryostream and the loop the most likely remained constant during the illumination cycle. Although the overall sample temperature is unlikely to change an individual chromophore will increase by several hundred Kelvin with the absorption of each photon. Molecular dynamics [253] and classical thermodynamics calculations [254] have indicated that this temperature increase diffuses on picosecond timescales. The ground-state barrier for isomerisation has been calculated by DFT to be around 122 $\mathrm{kJ\,mol^{-1}}$ for the isolated chromophore [8] and is measured in Chapter 5 to be around 90 $\mathrm{kJ\,mol^{-1}}$. Given that a barrier of this size would require temperatures of several hundred Kelvin to be surpassed, a ground state isomerisation seems unlikely. We therefore tentatively conclude the observation of low-temperature isomerisation supports excited-state isomerisation mechanism reaction proposed by Warren *et al.* [5] though is not conclusive.

3.6 Conclusions

Overall, a clear shift in activation energy is seen in the spectroscopic data. Ideally, to find the optimum crystal illumination times and to decide whether condition 3 is in the expected intermediate space, spectroscopic measurements should be repeated on crystalline samples. However, condition 4 clearly

shows an intermediate conformation that is thermally unstable and relaxes after thermalisation. It was noted that the largest signals seen in **4-4a** also occur on residues more prone to radiation damage [141, 255]. This data was collected using a low dose and a high multiplicity (a large number of images with a short exposure time) to decrease radiation damage risk and does not display any signs of damage from the merging and scaling statistics. Ultimately, to find out whether these intermediates are accessed in the room temperature reaction coordinates, it will be interesting to compare these structures with the intermediate structures found to have ultrafast TR-SFX (chapter 4) and spectroscopy (chapter 6) results.

4 | Ultrafast Optical Control Experiments with Time-Resolved Serial Femtosecond Crystallography

In this chapter, ultrafast TR-SFX is used to resolve the response of rsKiiro to femtosecond pump and pump-dump excitation. Pump-probe sub 1 ps structural differences show a translated and slightly twisted chromophore geometry close to the trans-state chromophore geometry. The addition of a dump pulse, known to quench product-formation of the cis-state (chapter 6), causes enhanced difference electron density signals which indicate an increased occupancy of the same ground-state intermediate observed with pump-probe excitation. This intermediate is assigned to a vibrationally excited ground state structure not an excited state product has been assumed by other TR-SFX studies. These results have implications across the field of TR-SFX and show that assignment should be tested in ultrafast crystallography experiments.

Note to the reader: This work was part of a large collaboration. The experiments were led by C. Hutchison and J. van Thor. This author led the crystallographic analysis with J. van Thor. The chapter is currently under preparation for publication as:

> C D.M. Hutchison, J M. Baxter, A Fitzpatrick, G Dorlhiac, A Fadini, S Perrett, K Maghlaoui, S Bodet Lefèvre, V Cordon-Preciado, J L. Ferreira, V Chukhutsina, D Garratt, J Barnard, G Galinis, F Glencross, R M. Morgan, S Stockton, B Taylor, L Yuan, M G. Romei, C Lin, J P. Marangos, M Schmidt, V Chatrchyan, T Buckup, D Morozov, J Park, S Park, I Eom, M Kim, D Jang, H Choi, H Hyun, G Park, E Nango, R Tanaka, S Owada, K Tono, T Nakane, D P. DePonte, S Carbajo, M Seaberg, A Aquila, S Boutet, A Barty, S Iwata, S G. Boxer, G Groenhof, J J. van Thor. Amplification of ultrafast structural motion in a photo-excited fluorescent protein (*in preparation for submission to Science*)

4.1 Introduction

As outlined in section 1.3, it is necessary to generate significant population transfer to the excited and photoproduct states to create detectable DED signals. In ultrafast measurements, this requires the use of an intense optical field for photoexcitation. This can be problematic as intense light can drive non-linear and multiphoton excitation events [256] which are not realistic of biological function [189, 257] and can create multiple excited states which are difficult to disentangle. As a result, the bleach rates and branching ratios to higher-order electronic excited states should be measured to determine the most appropriate peak-powers for use in TR-SFX experiments [211, 258]. Currently, there is only one published ultrafast TR-SFX experiment where the optical excitation conditions have been chosen with explicit consideration of optical pumping parameters: a study by Pande *et al.* on the photoactive yellow protein (PYP) [259]. Here, the optimal pump energy densities and pulse duration was calculated based on systematic studies of the photoconversion yields for crystalline samples [211]. Full knowledge of the photocycle [260, 261] and non-linear cross-sections as functions of the peak power, pulse length and second-order dispersion [258, 262] was applied to find the energy density which maintained low photo-bleach/multiphoton absorption with the highest photoproduct yields [211]. In crystals, a maximum single-shot yield of between 5-7% was measured at 0.25 mJ/mm^2 with an irreversible bleach close to ∼1% [211]. Pande *et al.* used these energy densities for the pump pulse and collected TR-SFX structures from 100 fs – 3 ps. They observed the formation of a new state after at 590 fs [259], although it now seems the timing should have been shifted by ∼100 fs due to systematic and unavoidable error in the time zero cross-correlation [114]. The comprehensive characterisation of spectroscopic cross-sections and the excited state behaviour made it possible to assign this state to a previously unobserved intermediate structure before the first step in the photocycle [259].

Two recent experiments investigating the dynamics of bacteriorhodopsin have highlighted the need for careful characterisation of laser pulse energies and control of the illumination scheme. Kovacs *et al.* [263] and Nogly *et al.* [264] both made sub 1 ps measurements to determine the initial steps involved in the retinal chromophore isomerisation. Nogly *et al.* [264] used a pump laser at 30 Hz whilst probing at 120 Hz with the XFEL (i.e. 1 laser shot followed by four X-ray pulses). They used an oval-focus to excite multiple positions in the LCP injection matrix and excite simultaneously, a sub 1 ps and an 8.3 ms time point. As the laser focus was poorly defined in focal-spot size (and therefore power density), it is unclear what the nominal pump energy density was for each time point, though it is estimated to be around 14 mJ/mm^2. The TR-SFX DED, showed a rapid onset of ultrafast protein motions in the first 1 ps which they assigned to a protein quake motion [264, 265]. This is a mechanism which has been proposed to disperse energy from a chromophore after absorption of a photon, analogous to macroscopic earth-quakes [266–268]. Conversely, Kovacs *et al.* [263] used a pump energy of 4 mJ/mm^2

in the TR-SFX experiment, but also made visible-TA measurements up to 0.25 mJ/mm^2 in addition to quantum chemistry calculations. They did not observe the same protein quake motions, although this could have been due to poorer crystallographic data quality [263]. Overall, they concluded the intense pump pulses used by then (and Nogly *et al.*) were driving multiphoton absorption pathways and possibly multiphoton-induced ionisation [263, 269, 270]. TA and quantum chemistry calculations suggested that double excitation would still lead to coupling of the same function modes, meaning the observed structural changes (which were very similar to Nogly *et al.*) were thought to still be representative of intermediates photocycle of bacteriorhodopsin. However, the peak TA energy densities were well below the TR-SFX pump powers and quantum chemistry calculations are known to be challenging on retinal proteins [271]. This means it is not clear if the observed DED by Nogly *et al.* and Kovacs *et al.* are actually directly relatable to the photocycle under physiological conditions.

These discrepancies firstly highlight the need for carefully controlled illumination regimes to avoid multi-photon effects but also raise the question of assignment and connectivity from structures in TR-SFX experiments. In ultrafast TA measurements, multi-pulse approaches can be used to make resolve ambiguous reaction schemes (section 1.5.4). In this chapter, we propose that the same multi-pulse methodology can be used for structural assignment in ultrafast TR-SFX experiments. Visible TA measurements on the protein developed in chapter 2, rsKiiro, show a stimulated emission with a Stokes peak at around 520 nm (C. Hutchison, unpublished). A dump pulse at 515 nm has been shown to pump the stimulated emission from the S_1 to the ground state with around 50% efficiency (C. Hutchison, unpublished). The energy density of the pump pulse was chosen by power titration measurements using flash photolysis. A pump-dump delay of 350 fs was chosen as electronic dephasing should have completed after interaction with the pump pulse (400 nm). This ensures that coherent Raman process are not being driven by the arrival of the second pulse [272]. Similar approaches of assignment may be applicable for experiments such as those on bacteriorhodopsin [270, 273] to clarify different assignment by different groups [263, 264], other reversibly switchable fluorescent proteins (RSFPs) [212, 274], and myoglobin [266].

This chapter combines the work of three major experimental efforts at three different XFELs facilities (beamlines): the LCLS (CXI) [35], SACLA (EH2) [275] and PAL-XFEL (NCI) [64]. Through this chapter, they will be referred to as LCLS, SACLA and PAL experiments. The nature of this work was highly collaborative and a summary of all authors who contributed is in section A.4. J. van Thor and C. Hutchison wrote the beamtime applications and led the experimental efforts. This author was involved in every part of the XFEL experiments and independently led the processing and interpretation of the crystallography data analysis with J. van Thor. Consequently, this chapter is structured to firstly, explain the overall experimental method/execution before focusing on the independent development of

a post-merging data analysis pipeline. The results of this analysis are then shown and discussed in context to the mechanism assignment of rsKiiro.

4.2 Experiment Overview and Implementation

These experiments aimed to resolve the trans/cis (off/on) dynamics of rsKiiro and to structurally resolve optical control from 200 fs – 100 ps. A 400 nm pulse was used as a pump and 515 nm used as a Stokes/dump pulse. Microcrystals were prepared using the method in section 2.2.4 to µm sized dimensions and densities of around 10^7 crystals/ml. The exact buffer composition and size were optimised based on sample injection technique (which was difference at each facility) to maximise hit-rates without reducing indexing performance. The sample was prepared in the off state by preilluminating with a laser (SACLA: 10 mW, 473 nm, PAL-XFEL: 100 mW, 488 nm) or LED source (LCLS: central wavelength 500 nm, 300 mW) at power density calculated to give full conversion before injection into the interaction region which was then confirmed by crystallographic analysis during the experiments. Crystal microslurries were injected using a gas dynamic virtual nozzle (GDVN) [120] or double flow-focusing nozzle (DFFN) [121] (LCLS), droplet injector (SACLA) or with lipid cubic phase (LCP) matrix injection in a syringe based system (PAL-XFEL) (section 4.2.1). The pump and dump laser delays were set after time-zero determination by cross-correlation scans with the X-rays. For pump-probe delays, a timing tool was used to account for inherent XFEL jitter (section 4.2.4) and sort each X-ray shot into 150 fs timing bins based on the relative X-ray pulse arrival time. Measured Bragg peak intensities were merged using the XFEL data processing program, CrystFEL [102]. The analysis of the merged crystallographic data led to the calculation of difference electron density maps which were then fitted with light-state coordinates through the generation of extrapolated electron density (section 4.3)

4.2.1 Sample Injection

At LCLS, a 75 µm inner-diameter GDVN tip (figure 1.4 (a)) was ground to remove the opaque coating that would normally prevent preillumination at the tip of the nozzle. Flow rates of around 10-30 $\mu l\,min^{-1}$ were generated using a high-pressure liquid chromatography pump which pushed a plunger system which then extruded crystals down the 100 µm inner diameter capillary. The capillaries were around 1.5 m in length and led to the interaction region where the nozzle was mounted. As the capillaries were semi-transparent the crystal slurry was illuminated on transit in to the CXI vacuum chamber. Hit rates fluctuated from 5-15% using the GDVN. A DFFN injection system (figure 1.4 (b)) was used at slightly lower sample flow rates (10-20 $\mu l\,min^{-1}$) and around 90 $\mu l\,min^{-1}$ ethanol flow. The DFFN nozzle was found to give very poor stability/reliability of the jet and so a majority of data was collected with the GDVN nozzle.

At the SACLA experiment crystals were delivered into the interaction region using a droplet-on-demand injector system (figure 1.4 (c)) which extruded droplets via a piezo-driver. The voltage and amplitude of the current pulse used to drive the piezo, as well back-pressure of the droplet injector could be varied to optimise the size/shape of the droplet being formed. The sample was back-loaded through the injector nozzle (diameter of 80 µm) which acted as a filter and prevented large particles blocking the injector. Around 300 µl of sample was loaded into the reservoir every 1-1.5 hours. More sample was not loaded as crystals began to settle over, which could block the injection nozzle. Timing relative to the XFEL master clock and the piezo-driver parameters were periodically adjusted and optimised as debris built up from the obliterated sample droplets. Ideally, hit rates of around 50-60% were maintained throughout the experiment as it was found this gave the best stability in indexing rates. This was tracked using real-time hit-finding in the program Cheetah [145]. It was found the 50% hit-rate was optimal for maintaining high indexing rates. When hit rates are too high multiple crystals can overlap droplets and shadow the crystals from illumination and can cause miss-indexing of patterns.

At PAL the microcrystal-slurries were resuspended in a lipid cubic phase (LCP) of monoolein. This was loaded into an adapted 100 µl syringe (Hamilton gas tight) which was driven with a syringe pump at a constant flow rate of 6.35 $\mu l\,min^{-1}$. The syringe needle had an inner diameter of 0.413 mm and extruded a thin jet at a speed of 6 $mm\,s^{-1}$. This led to a translation of around 200 µm/pulse at the fixed XFEL pulse repetition rate of 30 Hz. The only transparent area for preillumination was the base of the main body of the syringe before it interfaces with the needle fitting. As the transit-time through the preillumination area was on the order of the full-conversion time the injection speed was carefully controlled to maintain a constant preillumination population. This was also important to control for pump-laser scatter up up the monoolein/crystal jet. The syringe extrusion speed is much slower than the GDVN injection and is continuous - unlike the droplet injector. This can led to a laser scatter up the LCP matrix caused by total internal reflection of laser light which pre-exposes un-probed crystals to the 400 nm pump light. The choice of an injection speed is, therefore, key to the success of an experiment as we aimed to maximise pre-conversion and minimise the effect of pump-laser scatter.

4.2.2 X-ray Diffraction Data Collection

At LCLS's CXI hutch, diffraction data was acquired using the in vacuum Cornell Stanford Pixel Area Detector CS-PAD [117] with a detector distance optimised to 50.7 mm. The X-ray energy was set to 9.5 keV with pulse energy around 1 mJ at 120 Hz. To control for fluctuations in the SASE X-ray energy spectrum an upstream spectrometer was used to determine the average energy for each X-ray pulse. This was processed by the DAQ and written into the header of each image for use in downstream processing. To control the dynamic range of the detector a gain-mask was applied to avoid saturation of low-resolution components but still allow sensitivity to higher-resolution reflections. Each pixel of

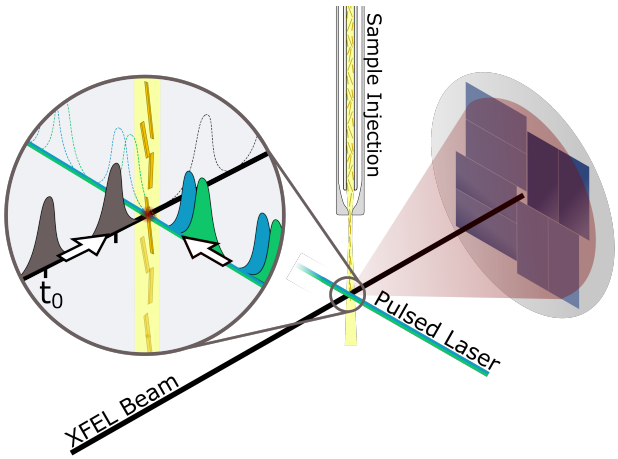


Figure 4.1: A crystal slurry is injected into the interaction and excited by a 400 nm (blue) pump pulse, followed by a dump pulse at 515 nm (green). X-rays probe the sample and diffraction is collected on the detector.

the detector has a high/low gain mode. In the high gain-mode pixels are relatively sensitive but have lower saturation thresholds/maximum-photon counts. In low gain-mode pixels exhibit a higher dynamic range but are less sensitive to individual photons.

The detector used at SACLA's EH2 hutch was the multi-port charge-coupled device (MPCCD) III-L [276]. The detector distance was set to a distance of 50.6 mm. The pulse energy was set to 10 keV with an average pulse energy of around 0.6 mJ at 30 Hz. As at LCLS, an upstream spectrometer measured energy of each X-ray pulse which was incorporated into the data processing pipeline. The MPCCD III-L has only a single gain mode meaning there was no need post-processing of the gain mask (like at LCLS). This detector does not require a vacuum to operate meaning it was operated unenclosed, at atmospheric pressure. To minimise the background scatter caused by air, a helium stream was used to flush the sample area and around the detector.

At PAL, a commercially available Rayonix (Illinois, USA) MX225-HS was used to acquire diffraction data. To increase the dynamic range and allow continuous 30 Hz data collection rates, a 4×4 pixel binning mode was set. An X-ray energy of 12.4 keV with a pulse energy of around 0.6 mJ was chosen as a compromise between X-ray photon flux, detector efficiency and to maximise the obtainable resolution. The detector distance was set to 142.9 mm.

4.2.3 Optical Laser Setups

At SACLA and LCLS an on-site synchronised Ti-Sapphire laser system was used to generate the pump and dump pulses. The the 400 nm laser pulse was generated using a 100 µm (barium borate) BBO cut at a phase-matching angle (29.2°) for type II SHG of 800 nm, the fundamental wavelength of the lasers used. An OPA was used to generate the 515 nm pulse and the two beams where combined using a 490 nm wavelength-cut long-pass dichroic mirror (DMLP490, Thorlabs). The two colours were

initially temporally-overlapped using a fast-photodiode with rise-times close to 20 ps. In the path of the dump beam there was an delay stage to set the relative delay between the pump and dump pulses. A cross-correlation measurement on a Ce:YAG screen determined the average delay between the X-rays and optical pulses. This was then jitter corrected to give a time resolution better than 200 fs (section (4.2.4). A chopper wheel cut the repetition rate of the laser to match the XFEL experiments (120 Hz LCLS, 30 Hz SACLA, 30 Hz PAL-XFEL) and was used to interleave dark and light data. At SACLA and LCLS four laser-on shots were followed by one laser-off whilst the PAL-XFEL experiment used a 1:1 ratio to minimise variation in scatter effects when using the LCP extruder. The pump and dump lasers were aligned to the X-ray focus using the fluorescence of a YAG screen (SACLA and LCLS).

The energy-density of the pump pulse was chosen to maximise the photoproduct yield. In the home lab, flash photolysis measurements titrated the energy density of a 150 fs, 400 nm pulse and measured the bleach and photolysis yield with increasing pulse energy density. It was found an energy density of around 4 mJ/mm^2 was optimal. At the facility, the focus size was chosen maximise overlap stability with the X-ray focus and minimise scatter to the next crystal. Focus size was determined using either a beamprofiler (SACLA) or knife-edge scan (LCLS). An energy meter measured the pulse energy which was then attenuated using a wave-plate and polarizer to around 4 mJ/mm^2 and 8 mJ/mm^2 for both the 400 and 515 nm beams respectively.

The crystals of rsKiiro were needle shaped which meant that had a tendency to self-align with the injection jet. Protein crystals are non-centrosymmetric which means they exhibit birefringent electric field decomposition along their optical axes [277] (not to be confused with the crystallographic unit cell axes). This can enhance or reduce the effective pumping of the electronic transition dipole depending on the orientation of the crystal relative to the laser polarisation axis [278]. As a result, to reduce the risk of the pump not overlapping with the chromophore transition dipole moment, the laser polarisation was oriented at 45° relative to the jet axis.

The optimised energy densities were measured in the home lab using a 150 fs pulse. At LCLS and SACLA, to match this pulse width and to allow coherence analysis, a home-built, portable FROG [279] and X-FROG [280] were deployed at the beamlines. The FROG measurements determined the pulse width and spectral phase of the optical pump and dump pulses. We made SHG-FROG measurements of the 800 nm (fundamental) and 515 nm beams and X-FROG measurements of the 400 nm pulses by sampling a small section of the beam offline from the experiment.

The SHG-FROG schematic is shown in figure 4.2 (a) the fundamental beam was split using a thin beam splitter. A linear stage delayed one of the pulses relative to the other. A focusing mirror was then used to focus both beams onto the BBO which generated the second harmonic of the beam which

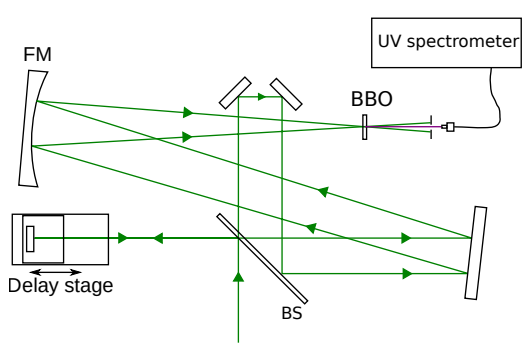
		λ [nm]	$\Delta\lambda$ [nm]	t [fs]	GDD [fs^2]	3rd Order [fs^3]
LCLS	Fundamental (after BBO)	800	15.5	100 ± 15	554 ± 50	2944 ± 600
	Pump	400	-	100	-	-
	Dump	515	6.8	71 ± 15	480 ± 40	2000 ± 400
SACLA	Fundamental (after BBO)	800	27	72	262 ± 12	-
	Pump	400	9	40	135 ± 15	-
	Dump	515	14	52	355 ± 175	184 ± 20

Table 4.1: FROG fitting results from LCLS and SACLA. Adapted with permission from Hutchison *et al.* (unpublished).

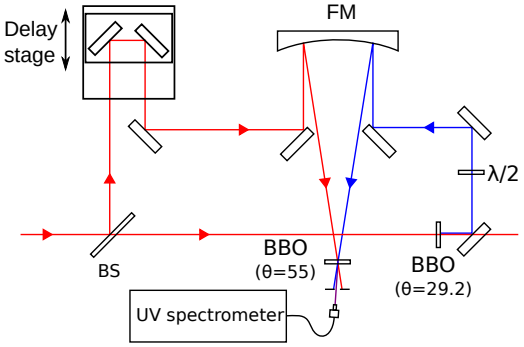
then co-propagated with the fundamental. When both beams were overlapped spatially and temporally, the two photon SHG process could occur using a single one photon from each arm of the FROG. This led to a cancellation in k vectors and the generation of a third beam which between the two collinear components. This central spot was collected using a fibre spectrometer. By scanning the delay stage a frequency/delay spectrogram was obtained that could be used to retrieve the pulse duration and spectral phase using phase reconstruction. A type II 20 µm surface deposited BBO with a 29.2° cut angle measured the 800 nm pulse and type II with 55° was used to measure the 515 nm. The orientation of the crystal with respect to the beams was adjusted to maximise phase matching in both amplitude and bandwidth. An example FROG trace and fitted pulse retrieved is shown in figure 4.2.

The XFROG (figure 4.2 (b)) measurement was required due to the limitations in finding a non-linear medium which could phase match the 400 nm pulses. The XFROG measurement is made similarly to the FROG measurement except the resolved frequency mixing is signal is not from two pulses of identical frequency. In our XFROG set-up, we split the test pulse (800 nm) and generated the second harmonic in one arm with the same 100 µm 29.2° BBO used in the main optical setup. This leads to a 400 and 800 nm beams being focused on the phase-matching medium. (another type II BBO with phase matching angle of 29.2°). The frequency-mixed output is collected using a fibre spectrometer and spectrally resolved as a function of delay. This trace (figure 4.2 (c)) is then analysed to extract the phase and pulse length.

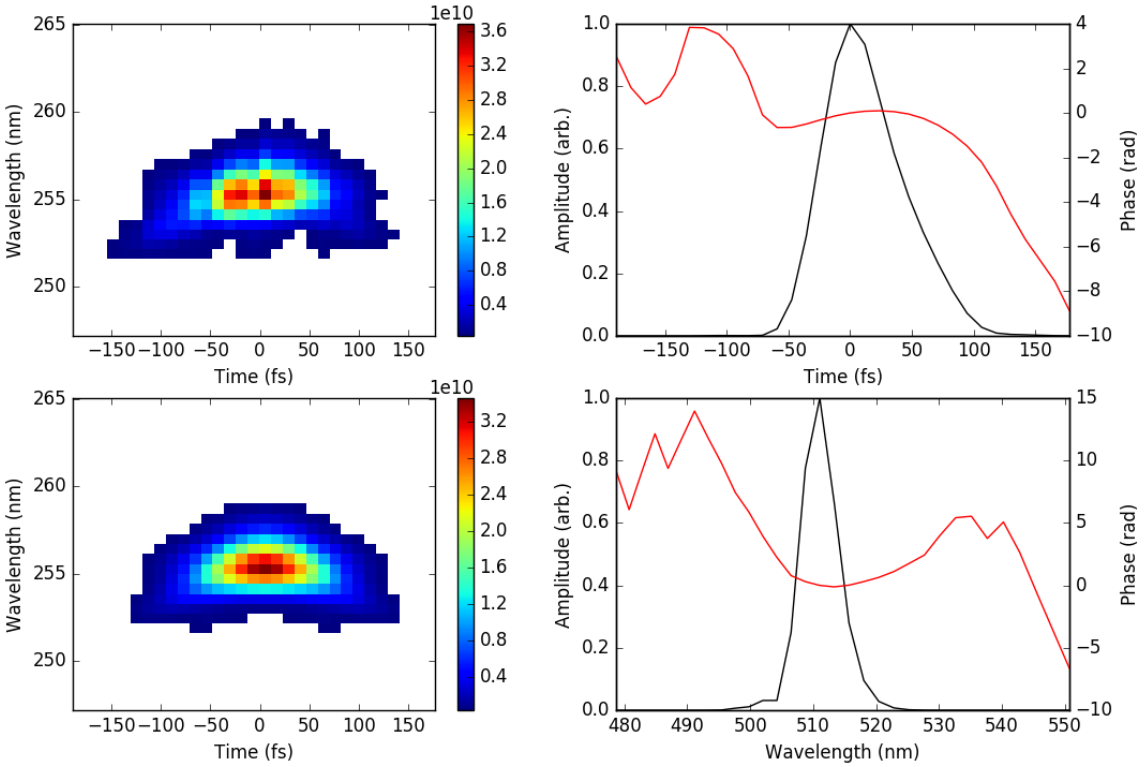
The spectrograms for the SHG-FROG and XFROG were fitted using a published MATLAB FROG-fitting tool (Adam Wyatt, MathWorks FileExchange). A summary of the fitted pulse lengths for experiments at SACLA and LCLS is shown in table 4.1.



(a) SHG-FROG



(b) XFROG



(c) SHG-FROG of the 515 nm pulse at LCLS. Raw spectrogram (top left) with, reconstructed spectrogram (bottom left), recovered pulse temporal profile (top right) and spectral profile (bottom right).

Figure 4.2: (a): Schematic for the second harmonic generation frequency-resolved optical gating (SHG-FROG). **(b):** Cross-correlation FROG (XFROG) **(c):** Measured and retrieved FROG spectrogram, pulse width and spectrum. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

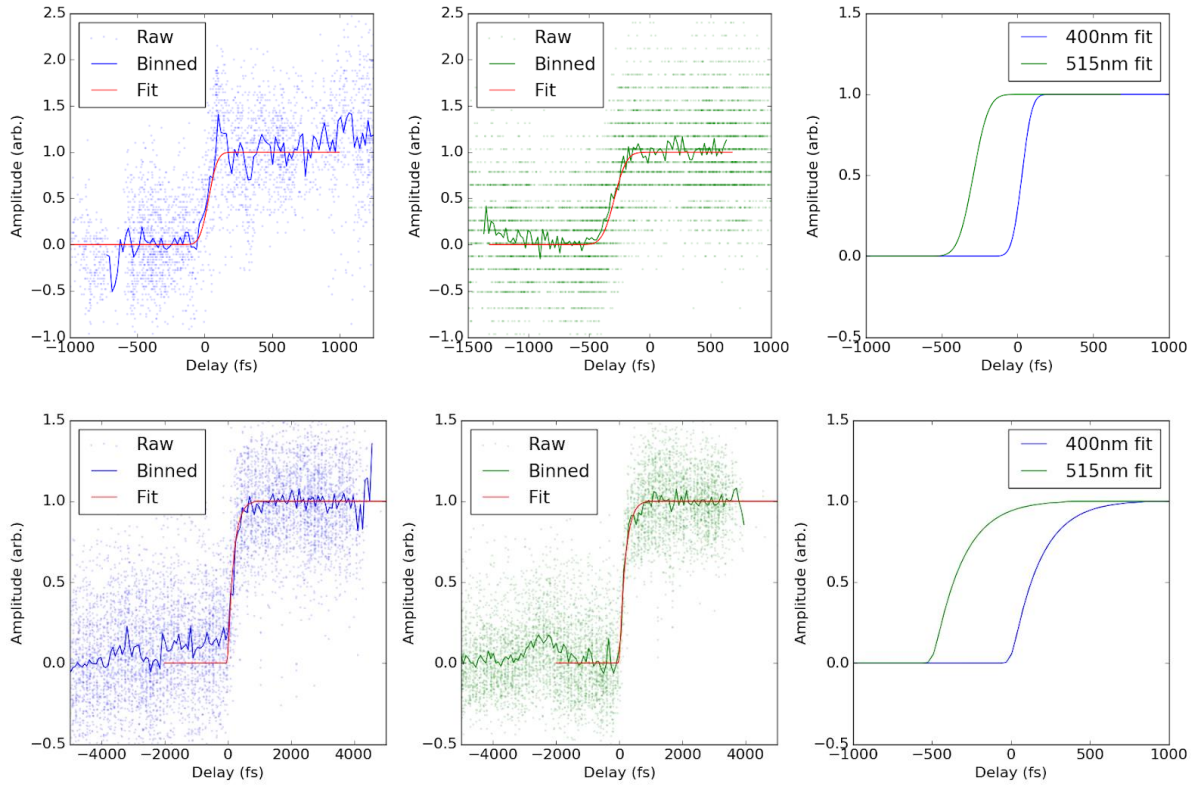


Figure 4.3: Cross-correlation time zero determination at LCLS (top row) and SACLA (bottom row). In the 400 nm pump (left column) and 515 nm dump (middle column) pulses, each show the raw photodiode signal after timing tool correction of the XFEL jitter was applied, binned trace and fitted sigmoid. The fit was generated using the methods found in [114]. The comparison of the fits (right column) shows the pump-dump delay. Note: the order of the optical pulses appears counter-intuitive (i.e the 515 nm pulse appears first) in the cross-correlation traces as in this measurement x-ray pulse is the pump and the optical pulse is the probe as opposed to the TR-SFX experiment when the X-rays are the probe hence the temporal order of the optical pulses is effectively reversed when scanning the same optical delay stage. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

4.2.4 Timing

At LCLS and SACLA a delay stage for each the 400 nm and 515 nm allowed adjustment of the average laser arrival time with respect to the X-ray pulses. Due to the spontaneous nature of XFEL SASE generation, there is a jitter from 0.1 - 1 ps [64, 114], in arrival time of the X-ray photons relative to the optical laser. Although the Ti:Sapphire amplifier is synchronised to the XFEL using a MHz radio frequency, the timing correlation of individual X-ray pulses needs to be determined. At PAL-XFEL, given the very low jitter (25 fs rms) and the longer delay times that were collected (1 ps, 100 µs and negative delays), it was not necessary to bin the data collected.

At LCLS and SACLA shot-to-shot timing measurements were made. Upstream of the experiment end-station, beam-splitters were used to take a small portion of the fundamental laser beam (800

nm) and small (>1%) portion of the X-rays. A timing tool either spatially [281] or spectrally [282] (depending on the facility) encoded the relative arrival time of the X-rays with respect to the laser pulses. To determine the precise delay between the optical pulses and the X-rays in the interaction region, a cross-correlation was made on a Ce:YAG screen. Absorbed X-ray pulses generate large charge carrier densities which modify the dielectric function of the semiconductor [283]. This leads to a transient increase in reflectively in the visible region [283] which changes the intensity of the of the optical probe pulse when it is temporally (and spatially) overlapped with the X-rays. This change in transmission/reflection can then be measured by a photodiode [281, 282]. Depending on the jitter of the facility a cross-correlation can accurately set the average probe delay to around 500 fs – 1 ps [114]. To accurately set the pump-dump delays each X-ray shot in the cross-correlation must be binned based on the X-ray pulse jitter within the cross-correlation scan [38, 114]. The cross-correlation and corrected/binned data are shown for LCLS and SACLA experiments in figure 4.3. For sub 1 ps pump-probe delays the output of the timing tool was used to sort and bin each image based on the measured X-ray-optical delay. This allowed the separate processing of data in 150 fs intervals from 0 – 1 ps if data collection redundancy allowed.

4.2.5 Serial Femtosecond Crystallography Data Merging

The software packages Cheetah [145] and CrystFEL [102] were used to analyse TR-SFX data. Cheetah was used as a 'hit' finder, to identify images that contained diffraction information that could be used for indexing, i.e. images where the X-ray beam hit a crystal. Cheetah was run as close to real-time as possible to give a semi-live hit-rate (percentage of X-ray shots leading to indexable images). This was used to optimise sample injection to no higher than 50-60 % by adjustment of flow rate, position or microcrystal slurry condition (i.e. crystal density, PEG conditions or protein concentration) during the experiment. This was important to avoid adjacent crystals from shadowing one-another from excitation laser, polluting merged datasets with un-pumped data. Additionally, if the hit rate was too high it was found the indexing rate (comparison of the number of hits successfully indexed to the number of hits) would also fall thereby decreasing the overall data rate.

Images that were identified as hits (figure 4.4) were passed to CrystFEL for indexing using a unit cell previously determined from MX experiments (chapter 3). Different experiments had different indexing algorithms available as specified below. Initially, pre-illuminated data was obtained to refine the geometry of the detector (where appropriate), the distance from the detector to the interaction region and the relative detector centre to the X-ray beam. The detector distance was refined by evaluating the distribution of the indexed unit cell dimensions. When correct, the unit cell distribution converged and became much narrower (figure 4.5).

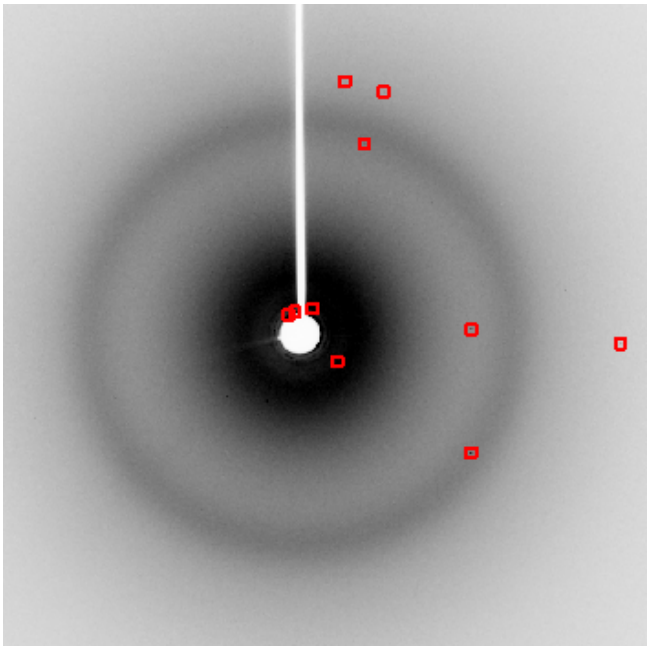


Figure 4.4: Example of peaks found (red boxes) by Cheetah and the applied mask around the beam stop

For each time-bin data were sorted into dark (absence of pump and/or dump laser) and light (presence of pump and/or dump laser) image-sets which were indexed and merged separately. Individual intensities were merged using both the Monte Carlo method of CrystFEL's `process_hkl` function, without refinement for partialities and with the `partialator` function using the 'unity' model with a single scaling iteration. It was found that both methods gave similar overall statistics but the `process_hkl` a gave a higher $CC_{1/2}$ and lower R_{Split}, especially in the highest resolution shells. It was also found that light-induced difference electron density were also stronger in the data merged with `process_hkl`, perhaps due to higher quality high-resolution structure factors. As a result, merging using only the `process_hkl` method was preferred and will be used unless stated otherwise in the following sections.

Data Merging: LCLS

Every 20-40 runs dark scans, in the absence of X-rays were taken to determine which pixels from the CSPAD detectors had bad/damaged pixels. These were then masked out before indexing algorithms were run. This was also performed after if any ice deposition occurred in the CXI vacuum chamber to account for any damage due to high-intensity diffraction from the accumulation. Nearly every run was manually masked before indexing and integration to eliminate shadowing from the nozzle on the images detected. Spots inside the nozzle shadow were excluded as the reported intensities may vary from the actual diffracted intensities and the determination of the final structure. Using CrystFEL, two indexing algorithms, DIRAX [151] and MOSFLM [284] were compared for performance in indexing the images, with the previously defined unit cell being fed to both. To refine and optimise the CSPAD metrology, the dark images were indexed first. The detector distance was refined by varying the `clen` variable and

reindexing. Steps of size 0.1, 0.05 and finally 0.01 and 0.001 mm were made to optimise the indexing results. The optimised geometry provided single narrow distributions with standard deviations of <0.5 Å for lattice axes and around 1° for angles, while the unoptimised geometry showed much larger unit cell distribution widths (>1 Å) and showed a bimodal distribution for the lattice c-axis. Performing an L-test gave a very high twinning fraction (16-17 %) when using the unoptimised geometry. Once the geometry was optimised with the `clen` variable and detector centre fixed this fell (to 6-7 %). The L-test fraction still remained higher than anticipated, however, given that suitable R-factors were obtained in refinement and merging, and a single-set of narrow lattice distributions it was not concerning.

Although the detector was moved and replaced on several occasions, it was not found necessary to individually optimise the detector distance for each run, as acceptable cell distributions and indexing results for the entire dataset were provided by a single optimised distance. It was also found that indexing results were relatively insensitive to indexing algorithm parameters such as integration radii or minimum criteria for SNR. An electronic gain-mask was used for a majority of runs to prevent low-resolution pixel saturation when acquiring high-resolution diffraction data. A gain mask, containing the gain setting of each pixel was applied to the images to correct the intensities across the detector before they were merged using `process_hkl`, as described above.

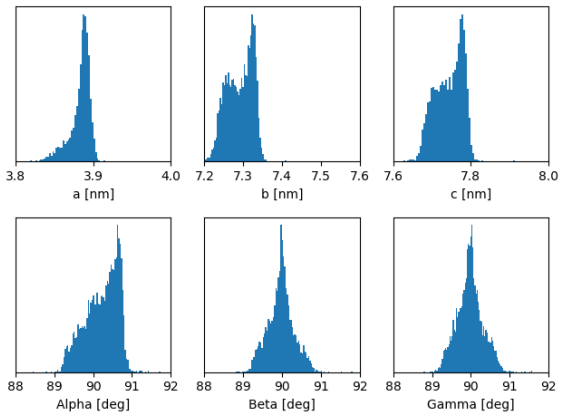
Data Merging: SACLA

The SACLA data was analysed using very similar steps to the LCLS data. Although there were some differences in the experimental set up which changed the data processing pipeline. For example, at LCLS, the GDVN nozzle caused significant shrouding on the upper half of the CXI detector panel. As the droplet injector was more stable than the GDVN nozzle, the distance between the X-ray focus and tip were further. This prevented shrouding meaning the only area that needed to be masked was the beam stop Additionally, the SACLA detector only operated in one gain mode which meant no gain mask was required to get the correct intensities. At SACLA there is also a dedicated real-time pipeline, developed by Nakane *et al.* [285], which gave very quick hit-rate feedback (within 1–2 seconds), meaning the jet could constantly be translated to account for any drift in the XFEL/droplet alignment. This occasionally led to hit-rates up to 80 % which could increase the number of multiple microcrystal lattices being indexed from single diffraction snap-shots. For peakfinding in CrystFEL, two indexing algorithms were tested at SACLA: TakeTwo [155] and XGANDALF [152].It was found that XGANDALF gave the best results so was taken forward.

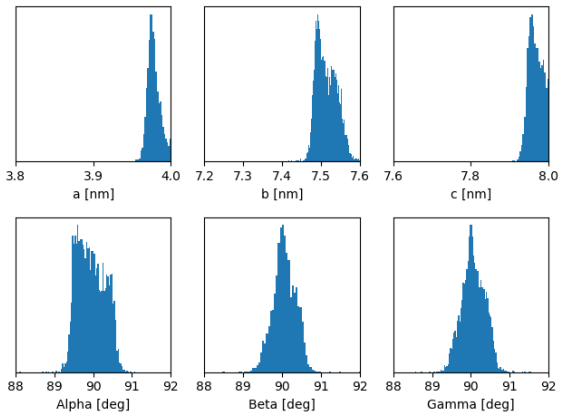
Data Merging: PAL

Occurring during September 2020, this experiment was mandatorily remote. Cheetah was not used for hit rate optimisation unlike experiments at SACLA and LCLS but it was used for estimating overall

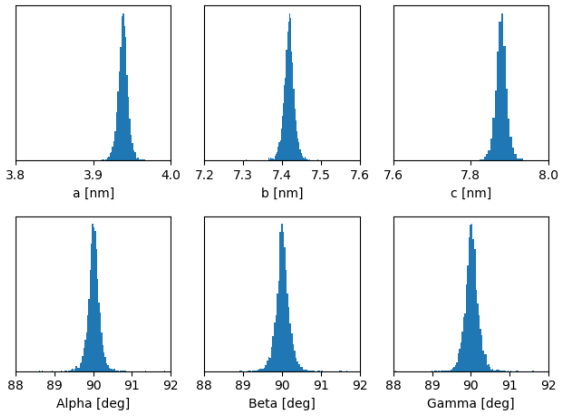
collection yields during the experiment, for peak finding and to apply the beam-stop mask. Initially, hit-finding parameters were optimised on dark-condition images and refined based on indexing results from CrystFEL. CrystFEL's internal hit and peak finder, `peakfinder8` [286] was tested in parallel with Cheetah but was found to give lower quality merged data, perhaps due to misidentification of noise as Bragg peaks. Unlike the detectors at SACLA and LCLS, the PAL experiment used a commercial Rayonix MX225-HS detector which has a geometry correction pre-applied and automatically subtracted a dark image meaning it was not necessary to manually check for hot/dark pixels in Cheetah. It also meant it was only necessary to refine the beam-centre and detector distance but not the geometry of individual detector panels. Two indexing programs: MOSFLM [284] and XGANDALF [152] were compared for indexing success and statistics. It was found that XGANDALF gave the best indexing results so was used in the analysis going forward.



(a) Distance = 141.0 mm



(b) Distance = 144.5 mm



(c) Distance = 142.9 mm

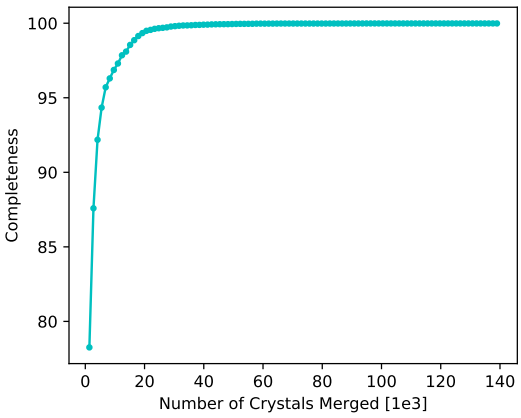
Figure 4.5: Detector distance refinement: Showing unit cell distribution changes with different detector lengths. Shown is a histogram of the unit cell lengths a, b and c and angles α, β and γ. The final refined dimensions were 39.36±0.06 Å, 74.15±0.13 Å, 78.70±0.13 Å and angles of 90±0.13°, 90±0.15°, 90.02±0.19°for the PAL experiment. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

The detector distance was refined in steps of 0.5, 0.05 then 0.01 mm. Figure 4.5 shows the distribution of unit cells tightening as the detector distance converges. The detector centre was checked each day and found to move by 0.2 mm on the second last day of data collection. The detector distance was

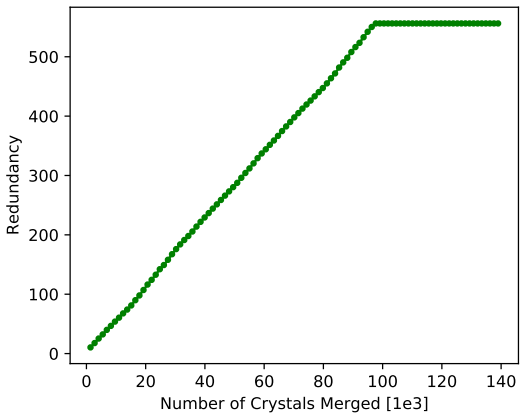
also re-refined at this point but found to be stable. Detector saturation levels were checked and the majority of the X-ray diffraction spots were within the dynamic range. There was no electronic gain mask applied to the detector during this experiment, unlike LCLS.

Occasionally a build-up of sample would occur around the injection syringe. This led to increased background scatter which may lead to lower accuracy in the measured spot intensities. Each run was checked for this by plotting the average pixel count of all images as a histogram and ignoring images with average pixel counts over 1000. However, It was found this had no observable effect on data quality so was not implemented through the experiment.

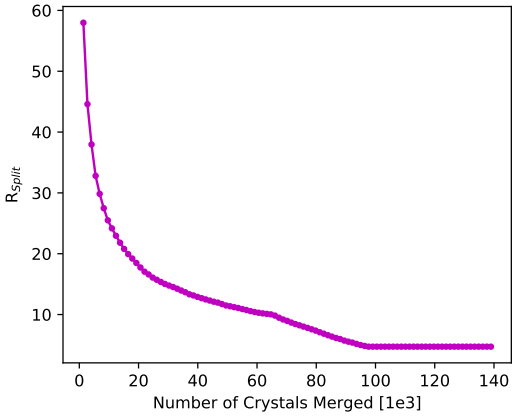
As of September 2020, there were no published crystallography structures from the NCI beamline at PAL-XFEL This meant it was not clear how many indexed patterns were required for convergence of merging statistics. This was tested by merging random laser-off images across the four days of experiments with an increasing cumulative numbers of images. Up to 140,000 images were merged in 140 steps. The results show that signal to noise always increases with high redundancy, whilst other merging statistics such as $CC_{1/2}$ plateau after merging of around 20,000 diffraction patterns (figure 4.6).



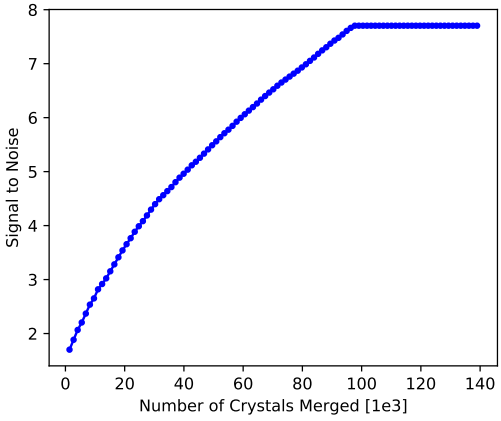
(a) Completeness



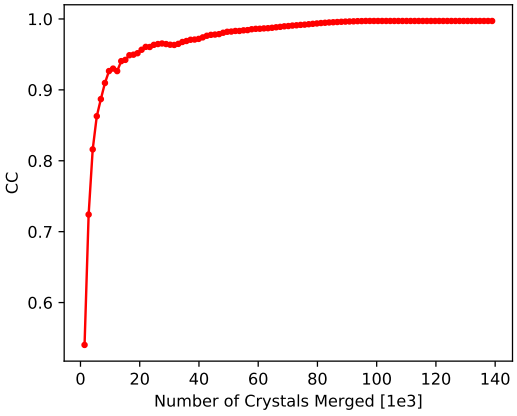
(b) Redundancy



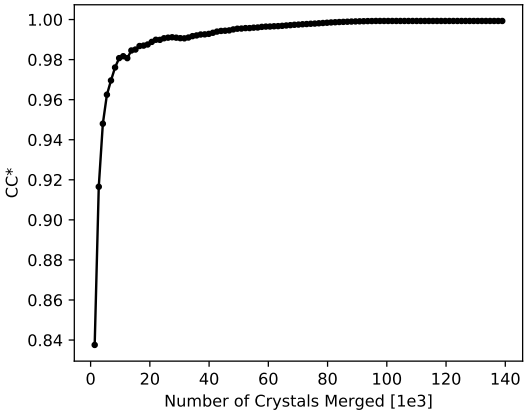
(c) R_{Split}



(d) Signal to Noise



(e) $CC_{1/2}$



(f) CC*

Figure 4.6: Laser-off data from PAL-XFEL: crystallographic measures of data quality plotted a function of number of merged crystals. A constant high-resolution limit of 1.45 Å was applied.

4.3 Data Processing of Merged TR-SFX Data

Over the three TR-SFX experiments a total of 32 independent pump-probe and pump-dump-probe crystallographic datasets were collected (table 4.2). These experiments aimed to firstly, observe the isomerisation pathway with high structural detail in the pump-probe data and secondly, to observe the structural effects induced by the dump pulse, both with and without the preceding pump. This required a way of accurately determining the 'light minus dark' electron density differences in real space so that transient model coordinates could be fitted. The processing steps are summarised in figure 4.7. Firstly, merged intensities (I_{hkl}) are converted to structure factors (F_{hkl}), the dark and light structure factors are then scaled the differences calculated (ΔF_{hkl}), difference electron density ($\Delta\rho$) is then generated from the Fourier transform of the ΔF_{hkl} with model phases (Φ). Intermediate-structures can be fitted using extrapolated structure factors (F_{Ext})and corresponding extrapolated electron density ρ_{Ext}. This section will justify the choices for each processing step and detail the data analysis pipeline that was established for the TR-SFX data.

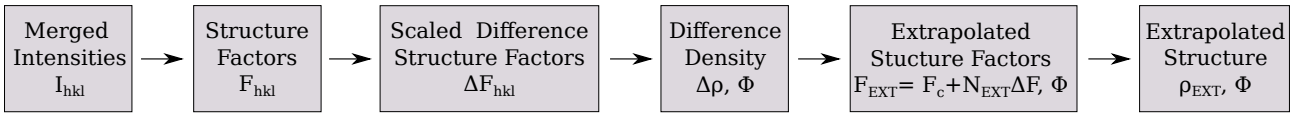


Figure 4.7: Data analysis pipeline for TR-SFX data as described in the main text.

Structure Factor Calculation from Intensities

As structure factors are real numbers which are the square-root of the measured intensities, by definition they cannot be negative. However, modern detectors measure the strength of a reflection by taking the pixel intensity and subtracting a background value. At higher resolutions, the reflection intensities become small relative to the background levels. This leads to an inevitable random error in the determination of reflection intensities and, in particular, can lead to the observation of unphysical negative intensities for weak reflections (when $I \sim \sigma_I$). To correct this either all negative intensities can be ignored or, French and Wilson scaling can be applied [172]. This corrects small and negative intensities based on a Bayesian statistical approach. Prior knowledge of the expected distribution of X-ray intensities is known [173] meaning reflection intensities with a magnitude close to 0 can be scaled appropriately using a fitted B-factor. The results of this can be evaluated using a Wilson plot to see if the calculated structure factor moduli follow the expected distribution. However, the expected intensities are calculated by assuming the relative distances and orientation of atoms in a crystal were random [173]. As protein crystals are well ordered (on the unit cell length) these assumptions hold for short distances but are unrealistic on longer length scales. This leads to the expectation that structure factor moduli will deviate from ideal Wilson plot behaviour at low-resolution limits.

In practice, French and Wilson scaling is implemented in the CCP4i package, Truncate [238] and

in PHENIX [176] within the reflection conversion function's `massage` option. Both implementations were tested on TR-SFX data and it was found that Truncate underestimated the scale factor for higher-resolution shells. The reasoning for this is unclear. Perhaps this is due to the anisotropic nature of the data or an artefact when merging reflections in CrystFEL. As the exact implementation of scaling in Truncate is not clear it was not possible to troubleshoot this issue. Instead, PHENIX's `massage` option was used. This gave correctly scaled data and so was taken forward. The source code for PHENIX shows the implementation used. The `massage` option takes weak/negative intensity measurements (I_{hkl}), their standard deviation, (σ_I) and treats them with equation 4.1 to calculate, positive structures factors, F_{Pos}:

$$F_{Pos} = \sqrt{\frac{I_{hkl} + \sqrt{I_{hkl}^2 + 2\sigma_I^2}}{2}} \tag{4.1}$$

This implementation is similar to that of French and Wilson scaling [172]. The data analysed from SFX were anisotropic and so the boosting of weak intensities was beneficial to scaling and refinement of the structure factors (evidenced by improved R-factors and clearer difference map features when using the massage option).

Difference Structure Factor Calculation: Weighting and Scaling

The calculation of difference structure factors requires scaling of reflections (section 1.4.3) followed by subtraction. Options for scaling are varied, and there are a range of implementations amongst CCP4i (`Kraut`, `overall` and `anisotropic/isotropic`) and PHENIX (`multiscale`, `overall`). After scaling, differences can be directly calculated or weighted by standard deviation which accounts for the error in each reflection measurement and weights it accordingly (section 1.4.3).

The TR-SFX data is high resolution (1.38-1.44 Å) and anisotropic meaning it was inappropriate to scale the data using an overall or isotropic scale factor. Instead, we choose to compare the differences between `Kraut`, `anisotropic` and `multiscale` methods. Kraut scaling was developed to improve the phasing of crystal structures with heavy atom phasing techniques [287]. The scaling of data with a Kraut scale (as implemented in FSHCAL/SCALE IT [177]) ideally used unmerged acentric reflections to account for any anomalous signal (i.e. processing the data in P1 space group). However, to improve the accuracy of the SFX data merging we assumed the crystal data obeyed Friedel's and merged data in the highest symmetry group possible ($P2_12_12_1$). As a result, Kraut scaling was not appropriate, although the calculated scale factors are shown in figure 4.8.

In SCALE IT, anisotropic scaling is implemented to minimise to average structural factor residuals between the dark and light structure factor sets. A scale factor (K) of:

$$K = C\exp(-(B_h h + B_k k + B_l l))$$

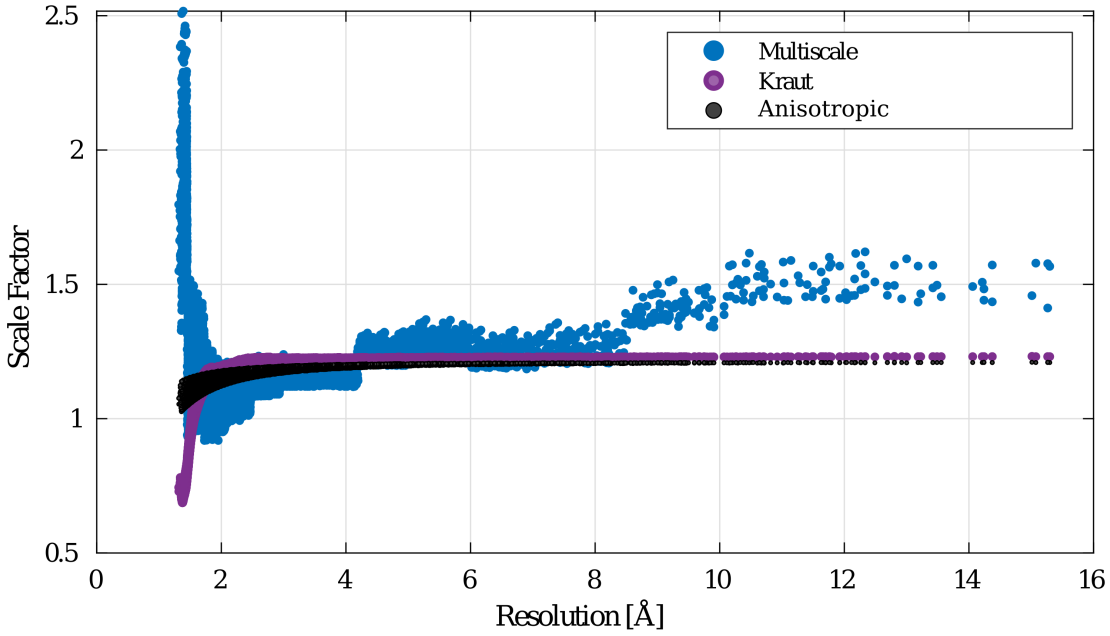
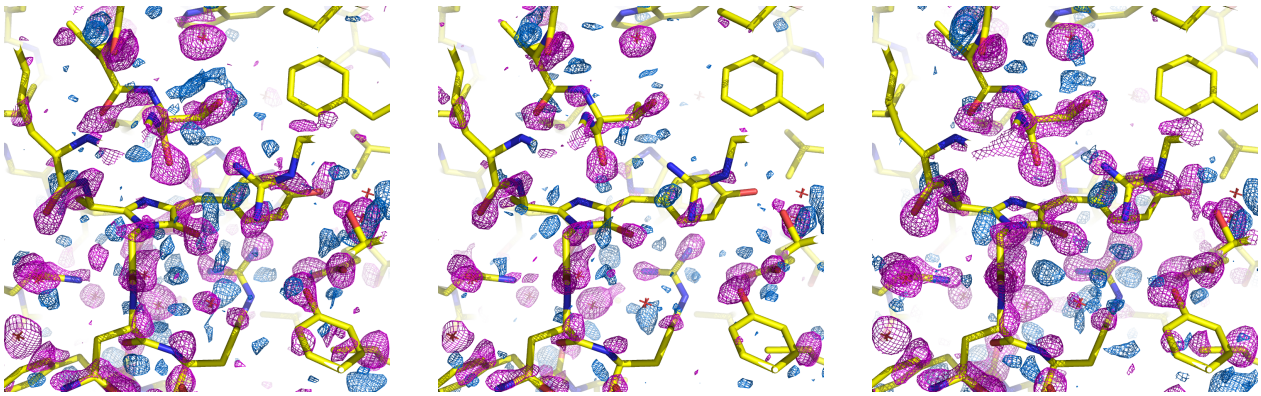


Figure 4.8: Scale factors for the LCLS dark dataset calculated using three different methods as described in the main text.

is used in the case of an orthorhombic space group. Where C is an overall scale amplitude and B_h, B_l, B_l are the B-factor scales implemented on the reflection indices, h,k,l. The calculated scale factor is plotted as a function of resolution in figure 4.8. One advantage of using SCALE IT over PHENIX was that more than two datasets could be scaled together in multiple scaling steps. This meant that in SCALE IT, calculated structure factors from the program SFALL [177] can be provided to use as an 'absolute' scale which data can be scaled against. This ensures normalised scaling when generating extrapolated structure factors (as explained in the next section).

PHENIX's scaling implementation is known as multiscaling. In the multiscale method, common reflections are scaled on a normalised absolute scale compared to the supplied model however, the reflections are binned by resolution which leads to the step-like scaling with resolution as seen in figure 4.8 [288]. There is no physical basis for why scaling by resolution bins would be advantageous. However, it was found that difference maps scaled using multiscale gave very sharp difference features with almost no erroneous signal in the solvent filled space between protein monomers. To calculate Q-weighted DED using the PHENIX multiscaling method edits were made to the PHENIX source code to write out the intermediate structure factors after scaling. Normally difference maps in PHENIX are calculated using the isomorphous difference function which only outputs the final difference map and phases in one file.

The results of the SCALE IT anisotropic-scaled Q-weighted and the PHENIX-multiscaled direct and Q-weighted DED maps are seen in figure 4.9. The unweighted multiscale differences show the strongest DED signals followed by the anisomorphous Q-weighted and then the multiscaled Q-weighted. Overall



(a) Anisomorphous Q-Weighted **(b)** Multiscaled Q-Weighted **(c)** Multiscale Unweighted

Figure 4.9: Difference electron density maps calculated using different methods of scaling and weighting as detailed in the main text.

the SCALE IT anisomorphous-scaled method was chosen for two reasons. Firstly, the ability to scale the dark and light data on a single absolute scale was essential for generation of extrapolated maps and comparison of different illumination conditions. It was not possible to this in PHENIX. Secondly, the anisomorphous scaling and the Q-weighting methods have the strongest physical foundations in terms of error propagation (with the Q-weighting) and scaling (unlike the resolution-binned multiscale method). This gave the most confidence that the final calculated difference density was due to physical differences in the structure and not an artefact of data processing

Difference Structure Factors to Difference Electron Density and Extrapolated Maps

The difference electron density is calculated from the structure factors difference, ΔF. The program FFT within CCP4i [177] is used to calculate the real-space difference electron density maps from the structure factor differences. The importance of choice in phase is discussed in section 4.3. Here we use the dark-model phases for the calculation of the final map. The success of scaling can be evaluated and the most reliable method chosen with examination of the DED. In particular, scaling was judged on the shape of the difference density features and the significance/frequency of signals observed in the solvent space between protein molecules.

The photoproduct yields of rsKiiro are expected to be around 20%. This means signals in the final DED will be due to partially occupied states. To extract small structural changes in the illuminated data, extrapolated electron density maps were generated and extrapolated coordinates then refined to these maps. Difference structure factors, ΔF were added in a linear combination in multiples of N_{EXT} to the calculated dark structure (F_{Dark}) to generate extrapolated structure factors, F_{EXT}. The negative electron density around areas of interest in the protein was then integrated. The point where negative density started to build up was used to determine a characteristic N_{EXT} value which was then used to generate the final extrapolated structure factors. Coordinates were refined against the extrapolated

structure factors (section 1.4.3) using 5 cycles of rigid body refinement in REFMAC [237] and further real-space refinement using coot [289]. The population transfer (PT) as a fraction can be approximated from the N_{EXT} value as $PT = 2/N_{EXT}$. Here, population transfer refers to the fractional occupancy of the light-induced state.

The sphere of integration radius and centre can be altered to include/exclude DED features of interest. In rsKiiro a majority of difference signals were located in close proximity to the chromophore: on the α helix and around the chromophore pocket. We, therefore, choose an integration sphere centred on the methyl bridge of the chromophore at a radius of 7 Å. This sphere then enclosed a majority of the key difference density features. It was found that minor alterations to the sphere location or radius did not significantly change the fitted extrapolated value. If the sphere was placed at an irrelevant position for example, far away from the chromophore, then it was found that the extrapolated calculation was unphysical as the calculated electron extrapolated electron density (ρ_{Ext}) would use too higher value of N_{Ext}. This would create noisy electron density which was uninterpretable.

Further Consideration of Phases

One of the most important choices in the evaluation of crystallography electron density is the phases used. Throughout this analysis we have used the dark-state phases for the structure factor differences. This was justified (in section 1.4.3) because light-induced structural differences in rsKiiro are mostly localised to the chromophore region so it is assumed the light and dark structure factors are very close in phase. However, with the generation of the extrapolated coordinates it is possible to generate a set of 'light' structure factor phases (Φ_Δ), calculated from a mixed occupancy model of dark and extrapolated coordinates. Figure 4.10 shows an exaggerated Argand diagram of a light induced structure factor difference, ΔF, generated from the light (F_L) and dark (F_D) structure factors. Clearly in this case, the phase of the difference structure factor is significantly different from the light and dark structure factors. This means the magnitude of ΔF systematically underestimates the full difference structure factor vector.

To test the effects of incorporating a light-phase set, the Q-weighted DED for a merged laser-on bin was phased using a dark:light-state model of 0.8:0.2 occupancy ratio. It was found there was almost no difference in the peak difference density values, positions or the overall average phases for the new maps. This is most likely because the photo-population is small and localised to the chromophore. This means the light-state structure has a very similar structure and therefore very similar phases. Furthermore, extrapolated maps generate new structure factors based on dark-state calculated structure factors and the observed structure factor differences. By fitting an extrapolated model and then re-using this model for phases we would essentially be biasing the maps to the light state structure. In the case where a state is fully occupied (i.e. 100% population transfer to a light induced state) accurate light-state model

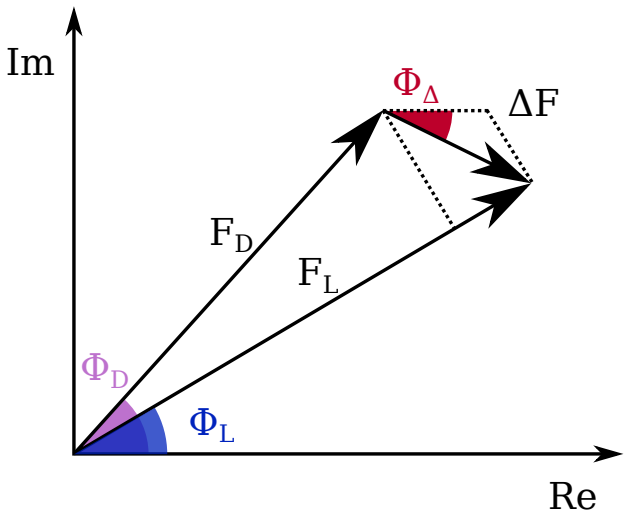


Figure 4.10: Complex plane of the structure factor triangle for light induced structural difference calculation.

can be determined and so can accurate phases. In the case of TR-SSX the low population transfers and synthesised maps were overall deemed less reliable than the use of dark phases. Therefore, it was decided not to calculate light-state phases. Instead, the DED were phased with the purely dark-state model.

Final Processing Pipeline

Bash scripts processed each light and dark bin to generate the Q-weighted DED maps based on the Bayesian statistical analysis by Ursby *et al.* [174, 175] (section 1.4.3) which were then interpreted by generating extrapolated electron density for which 'light' coordinates were refined too. The PHENIX reflection conversion function, using the `massage` option, converted the intensities to structure factors [176]. Calculated structure factors from SFALL (CCP4i [177]) using the ground-state refinement model were combined with observed structure factors into one file using CAD (CCP4i [177]). Both the light and dark structure factors were scaled on an absolute scale using `anisotropic` scaling in SCALE IT (CCP4i [177]). A relevant low and high-resolution limit was applied during the scaling step. This was chosen depending on the experiment (1.35, 1.44 and 1.42 Å at LCLS, SACLA and PAL-XFEL respectively). The low-resolution limit was set between 20 - 30 Å. Structure factors differences were Q-weighted as described. The FFT program (CCP4i [177]) was used to calculate the DED from the structure factor differences. The grid spacing was left unaltered. Extrapolated structure factors were calculated as: $F_{Ext} = F_{D,c} + N_{Ext}\Delta F$. FFT, using the `simple` option was then used to generate the extrapolated electron density. Negative electron density above 3 rms level and a 7 Å radius around the chromophore was integrated using published FORTRAN code [241]. The point where negative density started to build-up was used to determine a characteristic N_{Ext} value which was then used to generate the final set of extrapolated structure factors. Coordinates were refined against the extrapolated structure factors

using 5 cycles of rigid body refinement in REFMAC [237] and further real-space refinement using coot [289]. The population transfer was approximated from the N_{Ext} value as $PT = 200/N_{Ext}$.

4.4 Experiment Results and Discussion

Overall, a total of 682,650 indexed crystal diffraction patterns were merged into 32 pump-probe/pump-dump-probe delays (-100 ps – 100 µs) at three different experimental facilities. A summary of collected conditions is shown in table 4.2. The limiting resolution varied depending on the experimental setup but was better than 1.45 Å for the dark datasets of all experiments. The crystallography merging and refinement statistics are shown in tables A.5 – A.12. In total we have measured pump-probe (PP) crystallographic signals from 250 fs – 100 µs and pump-dump-probe (PDP) signals from 250 fs – 100 ps. We observe unambiguous DED with high signal to noise ratios which are reproduced at two separate experimental facilities at (SACLA and LCLS). The sub 1 ps PP data shows the onset of ultrafast oscillation which are enhanced in the PDP measurements. This allows us to assign the signals to the observation of ground-state vibrational coherence as opposed to the observation of an excited-state intermediate state.

Experiment	Conditions			
	Pump Probe	Pump Dump Probe	Dump Only	Dark
LCLS	0 – 1 ps, 250 fs, 400 fs, 550 fs, 700 fs, 850 fs, 1100 fs, negative	0 – 1 ps, 250 fs, 400 fs, 550 fs, 700 fs, 850 fs, 1100 fs, negative	1 ps, negative	Dark, No overlap
SACLA	550 fs, 3 ps, 100 ps	550 fs, 3 ps, 100 ps	-	Dark
PAL 2020	1 ps, 100 us, negative	-	-	Dark

Table 4.2: Summary of collected TR-SSX delays and conditions.

4.4.1 LCLS

At LCLS, sub 1 ps data was collected for both the PP and the PDP conditions which was further binned into 150 fs intervals using the timing tool (section 4.2.4). Additionally, a 1 ps dump-only, negative-delay and dark datasets were also collected (table 4.2). A dataset of all 0 – 1 ps delays for PP and PDP conditions were merged as separate dataset for calculation of extrapolated maps. The higher redundancy of the 0 – 1 ps meant the extrapolated density was less noisy compared to the binned extrapolated maps. The pump-dump delay was retrieved to be 350 fs by cross-correlation with the X-rays. The pulse width of the fundamental and dump pulses were resolved using the SHG-FROG to be 100 and 71 fs respectively (table 4.1). A successful measurement of the 400 nm pulse was not made. Instead, the pulse length was estimated from the SHG-FROG of fundamental, 800 nm pulse

duration of 100 fs. It was assumed that the 400 nm SHG with the same 100 µm BBO would be a similar duration (i.e. ∼100 fs). The dark data was limited in resolution to 1.37 Å as determined from the Fourier shell coefficient (FSC) in figure A.4. Refinement of the dark and negative merged data clearly shows a trans-state with structural features of the cryo-MX data (chapter 3) and a single chromophore conformation.

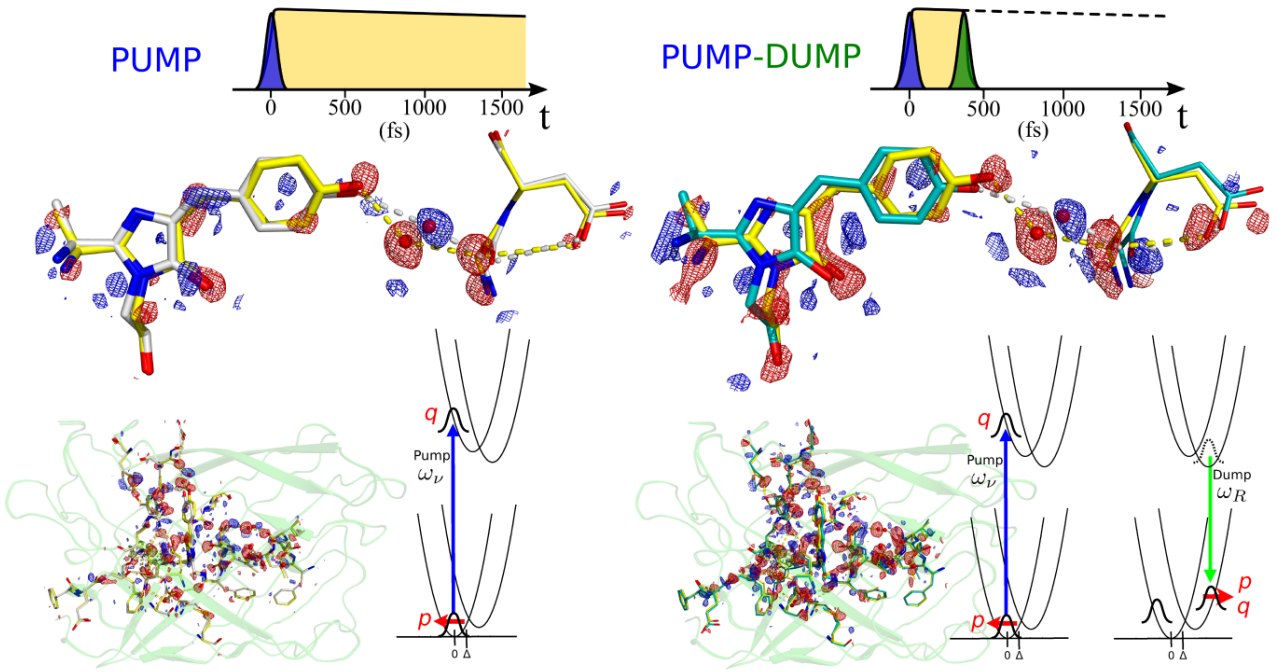


Figure 4.11: Femtosecond time resolved Pump-Probe (left) and Pump-Dump-Probe (right) TR-SFX experiment of rsKiiro off state with a trans neutral chromophore. Time-resolved data for delays between 250 fs and 1 ps were merged and Q-weighted Fo-Fo maps are shown at 3σ level and 1.3 Å resolution. Coordinates for ground state (yellow), Pump-Probe data (white, left) and Pump-Dump data (cyan, right) were refined from extrapolated structure factors, and occupancy refinement. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

The pump-probe delays merged between 0.2 - 1 ps show strong difference signals around the chromophore. Shown in figure 4.11 is the difference density and extrapolated coordinates of the pump-probe and pump-dump data. Signals up to 8.5 σ level are observed on the water hydrogen-bonded to the chromophore. Extrapolated structure factors were calculated with an N_{Ext} of 13.5 (figure A.5) indicating a population transfer of around 15 %. Refined extrapolated coordinates (white in figure 4.11) show a small movement of 69 pm the chromophore phenolic oxygen towards a planar geometry. Additionally, and very clearly represented in the difference density, is the movement of the hydrogen-bonded waters between the phenol and the Gly155 shown. Strong signals are seen emanating down the central α helix (figure 4.11 bottom left). This is clearly seen in the 1-D difference maps with signals smeared over the residues closest to the chromophore (figure 4.12).

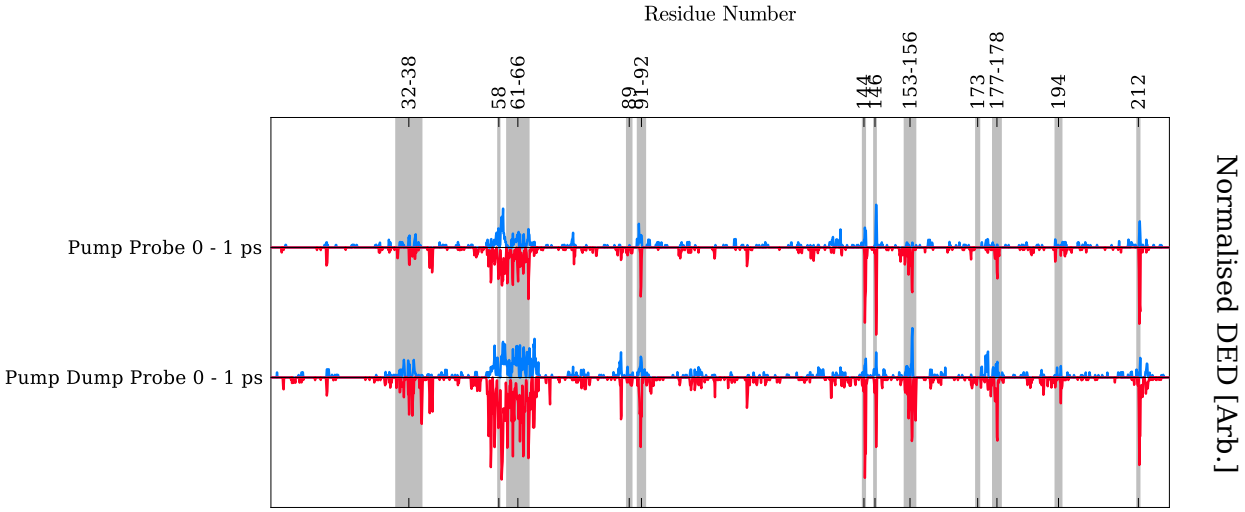


Figure 4.12: 1-Dimensional representation of integrated difference electron density within 2 Å at 3σ level along each atom of the protein chain, all plotted on the same scale (normalised difference electron density). Highlighted on the top axis are residue numbers of interest around the chromophore as discussed in the main text. Code to produced this figure is based on the published tool by Wickstrand *et al.* [250].

The pump-dump data shows signals at very similar locations but with significantly strong stronger amplitudes (figure 4.11 and 4.12). An extrapolated value of 12.8 was obtained, equivalent to a population transfer of around 16 %. The refined extrapolated coordinates show larger movements but in the same directions as the pump-probe data. The change of the hydrogen-bonded water between the chromophore and water positions are almost identical but features on the chromophore are exaggerated. For example the imidazoline oxygen shifts by around 0.2 pm in the PP coordinates but around 0.5 pm in the PDP coordinates.

DED maps were also calculated for each 150 fs bin from the 0 – 1 ps data. The signals for these bins were located in the approximately the same positions as the full 0 – 1 ps merged bin. The peak-heights of key DED features around the chromophore are plotted for each time-point in figure 4.13. These plots reveal clear oscillations in the strength DED peaks with delay. In particular, the strongest signals can be seen on the hydrogen bonded oxygen (positions 1 and 2 in figure 4.13). These signals are shown in detail in figures A.1 and A.1. To test if these features were due to changes in data quality, the R_{Split}, SNR and CC* were calculated for the highest resolution shell (1.55-1.50 Å) of each time bin. These were then plotted as a function of delay (figure 4.14). The trend in data quality for each time-bin (figure 4.14) and the difference density peaks (figure 4.13) show no correlation. As a result, the oscillations are assigned to ground-state coherent motion which is generated from the relaxation of the excited state wave-packet after it has undergone Franck-Condon decay [290].

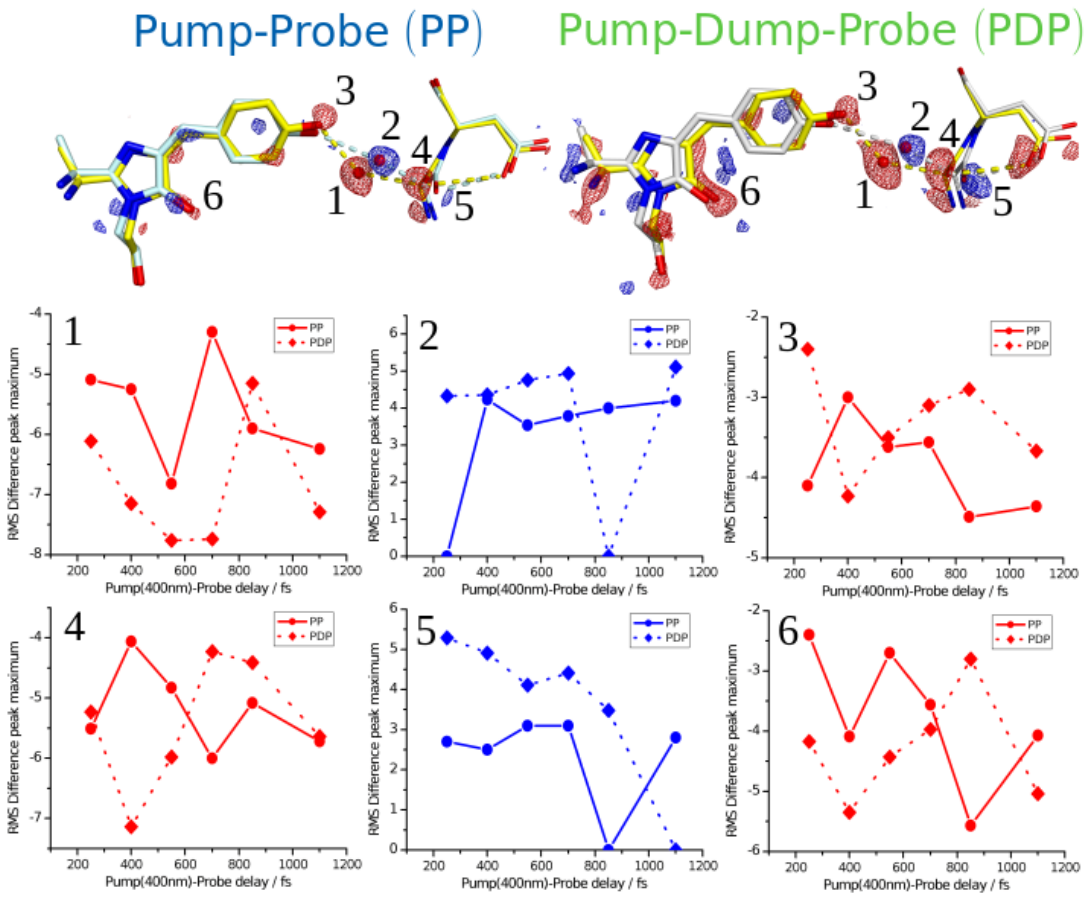


Figure 4.13: Top Panel: Q-weighted difference electron density between the dark and 400 nm 0-1 ps (left) and between dark and 400-515 nm 0-1 ps (right) contoured at $+3\sigma$ RMS (blue) and -3σ RMS (red). Labelled 1-6 are areas of difference signal reported in the graphs below. **Bottom Panel:** Difference density peak values at the locations shown above. Plotted as a solid line (with dots) and a dashed line (with diamonds) for the 400 nm 0-1 ps (pump probe, PP) and 400-515 nm (pump dump probe, PDP) respectively. Shown in blue is values of positive density, and in red negative. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

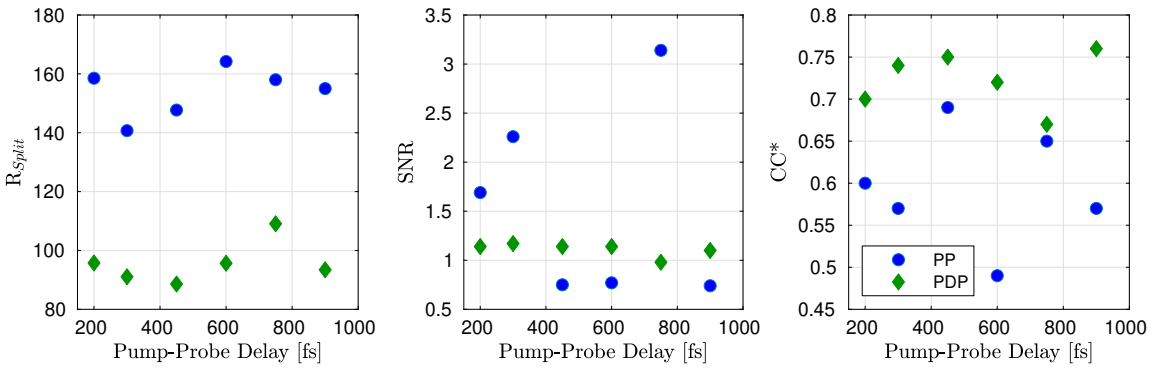


Figure 4.14: Crystallographic merging statistics: R_{Split}, SNR and CC* plotted for the highest resolution bin (1.55-1.50 Å) for each 150 fs time bin. The trends in data quality are not reflected in the oscillations of electron density seen in figure 4.13.

4.4.2 SACLA

At SACLA a 500 fs, 3 ps and 100 ps time points were collected for the PP and PDP conditions (table 4.2). Post-processing of the cross-correlation scans of the optical/X-ray pulses determined a pump/dump delay of 500 fs compared to the 350 fs delay at LCLS. The SHG-FROG showed a shorter than expected 515 nm dump pulse lengths of 52 fs as did the XFROG, with a 400 nm pulse duration of 40 fs (table 4.1). The limiting resolution of the SACLA dark data was measured to be 1.44 Å by FSC (figure A.4). This is slightly lower than at LCLS (1.38 Å) which is most likely due to the lower pulse energies of the XFEL (∼0.5 mJ at SACLA compared to ∼ 1 mJ at LCLS). Full merging statistics are found in tables A.10 and A.11. Other than the notable decrease in resolution, the data quality at SACLA is very similar to that collected at LCLS. The refined dark coordinates again show full conversion, as is consistent with the LCLS dark data and cryo-MX structures (chapter 3).

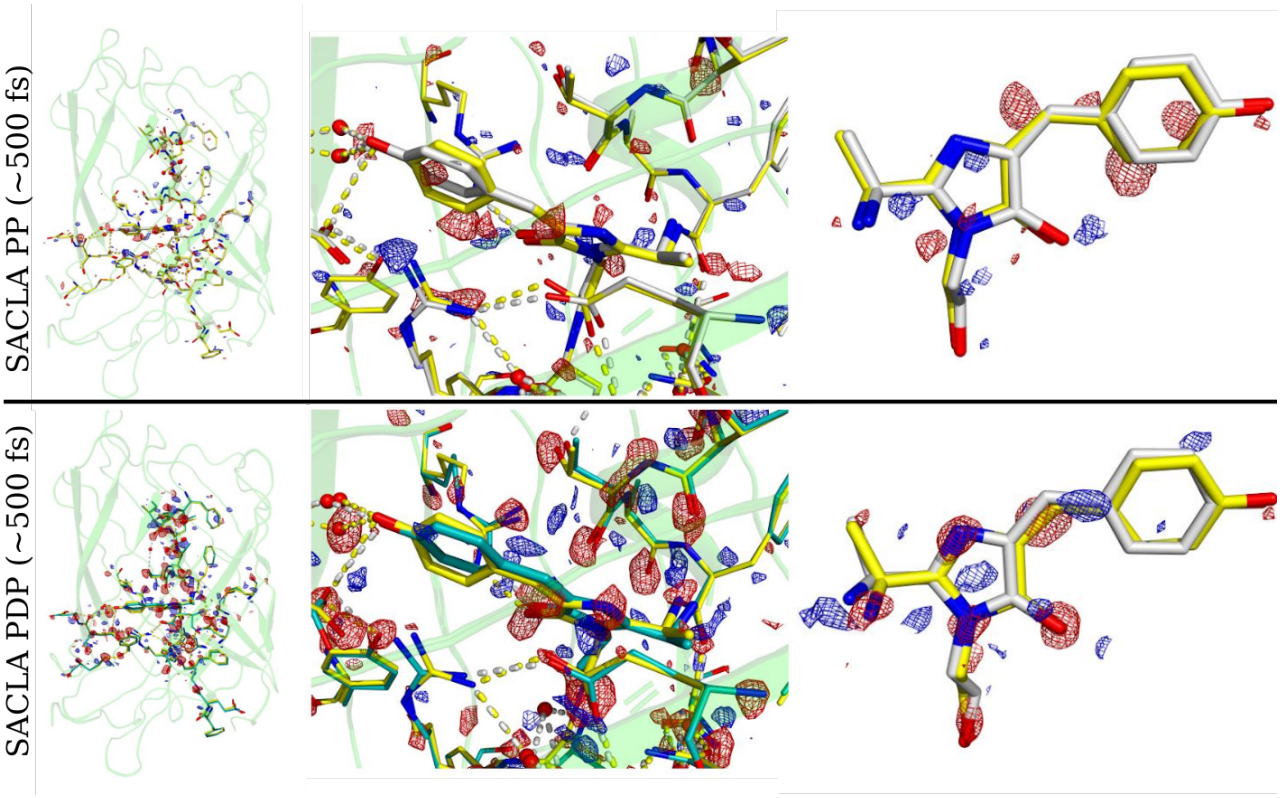
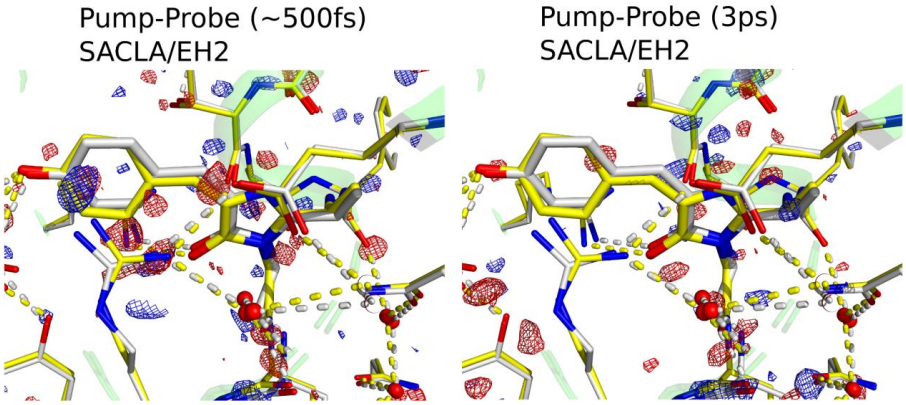


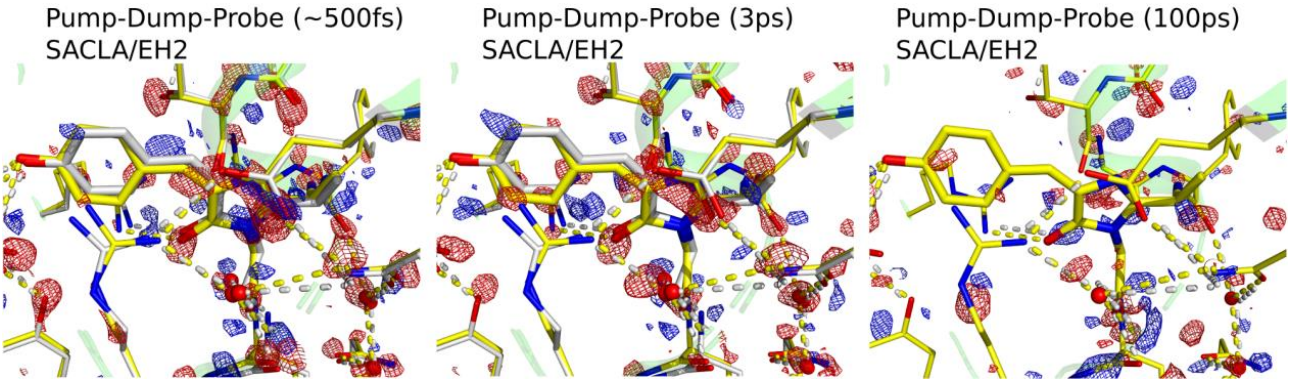
Figure 4.15: Q-weighted difference electron density. **Top:** differences between the pump-probe average delay 500 fs and dark (laser off). **Bottom:** differences between the pump-dump-probe (average delay 500 fs) and dark (laser off). The addition of the dump-pulse 350 fs after the pump, increases the electron density differences amplitudes, and larger shifts of the fitted coordinates to extrapolated structure factors (shown in grey and green), taken as N_{Ext}=18.8 for the pump probe and N_{Ext}=16.9 for pump-dump. The secondary structure (green) of the ground state coordinates (yellow) are shown with electron density contoured at +3σ RMS (blue) and -3σ RMS (red). Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

At SACLA, the sub 1 ps data was collected with lower redundancy than at LCLS so was not split into shorter time-bins. Instead, a single ∼ 500 fs bin was merged and processed for the PP and PDP

conditions. DED and extrapolated structures were calculated for this bin for direct comparison to the LCLS 0 – 1 ps PP and PDP datasets. The PP signals were in very similar locations to the LCLS PP signals (figure 4.15) but were slightly weaker with peak signals of 6.16 RMS compared to 10.32 RMS in the LCLS data. Differences in maximum resolution/data quality can not be excluded as the origin of reduced DED signals, however, the most likely cause is the shorter pump-pulse duration. At LCLS we estimated a ∼100 fs from the SHG-FROG of the fundamental pulse. The pulse length at SACLA was much shorter (40 fs, table 4.1). The shorter pulse duration will give lower population transfer to the excited state/photoproduct and therefore would yield lower DED signals. In support of this, the PP extrapolated value was calculated to be 18.8 (figure A.6) corresponding to a population transfer of 10.6 % which is lower than the LCLS PP population of ∼15 %. The refined extrapolated model also shows a very similar translation of the chromophore as observed in the LCLS measurements (figure 4.16).



(a) A comparison of the average 500 fs delay and 3 ps delay pump-probe data shows a clear decay of Q-weighted difference density features.



(b) A comparison of the average 500 fs delay, 3 ps and 100 ps delay pump-dump-probe data shows a moderate decay of the electron density signals, corresponding to the 50 ps decay time observed for the ground state intermediate induced absorption at 440 nm.

Figure 4.16: Comparison of Q-weighted difference electron density differences of the pump-probe in **(a)** and the pump-dump probe in **(b)** signals SACLA/EH2. Extrapolated structures for each condition are shown in grey, the secondary structure (green) of the ground state coordinates (yellow) are shown with electron density contoured at $+3\sigma$ rms (blue) and -3σ rms (red). Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

The PDP signals at SACLA, are in very similar location to the LCLS PDP (4.15). Signals are increased compared to the SACLA PP, with a maximum DED peak of 9.75 RMS, but are still weaker than the corresponding LCLS PDP 0 – 1 ps (10.46 RMS). This is most likely due to a decreased population transfer as a result of the shorter pump (as seen in PYP [258]) and dump pulses (table 4.1) but could also be due to the longer dump delay which was measured to be 500 fs compared to the 350 fs at LCLS. The extrapolated map calculation fitted a value of 16.9, corresponding to a PT of 12% which is again lower than the LCLS PDP PT of 16% (figure A.6). The refined coordinates also follow the same trend as the LCLS experiment with a similar but more pounced translation of the chromophore when compared to the PP data (figure 4.15).

At SACLA, in addition to the 500 fs delay, we collected later time points of 3 ps for the pump-probe then 3 and 100 ps for the pump-dump conditions (table 4.2). The PP delay of 3 ps (figure 4.16 (a) right) shows a decay in DED signals compared to the 500 fs time-point (figure 4.16). This is fits with the first 7 ps rate constant of the biphasic S_1 excited state decay from IR and visible TA measurements (chapter 6 and C. Hutchison, unpublished). The PDP data at delays of 3 and 100 ps shows clear decay of the light-induced difference signals in the 500 fs condition (figure 4.16). This also follows the decay signals observed in spectroscopic PDP IR and visible TA which show rate constants of 95 and 50 ps respectively.

4.4.3 PAL

At PAL we confirmed the 1 ps time-point behaviour before extending the pump-probe delays up to 100 µs. This experiment was compulsory remote (due to the ongoing COVID-19 pandemic) which meant full characterisation of the fundamental/pump pulses was not possible. The dark data was collected to a limiting resolution of 1.41 Å resolution as determined by FSC (figure A.4). As the jitter at PAL-XFEL is low (25 fs RMS [64]), so a timing tool was not needed to maintain accurate delays for the 1 ps and 100 µs time points. Statistics for the merged XFEL data are shown in table A.12.

The 1 ps data was taken as a control for comparison to LCLS and SACLA. Signals appear in similar locations on the chromophore and α helix but are notably weaker in the PAL data than either the SACLA or LCLS. It is not clear why the signals are weaker. It is possible a there was poorer coupling of the pump-pulse into the LCP jet compared with experiments using GDVN/droplet injection techniques (LCLS/SACLA). Additionally, the pulse duration was unknown and could be shorter than expected which may reduce the photolysis yields (as already discussed).

The main objective of the PAL-XFEL experiment was to confirm that isomerisation had occurred by 100 µs and was prepared in detectable concentrations. Looking at the Q-weighted difference density in figure 4.18 there are clear difference density signals which correspond to population transfer from the trans to

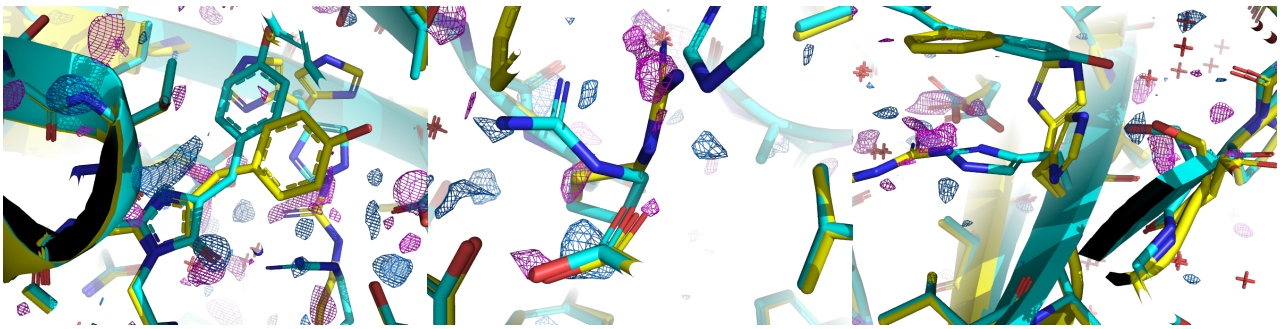


Figure 4.17: Q-weighted difference electron density between the 400 nm 1 ps and neg 1 ps laser-on, dark data. The ground-state coordinates are shown in yellow and cis, ground-state photoproduct coordinates are shown in green. Electron density contoured at +2 rms (blue) and -2 rms (magenta). Signals are very similar to the 1 ps delays observed at SACLA/LCLS.

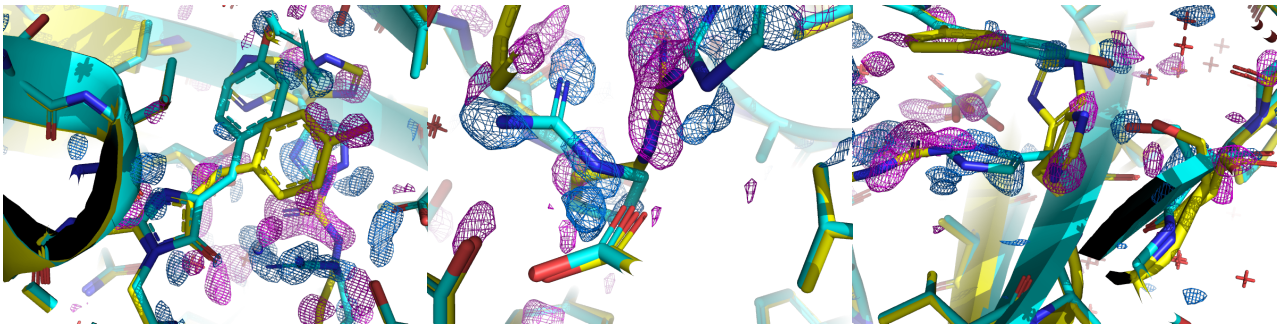


Figure 4.18: Q-weighted difference electron density between the 400 nm 100 µs and neg 1 ps laser-on, dark data. The ground-state coordinates are shown in yellow and cis, ground-state photoproduct coordinates are shown in green. Electron density contoured at +2 rms (blue) and -2 rms (magenta). Signals clearly show emergence of the photoproduct after 100 µs.

the cis-state. The signals are very similar in location to cryo-MX DED (chapter 3) especially close to the chromophore, Arg66 and His194 residues, all of which undergo rearrangement on isomerisation. To confirm this, occupancy refinement was carried out by varying the cis/trans occupancy of model coordinates and calculating the R_{Work}/R_{Free} for each 1% step. This showed a clear increase of 10-15 % in cis occupancy between the dark and 100 µs data (section A.7).

4.5 Conclusions

The dynamics of rsKiiro have been resolved using pump-probe TR-SFX from 250 fs – 100 µs at atomic resolution. Extrapolated structures generated from the DED have clearly resolved the adoption of a new geometry with clear movements to waters in the hydrogen bonding network of the phenol ring and structural changes of the α helix. The pump-dump condition was collected with delays from 250 fs – 100 ps and clearly shows stronger difference density signals across all conditions compared to the pump-probe data. The extrapolated pump-dump intermediate coordinates are very similar to the pump-probe but show a higher occupancy exaggerated translation movement of the chromophore.

To assign the extrapolated structure, visible transient absorption (TA) spectroscopy on solutions (C. Hutchison) and mid-IR TA (this author, chapter 6) were made. The presence of a dump pulse 350 fs after the pump was shown to cause an abrupt decrease of 50-60 % in the S_1 excited-state induced absorption spectrum with complete suppression of the stimulated emission signals. Flash-photolysis measurements on crystals scanned the dump delay and monitored monitored the cis-anion product. They confirmed the dump pulse significantly reduced the yields of formation of the product cis-anion state when delays were between 0 – 1 ps. This is unequivocal evidence that the isomerisation reaction occurs as an excited state reaction [5, 219], not a ground-state one [6, 7]. Furthermore, the cryo-kinetic spectroscopy measurements in chapter 3 suggest the existence of local ground-state minima likely close to the global neutral trans/anionic cis minima. These cryo-trapped states were shown to have a slightly different chromophore geometry to the global ground states. We, therefore, suggest that the TR-SFX measurements cannot resolve the excited-state electron density (at our experimental bandwidth) and can only resolve the ground state structural differences. This means we assign the pump-probe and pump-dump probe extrapolated coordinates to the trans ground-ground state local minima - equivalent to the cis-state intermediate which was resolved in chapter 3.

One of the consequences of using intense ultrafast optical excitation is the generation of vibrational coherences in the excited and ground states. In the LCLS TR-SFX measurements, the experimental bandwidth is limited to ∼4.4 THz as set by the pump pulse duration (100 fs) and 150 fs binning width. Under conditions of near-resonance pump frequencies and short-lived excited states, ground-state coherences will dominate in amplitude [290, 291]. These signals will decay on 1–2 ps time-scales [291–293] but could be enhanced with the addition of a dump pulse [290]. To interpret the sub 1 ps time-scale signals we therefore need to assign the excited/ground state wavepacket motion in the context of vibrational coherent motion. These assignments are beyond the scope of this thesis and are ongoing at the time of writing (S. Perrett and J. van Thor).

In conclusion optical-dumping of the electronic excited state have shown that, for rsKiiro, it seems TR-SFX is unable to resolve excited state intermediates. Instead, the observation of the ground state species indicates the formation of a coherently oscillating ground state. This has important ramifications for the field of TR-SFX and for previously published ultrafast crystallography experiments [212, 235, 241, 259, 263, 264, 294]. We have shown that more work must be done to confirm structural assignments in TR-SFX experiments.

5 | Serial Synchrotron Crystallography Feels the Heat

Time-resolved serial crystallography at synchrotrons (TR-SSX) is a relatively new technique built upon the methods developed for structure determination at X-ray free electron lasers. In this chapter, the trans-cis reaction of rsKiiro is probed at 10 ms using TR-SSX. The precise control of isomerisation yields is demonstrated through selective excitation with varied 400 nm laser intensities. High-resolution crystal structures at 1.75 Å resolution are obtained for energy densities between 0.1–15 mJ/mm^2. Direct comparison of crystallographic isomerisation yields to millisecond flash-photolysis measurements show that X-ray induced heating of an estimated 83–112 K, causes thermal reconversion over a 91 $kJ\,mol^{-1}$ ground state barrier during fixed target TR-SSX measurements.

Note to the reader: Experiments in this section were in collaboration with the serial crystallography EMBL beamline P14.EH2/T-REXX at Petra III, Hamburg. Experiments and analysis were led by this author. This chapter is currently being prepared for publication as:

> Baxter, J. M., Hutchison, C. D. M., Fadini, A., Cordon-Preciado, V., Morgan, M. R., Agthe, M., Tellkamp, F., Mehrabi, P., Pfeifer, Y., von Stetten, D., Pearson, A. R., & van Thor, J. J. Precise control of Photoisomerisation population transfer of a Fluorescent Protein in the presence of thermally driven barrier crossing shown by Quantitative millisecond time-resolved Serial X-ray Crystallography. *(in preparation for submission to JACS).*

5.1 Introduction

The development of XFELs has encouraged significant progress in data processing of stationary diffraction patterns (section 1.4.1) and catalysed the advancement of many rapid sample delivery mechanisms (section 1.3.2). These new high-throughput sample delivery techniques can be utilised at synchrotron sources to easily obtain many thousands of stationary diffraction snapshots [79]. Although this requires the production of large quantities of microcrystals compared to normal rotation techniques,

it allows for much shorter X-ray exposure times which significantly reduces the radiation dosage per crystal [144]. In turn, this allows for structure determination from particularly radiation-sensitive [295] and smaller crystal samples which [296], opens up the opportunity for dynamic room-temperature crystallography using synchrotron serial crystallography (SSX) [28].

Conventionally, macromolecular crystallography measurements are made at around 100 K as low temperatures significantly suppress radiation damage for a given radiation dosage [297]. However, cryogenic temperatures hinder the ability to study structural dynamics as cooler temperatures reduce conformational diversity [68]. Moreover, protein flexibility is reduced which can slow down or prevent the completion of reactions as well as introduce structural artefacts [298]. in time-resolved serial synchrotron crystallography (TR-SSX) reactions can be initiated with light, gas diffusion [116] or rapid buffer addition/mixing strategies [90]. With the current brilliance and detector technology available at micro-focus beamlines, TR-SSX is limited to a time-resolution of the order of milliseconds. However, with planned synchrotron upgrades likely to improve brilliance by 1-2 orders of magnitude [299, 300] and the development of quicker/more efficient detectors [301] it seems likely that routine microsecond measurements are not far away.

TR-SSX can provide longer-time point structures which complement the fast dynamics that can be captured at XFELs. For example, some of the first TR-SSX studies were on bacteriorhodopsin and span from 5 ms to 200 ms [138]. In combination with the early SFX experiments on the same sample, [264, 302] a near-complete movie of the entire rhodopsin photocycle has been captured. Additionally, the greater availability of synchrotron beamlines (and therefore easier access to beamtime) means that SSX experiments can be used to test crystal characteristics and the feasibility of an SFX experiment. However, TR-SSX is still a relatively new technique and there are open questions about the effects of intense X-ray exposure when not in the diffract-before-distruct regime as in TR-SFX. The experiments in this chapter will assess the level of X-ray induced heating in TR-SSX measurements which has long been considered a factor in MX measurements [143, 303–307] but is yet to be fully characterised in SSX. As the increasing brilliance of synchrotrons will lead to shorter X-ray exposures for structure determination, heating could become a major problem in TR-SSX, especially as X-ray exposure times become significantly shorter than thermal diffusion rates (on the order of µs for small crystals [305]). This highlights the need for understanding the heating effects of X-rays in TR-SSX measurements as higher time-resolution will be achieved with with higher X-ray fluxes.

In this chapter, a fixed target sample delivery system (section 1.3.2) is used to obtain a 10 ms time-point structure of the trans/cis reaction of rsKiiro. The energy density of 400 nm laser excitation is varied to precisely control the cis product yields. The crystallographic structures are interpreted using occupancy refinement to measure the population transfer. Light-induced structure factor differences are calculated

and show that the crystallographic data can detect even very small changes in isomer concentration. The product formation yields in the TR-SSX experiment are directly compared to flash-photolysis spectroscopy data which is used to determine the temperature increase caused by X-ray induced heating in fixed target TR-SSX experiments.

5.2 Methods

5.2.1 Flash-Photolysis

Crystalline samples, with absorption between 0.5-1 OD (at 505 nm), were prepared by crushing a microcrystal slurry between two glass cover-slips (Hampton Research HR3-254). Samples were pre-converted to the trans state, by illuminating with the unfocused beam a 100 mW 488 nm CW diode laser for 1 second. The trans/cis reaction was driven by a flash (2 ms TTL gated) from a 405 nm CW diode laser, focused to 40×29 µm (FWHM) spot. A series of neutral density filters (3 to 0.05 OD) was used to attenuate the pump beam at energy densities between 0.07 – 56.8 mJ/mm^2. A third CW diode laser at 505 nm was used to prove cis-anion absorption. The change in sample optical density was monitored using an amplified photodiode (Thorlabs APD440A) and digital oscilloscope (picoTechnology, PicoScope 4000). The power and focal spot size of the probe were chosen to minimise actinic behaviour, whist still allowing sufficient signal to noise on photodiode and overlap with the pump. Before and after flashing with each sample was fully cycled between the cis-trans state to determine the total photo switchable populations and irreversible bleach.

Modelling Flash Photolysis Product Formation

The photodiode signal (figure 5.1) is directly proportional to the cis anionic population and so can be used to calculate the proportion of population change (N) induced by each optical flash using the Beer-Lambert law:

$$N = \frac{[Cis]}{[CisTotal]} = \frac{log_{10}(I_0/I_{[Cis]})}{log_{10}(I_0/I_{[CisTotal]})}$$

Where [Cis] is the concentration of the cis conformer and [CisTotal] fully converted concentration (i.e. total sample concentration), I_0 is the initial intensity (or full trans), $I_{[Cis]}$ is the population due to the 405 nm flash and $I_{[CisTotal]}$ is the fully converted signal level of the cis state. On ms times scales the cis-trans isomerisation can be considered a single-step process:

$$\text{Trans} \xrightarrow{k_{Switch}, k_{Th}} \text{Cis}$$

With the photolysis switching rate, k_{Switch} and the thermal recovery (trans to cis) rate constant, k_{Th} contributing to the reaction. The switching rate constant depends on the power density per an unit area (P), the proportion of absorbed photons in a sample of absorbance, A, the quantum yield of switching, ϕ and the thermal rate, k_{Th}:

$$-\frac{d[Trans]}{dt} = \frac{P}{l}(1-10^{-A})\phi + k_{Th}[Trans]$$

For the case where OD << 1 a Taylor expansion can be used to approximate the absorbed fraction as:

$$1-10^{-A} \approx Aln(10)$$

Using this approximation and the Beer-Lambert law, the switching rate can be rewritten as:

$$-\frac{d[Trans]}{dt} = P\epsilon_{Trans}\phi ln10[Trans] + k_{Th}[Trans]$$

The rate of thermal recovery of Trans state during the flash-photolysis measurement is small (5.0×10^{-4} s^{-1}) compared to the the rate of photo-induced switching so can be set to zero. The rate with time can be rewritten in terms of energy density using: dt/dE = 1/P:

$$-\frac{dt}{dE}\frac{d[Trans]}{dt} = -\frac{d[Trans]}{dE} = \epsilon_{Trans}\phi ln10[Trans]$$

Giving an integrated rate equation of:

$$[Trans] = [Trans]_i \exp(-\epsilon_{Trans}\phi ln10E)$$

Assuming the total concentration of cis and trans states is constant, the fraction of cis population can be calculated as:

$$N = 1 - [Trans]_i \exp(-\epsilon_{Trans}\phi ln10E) = 1 - [Trans]_i \exp(-kE)$$

Where k is a rate constant in the units of mJ/mm^2. The data were fitted using the orthogonal distance regression package in scipy [308]. This minimises residuals by accounting for errors in both the dependent and independent variables when fitting which was appropriate due to the significant errors in both the energy density and the measured intensity.

Error Propagation

Low energy densities are very difficult to measure due to the low repetition rate of the flash laser (20 Hz) and the low-intensity pulses. This meant the measured energy density had a sizeable error. Furthermore, the percentage conversion was calculated by taking the mean intensity of the probe before the flash, after the flash and then after full re-conversion. These intensities were determined by averaging the photodiode signal over a window. Longer windows averaged over more points so gave lower standard errors of measurement- which was favourable. However, as the probe wavelength was actinic it caused re-conversion over time which therefore favoured shorter averaging-windows. As a result, the window size and the probe laser intensity were careful chosen to minimise the standard error in this experiment. The error associated with the conversion measurement is based on the population change (N) of a flash which was calculated from the photodiode signal using the Beer-Lambert law:

$$N = \frac{[Cis]}{[CisTotal]} = \frac{log_{10}(I_0/I_{[Cis]})}{log_{10}(I_0/I_{[CisTotal]})}$$

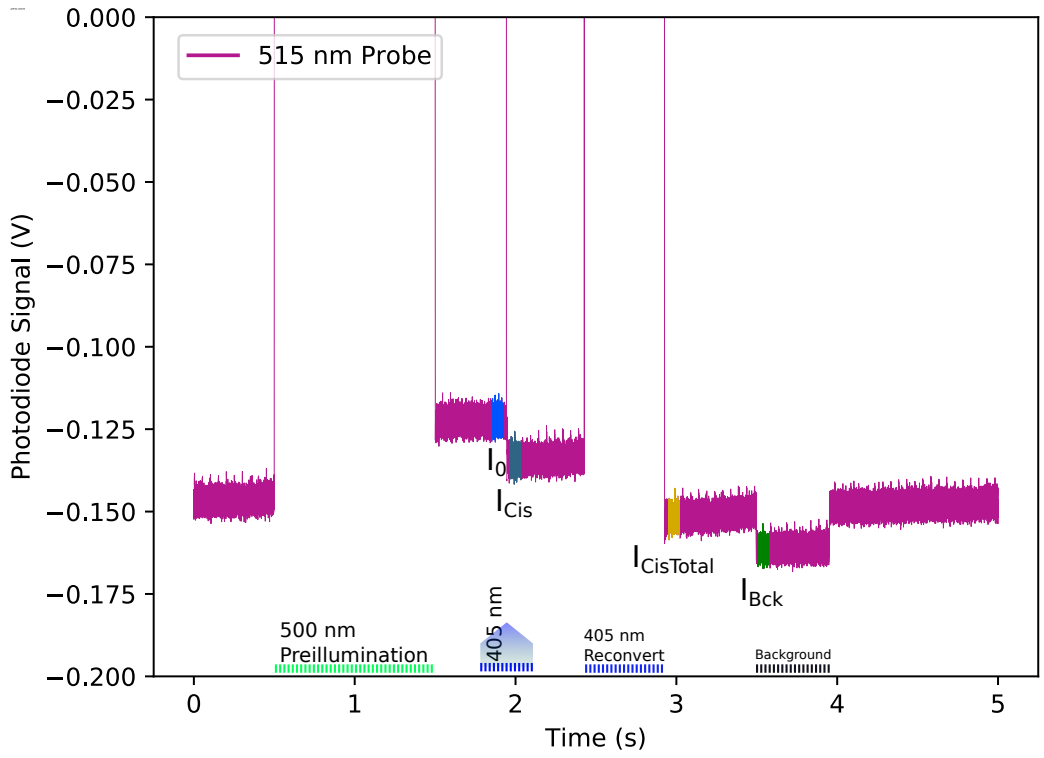


Figure 5.1: Plot of photodiode signal over one cycle of a flash photolysis measurement.

Where [Cis] is the concentration of the cis conformer and [CisTotal] fully converted concentration (i.e. total sample concentration), I_0 is the initial intensity (or full trans), $I_{[Cis]}$ is the population due to the 405 nm flash and $I_{[CisTotal]}$ is the fully converted signal level of the cis state. The standard error in the conversion, δN was calculated as:

$$\delta N = \sqrt{\frac{(I_{[CisTotal]}I_{[Cis]}\delta I_0 log_{10}(I_{[Cis]}/I_{[CisTotal]}))^2 + (\delta I_{[Cis]}I_0 log_{10}(I_0/I_{[CisTotal]}))^2 + (\delta I_{[CisTotal]}I_{[Cis]}I_0 log_{10}(I_0/I_{[Cis]}))^2}{(I_{[CisTotal]}I_{[Cis]}I_0 log_{10}(I_0/I_{[Cis]})^2)^2}}$$

Where $\delta T_{[CisTotal]}$, $\delta I_{[Cis]}$ δI_0 are the standard error in the respective values, calculated as: $\delta X = \frac{stdev(X)}{\sqrt{n}}$ where $stdev(X)$ is the standard deviation in value X over n observations. The energy density in mJ/mm^2 was calculated as:

$$E = E_{max} \times 10^{-ND}$$

Where E_{max} is the energy density with $ND = 0$ where ND is the optical density of the neutral density filter being used. The standard error in energy density, δE can be calculated as:

$$\delta E = \sqrt{10^{-2ND}(E^2\delta ND^2 + \delta E^2)}$$

5.2.2 Ground state Thermal Activation Barrier Measurement

To analyse the effect of increased X-ray heating we measured the ground state activation energy barrier following the measurement of recovery rates from 295–320 K. The sample temperature was controlled by varying the temperature of a nearby water bath (Julabo Refrigerated and Heating Circulator). This

water bath was connected to the sample cuvette holder (Hellma Quartz Suprasil Micro Absorption Cell 1.000) by plastic tubing. Once the water bath was set to a specific temperature, its temperature and that of the sample were monitored using a RS PRO RS42 digital thermometer until both were stable at the desired value. An Agilent 8453 spectrophotometer was used for the measurements. After blanking with buffer, measurements were carried out in the dark on 0.75 µmol protein solution. Once each specific stable temperature was reached in the sample, the protein was driven completely to the off (trans) state by illumination with a 500 nm LED. Thermal recovery of the on (cis) state was then recorded by taking the absorption at 395 nm and 481 nm every 30 s with a spectrophotometer integration time of 0.5 s. To prevent photodamage throughout the measurement, a 280 nm long-pass filter was used.

5.2.3 Serial Crystallography Data Collection and Reduction

A fixed-target sample delivery system using lithographed silicon chip, described in section 1.3.2 [127], was set up at the European Molecular Biology Laboratory (EMBL) end station P14.EH2 (T-REXX). To prepare the protein in the trans conformer, a 473 nm, 10 mW continuous wave (CW) laser was focused to around 1 mm (FWHM) on the chip and rastered over a block of 20×20 sample positions in a 5 by 5 pattern for a total of 25 seconds. To drive the reaction from the trans to the cis state, a flash of 405 nm light was combined co-linearly with the 473 nm laser beam and focused to 72.8 µm by 22.5 µm (FWHM) as measured by a fitted knife-edge scan [309]. The 405 nm laser output was TTL gated with 2 ms pulses and locked to the repetition rate of the fixed-target stages (20 Hz). The data collection proceeded with a 405 nm flash followed by a further 10 ms of X-ray exposure time with a photon-flux of around 2×10^{12} photons/s, X-ray beam-size of 15×10 µm (top hat), and wavelength 0.9801 Å. Diffraction images were collected using a Dectris Eiger 4M detector triggered by the fixed target stage controllers [127] (figure 5.2).

Batch grown crystals around 2×2×15 µm in size were produced in conditions of 25% poly-ethylene-glycol 3350 in 0.2 M lithium sulfate, 0.1 M Tris-HCl pH 8.5 with a final protein concentration of 15 $\mathrm{mg\,ml^{-1}}$ (similar to [1] and as described in section 2.2.4). The X-ray diffraction images were indexed using XGANDALF [152] and intensities merged using CrystFEL's Monte Carlo method without consideration for partialities [286]. Conversion of intensities to structure factors was done using PHENIX's reflection file converter and massage option [176, 310]. Coordinates were refined using REFMAC5 [237] initially using rigid-body type (for occupancy refinement) and then anisotropic B-factor refinement for the final structures. Difference electron density maps were calculated from weighted structure amplitudes based on the Bayesian statistics analysis by Ursby *et al.* [174, 175] in the same method described in chapter 3.

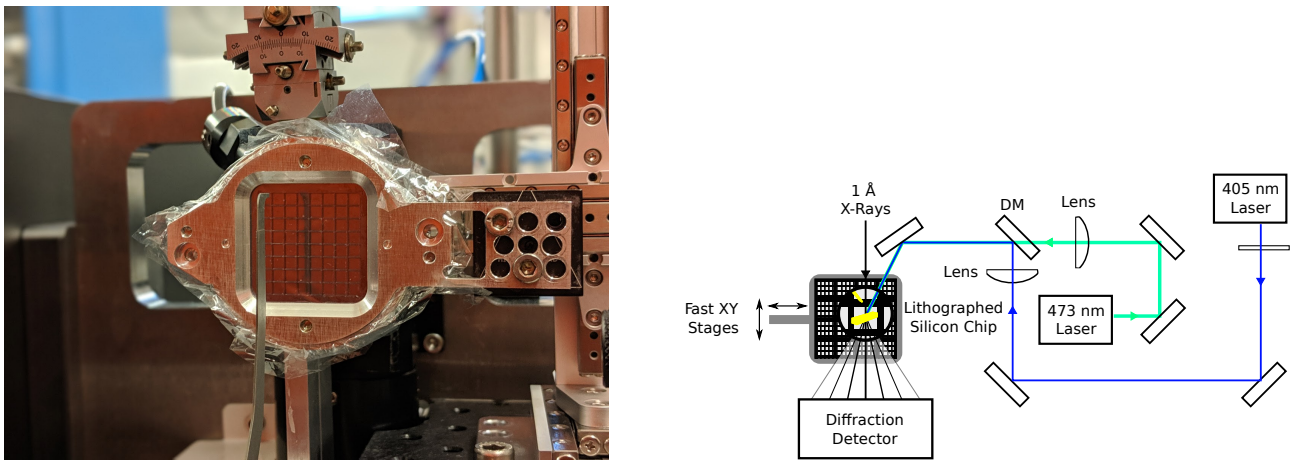


Figure 5.2: Photo and diagram of the setup described in the main text for TR-SSX experiments at EMBL beamline P14-2.

5.2.4 Crystallography Analysis

To analyse the occupancy of the different illumination energy densities of the protein three methods of crystallographic analysis were used: minimisation of R-factors, minimisation of difference electron density and regeneration of the end state through extrapolated structure factor addition [239, 259]. Difference electron density peaks were reported using Coot [289].

R-Factor Minimisation

The R-factor minimisation method involved varying the occupancy of a coordinate set from fully cis to fully trans occupancy in steps of 1%. The models of the pure-cis and pure-trans structures were initially generated from cryo-MX data. These 'pure' models were first refined against the 0 mJ/mm^2 (pure-cis) and 14.43 mJ/mm^2 (pure-trans) to account for changes in hydration and unit-cell size between the cryo-MX and room temperature data. Multiple conformations were removed from each model and the two coordinate sets combined into a single model using PHENIX [310]. The occupancy was altered for the cis and trans states from 0% cis (100% trans) to 100% cis (0% trans). Each generated coordinate set was refined using 5 cycles of rigid-body refinement to find the best fitting occupancy ratio which minimised R_{Work} and R_{Free}.

Minimisation of Difference Density

The second method involved the minimisation of difference density around the chromophore. Weighted difference electron density maps of observed structure factors (mF_o) minus calculated structure factors (nF_c) maps were generated for each occupancy step automatically by REFMAC. The occupancy with the lowest integrated negative electron density in the chromophore region was used to determine the occupancy assignment and therefore population transfer from the initial trans to the cis state.

Population Transfers from Extrapolated Maps

Thirdly, extrapolated structure factors and maps were calculated using the Bayesian weighted difference electron density according to Ursby *et al.* [174, 175]. Dark state (cis) coordinates were used to calculate structure factors (F_C) and phases. The characteristic N_{Ext} was generated using the same methods as in section 3.2.3. The population transfer (PT) was calculated as $PT = 200/N_{Ext}$ [259].

5.2.5 Modelling of X-ray Photolysis Yields to Determine X-ray Induced Heating

The TR-SSX measurement can be considered as a three-step process:

$$Cis + Trans \xrightarrow{Preconvert} Cis + Trans \xrightarrow{Flash} Cis + Trans \xrightarrow{X-ray\,Probe} Cis + Trans$$

As mentioned in section A.4.2 the initial preconversion step and thermal recovery before measurement should lead to an initial cis population of 18.6 % at maximum which was constant through the experiment. The population transfer caused by the flash can be modelled identically to the flash photolysis but with a starting population of 18.6% cis:

$$[Trans] = 0.814 \exp(-k_{Switch}E)$$

The thermal recovery occurs at a rate:

$$\frac{d[Trans]}{dt} = -k_{Th}[Trans]$$

$$[Trans] = [Trans]_i \times \exp(-k_{Th}t) = 0.814 \times \exp(-k_{Switch}E) \times \exp(-k_{Th}t)$$

The cis state and percentage conversion can then be modelled as:

$$N = 1 - [Trans] = 1 - b \times \exp(-k_{Switch}E) \text{ where } b = 0.814 \times \exp(-k_{Th}t)$$

Using the Arrhenius equation the temperature increase can be extracted as:

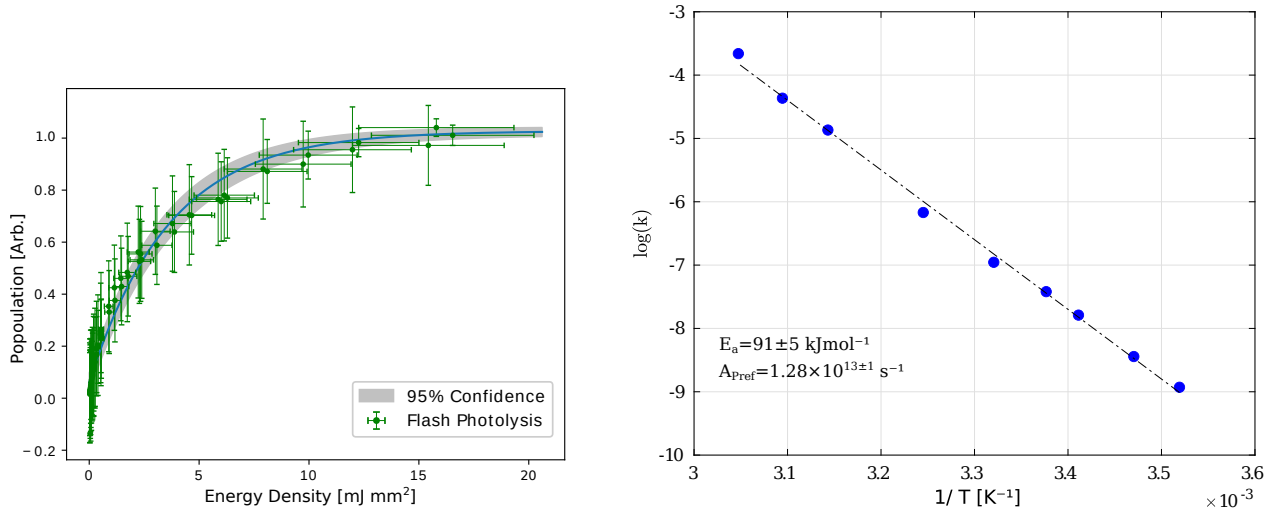
$$T = -\frac{E_a}{R} \times \left(\log \frac{-\log \frac{b}{0.814A}}{t} \right)^{-1}$$

5.3 Results and Discussion

The determination of the ground state energy activation barrier through measurement of the ground-state thermal recovery over a range of temperatures allows explicit calculation of the temperature increase due to X-ray photon absorption in the TR-SSX experiment. This calculation, however, relies on the accurate determination of populations from the crystallographic data. To access this accuracy, the crystallographic rate constants can be directly compared to the equivalent flash photolysis rate measurements as will be done in this section.

5.3.1 Flash Photolysis of Crystals and Thermal Activation Barrier Measurement

The flash photolysis measurements of rsKiiro crystals shows a plateau above energy densities of 10 mJ/mm^2 (figure 5.3 (a)). It is assumed this is due to the maximum product formation of 100 % trans neutral isomer. As absorption is probed at 515 nm, only the cis anion concentration is monitored and a unimolecular reaction model implies the creation of the trans-neutral state. Product formation of a cis-neutral or anionic-trans state would both show significantly blue-shifted electronic absorption spectra from 515 nm so would not interfere with the probe measurement [311].



(a) **(b)**

Figure 5.3: (a): Normalised concentration of the trans-neutral state formation plotted against energy density. The blue line shows the fitted model and grey shaded area is the 95th confidence percentile. Error bars are based on standard errors calculated as shown in the methods section. **(b):** Arrhenius plot derived from fitted rate constants – the insert shows the magnitude of the fit error. This plot was used with permission from Baxter *et al.* (unpublished).

Up to 32 cycles of photolysis measurements were made on a single spot of the crystal sample before being translated to a fresh area. A bleach rate was characterised by the total change in switched optical density ($I_0/I_{Cis\ Total}$) between cycles. The rate was found to vary with energy density but was always lower than 0.1 % per cycle. This suggests that rsKiiro is a highly stable photoswitcher. It also allowed multiple photolysis measurements to be reliably made on a single sample position without having to account for shot-to-shot bleaching.

The average optical density was 0.95 at 405 nm. This corresponds to a density of around 7.1 $\times 10^{15}$ protein molecules/cm^3 which is equivalent to a crystal of thickness around 3 µm. The average crystal dimension in the TR-SSX experiment was 6 µm. However, the needle shape meant crystals were selectively oriented with the longest dimension perpendicular to the laser beam/X-rays. This meant the mean path length through the crystal will be closer to the shortest dimension (2 µm) which is

comparable to the optical density of the spectroscopic crystal samples. The photokinetics of the two samples should therefore be similar and can be directly compared without scaling for the effects of different optical densities.

The data were modelled as described in the section 5.2.1 to extract an energy dependent rate constant. The model was found to closely describe the data, confirming the reaction followed first-order rate laws with no indication of bi-exponential behaviour. A rate constant, k, of 0.274 mm^2/mJ was fitted for the trans to cis reaction (off to on reaction). Conversion rates were also measured for a higher energy region from 15-57 mJ/mm^2 (not shown). In this region the population formation started to decrease from the plateau of 100 % seen in figure 5.3. This could be caused by the high photon flux creating a photoequilibrium with the neutral cis state (before the µs deprotonation step) which likely absorbs close to 400 nm [233]. Absorption of this species could reform the original trans-neutral state as is seen in mutant protein JvT75 at low pH in Chapter 7.

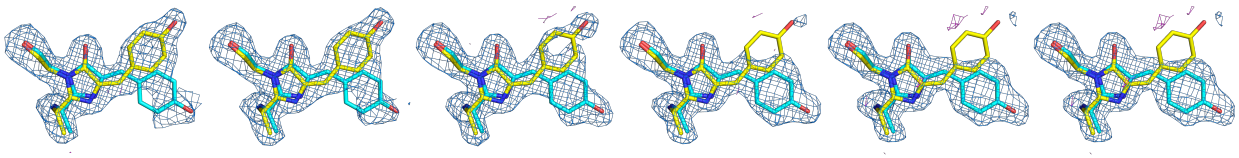
The cis-anionic is the most stable ground-state conformation meaning the trans-neutral state reconverts with a rate constant of 5×10^{-4} s^{-1} at 298 K. The activation energy of thermal reconversion was measured by fitting using the Arrhenius equation to thermal recovery rates between 292-340 K (figure 5.3 (b)). An Arrhenius prefactor of $1\times10^{13\pm1}$ s^{-1} and associated activation energy of 91±5 $kJ\,mol^{-1}$ were measured. Interestingly, this is a much larger barrier than in isolated HBDI which lies between 49-55 $kJ\,mol^{-1}$ depending on the protonation state [311]. This highlights the stabilising effect of the chromophore environment on the chromophore as compared to solution.

Finally, the laser-induced heating effects in the flash photolysis must be considered. At higher energy densities heating will be more significant. We would therefore expect a catalysed rate of trans-cis isomerisation to be observed at the highest energy densities. Looking at figure 5.3 (a) it is clear that the higher energy densities do not show significant deviations above the first-order fitting. This provides strong evidence that the conduction of heat away from the probed sample is efficient enough that only small amounts of the cis state will be thermally catalysed and react over the 91 $kJ\,mol^{-1}$ barrier.

5.3.2 Crystallography

In total, 79,976 indexed diffraction images were merged into 6 different datasets at laser energy densities of 0, 0.10 ,0.53, 1.78, 6.74, 14.43 mJ/mm^2 respectively. This gave an average indexing-rate of around 25% which is reasonable for SSX experiments using fixed target injection. Compared to cryo-MX structures of rsKiiro in the cis-dark state TR-SSX structures show a marked decrease in resolution to 1.70 Å and an increase of unit cell volume of 5%. The increase in unit cell size is in line with the expected thermal expansion when heating from 100-298 K [312]. Chromophore omitted F_o-F_c maps can be used to give an unbiased assessment of the chromophore isomerisation state. Looking at figure 5.4,

they show a clear progression of the occupancy from trans to cis with higher energy density illumination. The un-flashed data (0 mJ/mm^2) shows some signs of cis conformation. This is attributed to laser and X-ray induced heating driving isomerisation over the ground state barrier and will be discussed in section 5.3.4.



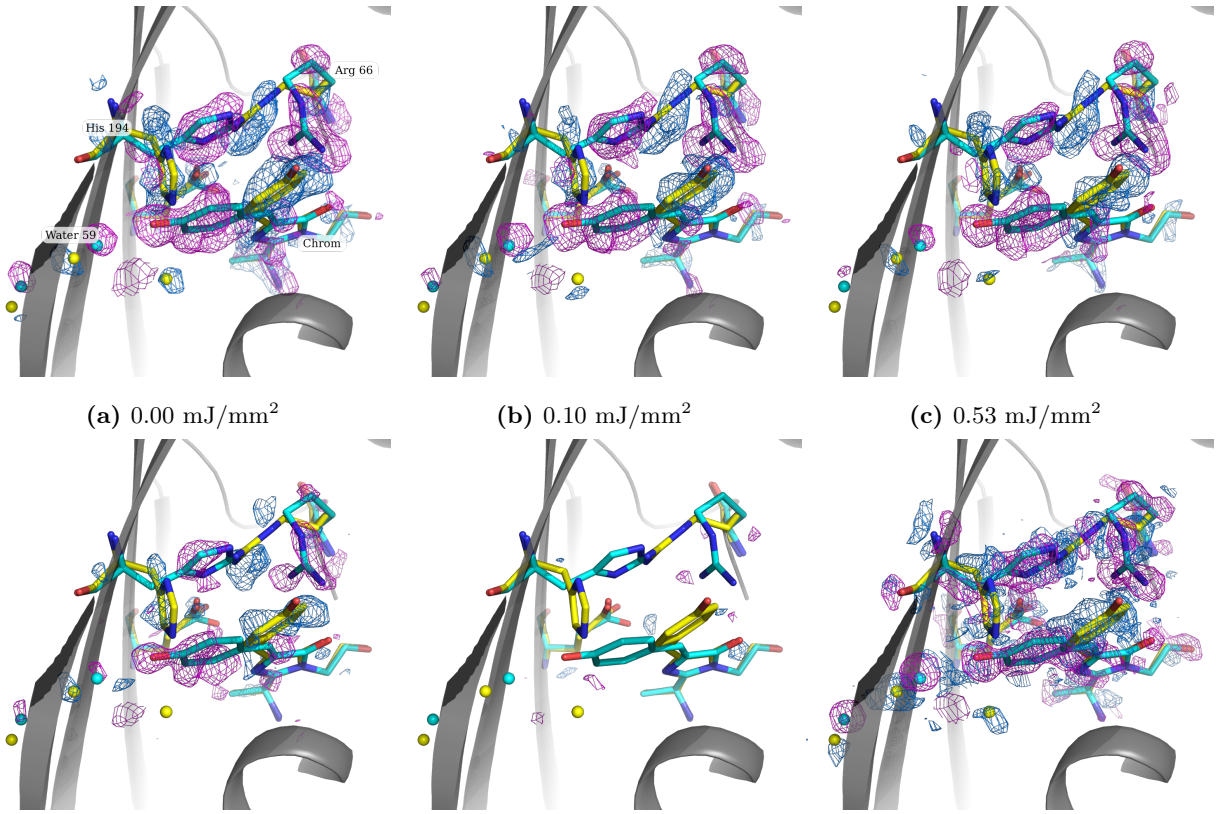
(a) 0.00 mJ/mm^2 **(b)** 0.10 mJ/mm^2 **(c)** 0.53 mJ/mm^2 **(d)** 1.78 mJ/mm^2 **(e)** 6.74 mJ/mm^2 **(f)** 14.43 mJ/mm^2

Figure 5.4: 2F$_o$-F$_c$ chromophore-omit maps of varying laser powers plotted at a sigma level of 3, calculated by phasing a 50:50 model of the cis and trans structures with the chromophore occupancy set to 0. Shown in blue is the positive density and in magenta the negative density.

A summary of merged crystallography statistics are shown in table A.16 for all laser energy densities. Overall, the data collection statistics show good quality across all measures with similar signal to noise ratios and comparable Wilson B-factors. A uniform high-resolution cut-off was chosen at 1.70 Å which corresponded to a value of 0.5 in Fourier shell correlation [313]. The lowest resolution dataset was the 0.1 mJ/mm^2 bin which was most likely because it contained the lowest number of indexed images. This suggests that the overall resolution would have been increased with higher multiplicity and was not fundamentally limited by the crystals. Refinement gave strong model-electron density agreement with R-factors appropriate for this resolution and the experiment type.

Q-Weighted F$_o$(E)-F$_o$(14 mJ/mm^2) difference electron density (DED) maps were calculated using a 50% cis, 50% trans model for phases (figure 5.5). The difference density peak amplitudes are inversely correlated with the higher energy densities, indicating that higher 405 nm photon fluxes cause increased population transfer from the trans to cis states. The differences density features in the TR-SSX data exactly matches the cryo-MX difference signals (figure 5.5 (f)) only differing in amplitude. The strongest signals are seen on the chromophore, His 194, Arg 66 and Water 59. The agreement of the TR-SSX and the cryo-MX signals suggest the isomerisation process has fully has taken place after 10 ms. Therefore, rsKiiro likely undergoes similar photochemistry to other similar FP such as Dronpa, the most widely studied reversibly switchable fluorescent protein (RSFP) [5, 6, 8, 233].

Heating is of particular concern to crystallography as it increases the rates of radiation damage. X-ray induced heating is caused by direct energy deposition of X-ray photons into the sample. Typically, a protein crystal will absorb around 2% of X-rays photons primarily through the photoelectric effect (1.5%) and Compton scattering (0.15%) [141]. This absorbed energy leads to stronger thermal vibrations as well as chemical reactions which are caused by the generation of radical species [143]. At cryogenic



(a) 0.00 mJ/mm^2 **(b)** 0.10 mJ/mm^2 **(c)** 0.53 mJ/mm^2

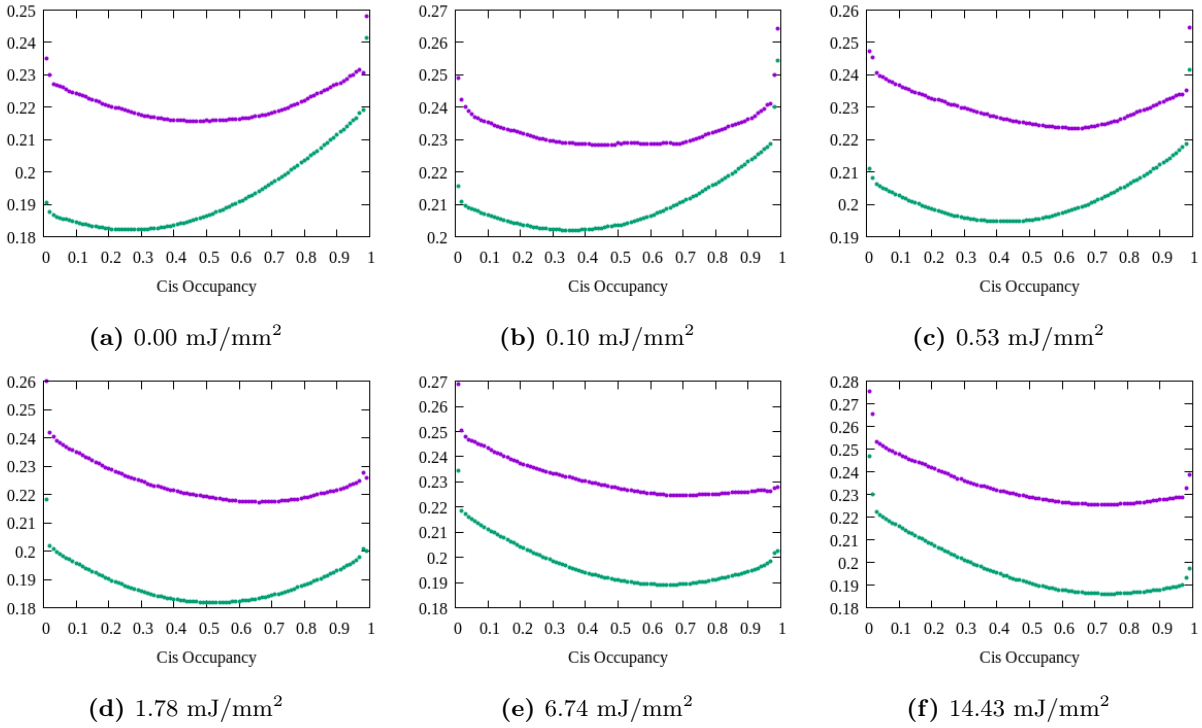
(d) 1.78 mJ/mm^2 **(e)** 6.74 mJ/mm^2 **(f)** Cryo-MX

Figure 5.5: Weighted difference electron density plotted at 3 rms level (positive in blue, negative is magenta). Cis coordinates are shown in cyan and trans coordinates are in yellow. Figures (a)-(e) show differences from the highest energy density (14.43 mJ/mm^2) structure factors: $F_o(14.43) - F_o(E)$. For comparison, (e) shows the difference of two steady state structures collected using MX.

temperatures the mobility of these radicals is low and so self-recombination occurs at high rates. At room temperature, radicals propagate much more quickly through the crystal structure, causing higher levels of radiation damage [142]. This means the tolerable dosage for room temperature measurements is much lower than in cryogenic measurements. The diffraction weighted dosage per crystal in this experiment was calculated using RADDOSE3D [191, 314], to be a maximum of 0.12 MGy per crystal. This is well below the estimated room temperature tolerable dose limit of 0.38 MGy suggested by de la Mora *et al.* [144]. Additionally, the phased electron density or rsKiiro shows no notable signals which would be associated with specific radiation damage by the X-ray photons. Further evidence of this is seen in the merging statistics such as the Wilson B-factor which are acceptable. This means the decrease in resolution 1.16 Å to 1.7 Å (Cryo-MX to TR-SSX) is most likely due to decreased crystal size/collection time preventing weaker reflections from being indexed/detected due to signal to noise limitations.

R-Factor Minimisation

R-factors were calculated in 1% steps from 0% to 100% for the cis-state which corresponded to 100% to



(a) 0.00 mJ/mm^2 **(b)** 0.10 mJ/mm^2 **(c)** 0.53 mJ/mm^2

(d) 1.78 mJ/mm^2 **(e)** 6.74 mJ/mm^2 **(f)** 14.43 mJ/mm^2

Figure 5.6: Plots of R_{Free} (purple) and R_{Work} (green) against the cis occupancy level for differing illumination powers.

0% occupancy of the trans-state. The 5 cycles of rigid-body refinement ensured that any differences in unit cell size were accounted for and did not skew the R-factor minima. As seen in figure 5.6, all energy densities were refined to at least 0.20 and 0.23 for the R_{Work} and R_{Free}, respectively. The 0.10 mJ/mm^2 bin showed the poorest rigid body refinement statistics (figure 5.6 (a)), most likely due to the lower redundancy (table A.16).

The population transfer is taken from the occupancy with the lowest R-factors for each condition. The extracted populations (figure 5.9), follow the same general trend as the flash photolysis data although are offset, showing populations greater than 0 for the dark data (0 mJ/mm^2). There is also a notable difference between the populations determined using R_{Free} and R_{Work} which is most likely due to the smaller number (5%) of R_{Free} reflections being less sensitive to subtle changes in structure.

Minimisation of Difference Density

The minimisation of DED is illustrated in figure 5.7. For each 1% occupancy step model structure factors were calculated (F_C). A nF_o-mF_c DED map was then calculated and the negative electron density was integrated within a 6 Å sphere centred on the chromophore. The minimum of integrated electron density was taken as the population transfer. Clearly, the general trend (figure 5.9) shows that higher energy density flashes give larger population transfers. A clear difference in population transfer is seen even between the dark and lowest flash energy (0.00 and 0.10 mJ/mm^2), where a difference of 3% in occupancy is observed.

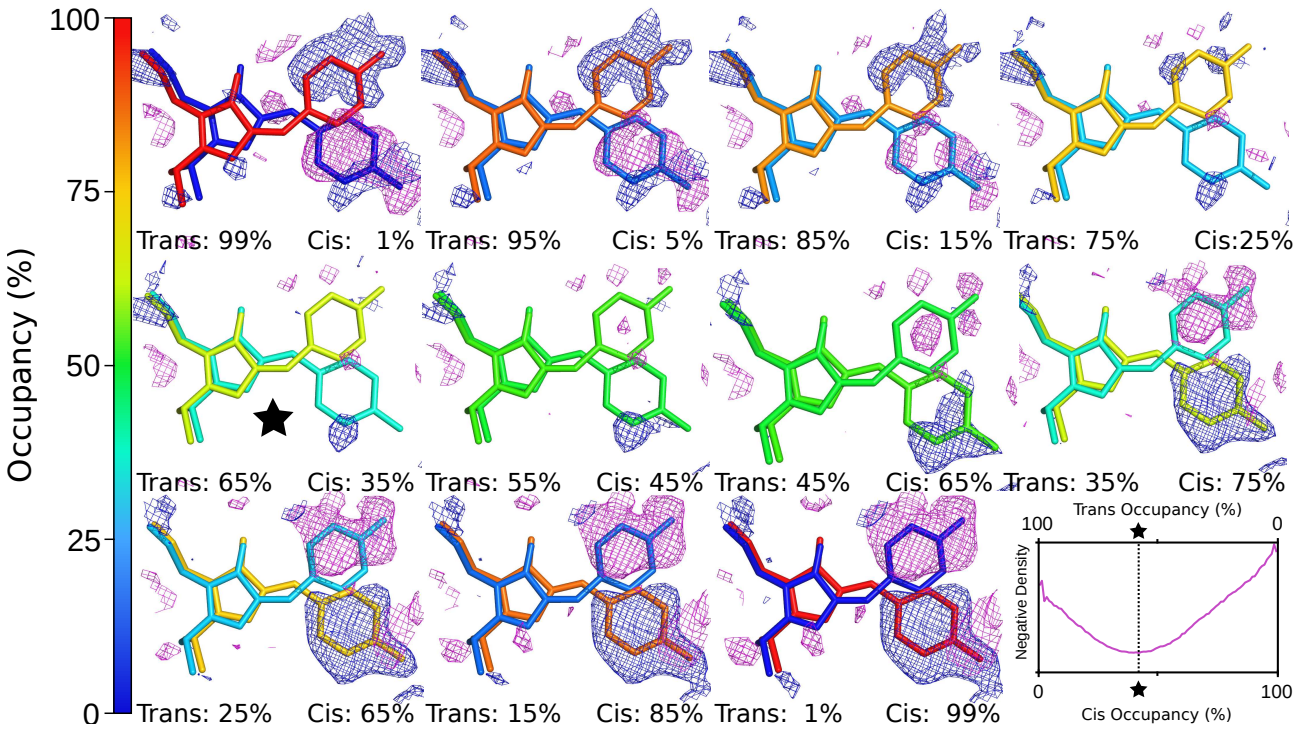
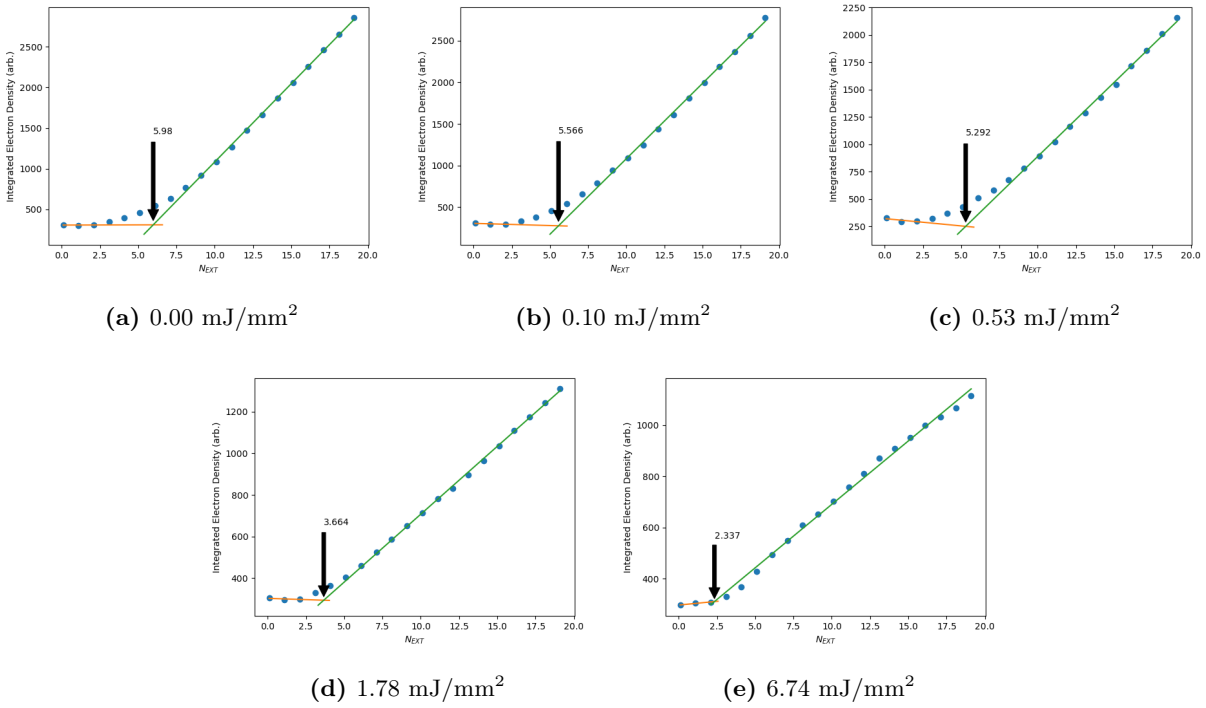


Figure 5.7: Weighted difference electron density of observed structure factor for 0.53 mJ/mm^2 energy density $mF_{Obs}(0.53)$ minus calculated structure factor nF_C(Occupancy). Plotted at 2.0 r.m.s level with negative and positive features in magenta and blue respectively. The percentage indicates the cis occupancy fraction. In the bottom right the integrated negative electron density around the chromophore is plotted as a function of the cis occupancy with the minimum indicated.

This method was found to be more robust in extracting population changes compared to the R-factor minimisation or extrapolated-map calculations. In particular, the dependence on the model coordinates was much lower. This is most likely due to the site-specificity of integrating only around the chromophore region. Using perturbed coordinates with intentional errors the density minimisation method still gave population transfers within 10% (compared to using fully-refined structures) as long as the errors were not in the chromophore region. In comparison, the R-factor and extrapolated-map calculations were significantly affected by the model coordinates and so it had to be ensured that the starting model was fully refined.

Extrapolated Map Calculation and Population Determination

Q-weighted DED maps (figure 5.5) clearly show greater product accumulation at higher energy densities as difference signals become more pronounced. Extrapolated structure factors were generated by adding multiples of Q-weighted structure factor differences to the calculated structure factors for the pure cis state. As the cis structure was refined against the highest energy density data (14.43 mJ/mm^2) the population for this state could not be calculated. The characteristic extrapolated map multiple, N_{Ext}, was calculated from the sharp increase in integrated electron density which was determined by fitting two linear regions to the integrated electron density around the chromophore (figure 5.8). Again, the



(a) 0.00 mJ/mm^2 **(b)** 0.10 mJ/mm^2 **(c)** 0.53 mJ/mm^2

(d) 1.78 mJ/mm^2 **(e)** 6.74 mJ/mm^2

Figure 5.8: Plots of integrated electron density in arbitrary units against multiples of added difference electron density (N_{Ext}) for different illumination powers.

expected increasing cis population at higher flash energies is seen, however, this method was found to be extremely sensitive to starting phases and coordinates and was not as robust as the R-factor minimisation or density minimisation.

5.3.3 Comparison of Flash-Photolysis and Serial Crystallography Photolysis Yields

The three methods of crystallographic population analysis show variation in rates (table A.14) but broadly the same trend and all suggest residual cis population in the un-flashed data. Figure 5.9 shows a direct comparison of the crystallographically obtained population transfer with the 95th percentile of the flash photolysis fit. The difference density minimisation method most closely follows the flash photolysis trend with the fitted rate constant varying by 14%. The R_{Free} determination is seen to greatly overestimate the product concentrations. As mentioned above this is likely due to the lower number of R_{Free}-flagged reflections. The R_{Work} determined populations are consistently lower than the difference minimisation method. This leads to a poor fit in determining the rate constant as the photolysis plateaus at 68% conversion instead of 100%. This discrepancy could be caused by a form of model bias. If the cis structures were not fully-refined or had significantly different Debye-Waller factors compared to the trans-state the R-factors may shift in the cis direction and inaccurate populations would be determined. Determination of population transfer with extrapolated maps showed a similar rate constant to the flash photolysis and is very similar in trend to the difference density minimisation. This is in part due to the poor rate fitting as the highest energy density structure was excluded due to

the 100 % bias.

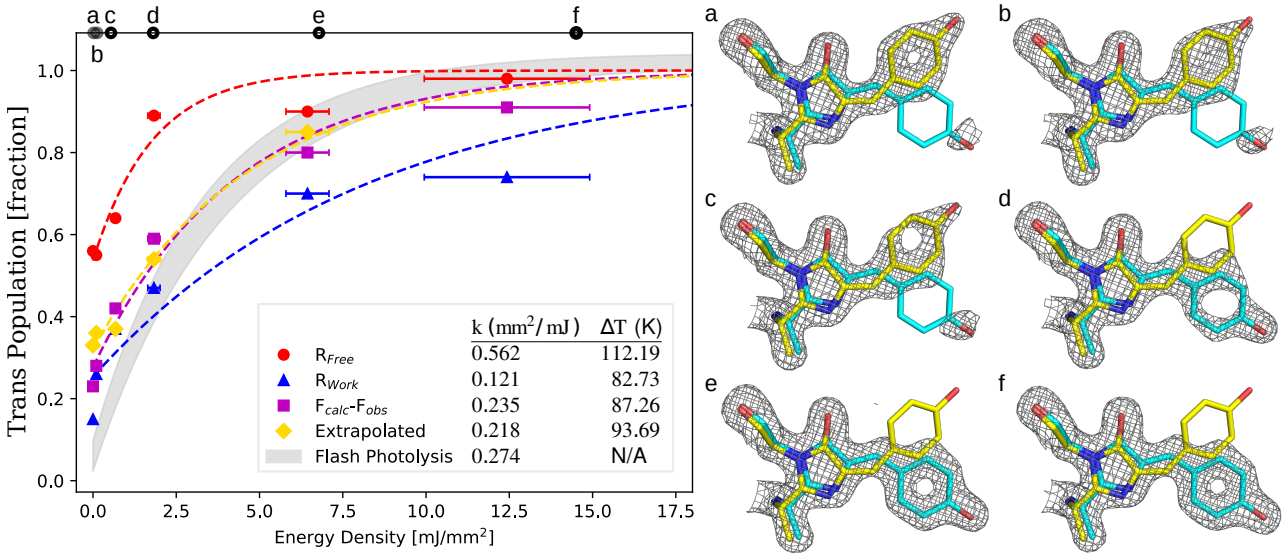


Figure 5.9: Left: Energy density plotted as a function of population fraction for multiple occupancy determination methods. Figures (a)-(f) show phased electron density at 1.5 RMS level using occupancies determined by minimisation of $2F_o$-F_c density as described in the main text.

5.3.4 Measurement of X-ray and Laser Induced Heating in TR-SSX

X-ray induced heating, incomplete laser conversion and thermal recovery before probing are responsible for the non-zero cis population offset of the crystallographic data compared to the flash photolysis. The incomplete laser conversion and thermal recovery are estimated to cause a maximum dark-state offset of 18.4% (section A.4.2) in the cis population. High levels of laser scatter over adjacent wells could also be responsible for residual cis-state populations, however, this is an energy density dependent process and would be expected to cause increased rate constants compared to the flash photolysis rates. As only the R_{Free} population determination shows a significantly higher rate constant the scatter induced conversion is assumed to be small. Instead, it is proposed that the initial cis state population at 0.00 mJ/mm^2 is created X-ray induced heating effects.

Laser Induced Heating:

An estimate of laser heating ΔT can be made by assuming absorbed photons that do not lead to switching (of yield 34 % (section A.4.1) switching, 5 % fluorescence (C. Hutchison, unpublished)) are dissipated as heat:

$$\Delta T = \frac{(1 - 10^{A(t,I)})(1 - 0.39)IA}{c_{Cryst}\rho_{Cryst}V_{Cryst} + (NV_{Cryst})c_{Solv}\rho_{Solv}}\Delta t$$

Where I is the photon flux (W/m^2), A the area of the crystal, c_{Cryst} and c_{Solv} are the specific heat capacities of the crystal and the solvent, ρ_{Cryst} ρ_{Solv} are the density of the crystal and solvent. N is multiple of crystal volumes of buffer the crystal is in contact with. For example, a value of 0 is a

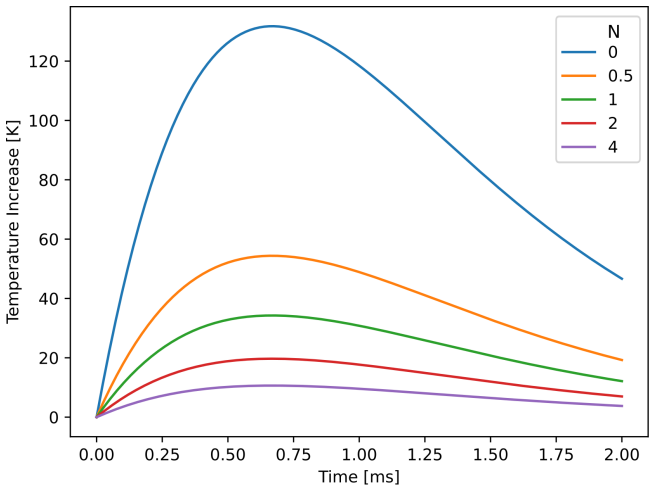


Figure 5.10: Temperature increases due to laser induced heating from the 400 nm, 2 ms pulse for the 14.43 mJ/mm^2 energy density. Plotted is the different volumes of buffer in contact with the crystal.

completely dry crystal with no surrounding buffer contact, whilst 1 is a crystal in a meniscus of the same volume as the crystal. This equation assumes the thermal diffusion time from the crystal to the buffer occurs much quicker than the rate of heating. Essentially we are assuming that a highly solvated (wet) crystal has a larger thermal mass to dissipate the laser-induced heating over. The absorption, A(t,I) is time dependent as the crystal is flashed and the cis-trans isomerisation occurs.

$$A(t, I) = A_0 exp(-k(I)t))$$

Figure 5.10 shows modelled temperature increases given different degrees of crystal wetness. N=0 corresponds to completely dry crystals whilst N=4 corresponds to a crystal in a drop 4 times the crystal volume. These calculations show that the solvation level of the crystal has a large effect on the thermal mass of the system. Previously we have found that dehydration of rsKiiro crystals tends to decrease the diffraction quality and increases the unit cell size. As the TR-SSX structures presented here show strong merging statistics and only expected thermal expansion of the unit cell it does not seem the structures are dehydrated. This suggests the crystals are suspended in a drop of mother-liquor when situated in the fixed-target-chip. Given the likely solvated nature of the crystal, we conclude that that laser-induced heating alone is not sufficiently high to account for the population changes observed. We assume the crystals are in a drop of sufficient size so that heating effects remain below 10 K (N>4). Additionally, given the effect of laser-induced heating is proportional to the energy density, higher rates of population transfer would be expected if the heating was significant. As discussed for the effects of laser scatter, only increased rate constants are seen in the R_{Free} determination of occupancy. This implies that laser-induced heating is not a major driving force over the ground state isomerisation barrier and suggests X-ray induced heating is likely to account for a significant proportion of the residual population.

X-ray Induced Heating

Helliwell [303] derives the rate of adiabatic temperatures increase due to X-ray induced heating without cooling ($\dot{\theta}$) as:

$$\dot{\theta} = \frac{I_0 \mu E \times 1.6 \times 10^{-19}}{\rho c_p}$$

Where I_0 is the photon flux per an area, μ is the X-ray attenuation coefficient, E is the X-ray energy per unit time, ρ the sample density and c_p the specific heat capacity. This can be rewritten and calculated as:

$$\Delta T = \frac{D}{c_p} = \frac{0.16 MGy}{1.8 \times 10^3 Jkg^{-1}K^{-1}} \approx 90K$$

Where D is the maximum dosage in Gy calculated using RADDOSE3D and c_p is the specific heat capacity of the protein. As the measurement of specific heat capacity have not been made for rsKiiro we used a value of $c_p = 1.8 \times 10^3$ $\mathrm{J\,kg^{-1}\,K^{-1}}$ determined from Lysozyme at 300 K [305]. Using this equation, a temperature increase of 90 K is calculated due only to X-ray absorption in the crystal. However, these calculations do not take into account the accumulative effect of X-ray heating of the silicon chip and the rate of heat transfer between the crystals and the sample holder. Although the mass attenuation length of the Silicon is comparable (7.9 $\mathrm{cm^{-1}}$ at 10 keV [315]) to the protein crystals (2.2 $\mathrm{cm^{-1}}$ calculated with RADDOSE3D [191, 314]), the chip is much thicker (mm depth) meaning it could cause appreciable heating of the sample holder. However, as it is very difficult to determine the thermal coupling between the protein crystals and silicon chip we consider the heating described below due to both X-ray absorption in the crystal.

As shown in section 5.3.1, the thermal activation barrier of the ground state trans-to-cis reaction was measured to be 91 $\mathrm{kJ\,mol^{-1}}$. The increase in cis population seen in the dark crystallography data was modelled as a three-step process with contributions from the thermal recovery, incomplete pre-conversion and the X-ray induced heating (section 5.2.5). Temperature increases between 83 – 112 K were extracted (31 – 181 K at 95 % confidence interval seen in table A.14). This observed temperature increase falls within the adiabatic calculation and is sufficiently high to surpass activation barriers in many biomolecular reactions in proteins. Warren *et al.* [307] directly measured the heating effects of X-ray energy deposition in microcrystals for cryo-MX experiments and found that temperature increases were much lower than the adiabatic estimates. Crystals in cryo-MX experiments are well thermally-grounded with efficient heat removal by the cryo-stream, cryoprotectant, and crystal mounting loop. Our observations imply that the fixed target system does not allow the same regulation of crystal temperature and the effect of temperature rises on a protein system should be considered carefully during the design of TR-SSX experiments.

5.4 Conclusions and Outlook

This chapter has shown that small population changes follow flash-photolysis measurements and can be accurately extracted using TR-SSX. This implies that TR-SSX does not significantly perturb the protein system during measurement and can be used to determine accurate ground-state structures. However, the determined temperature increases of around 90 K causes isomerisation over the ground-state barrier is particularly concerting for measurement of dynamics temperature-sensitive systems. In part, this is because increased temperatures exponentially decrease the lifetimes of intermediates that follow Arrhenius behaviour. Depending on the size of the barrier, this means even low levels of temperature increase can speed reactions beyond the time resolution of TR-SSX. For example, in PYP the PR_2 to PB_1 state would decrease in a lifetime by an order of magnitude given an increase in temperature of only 21 K from room temperature [316]. Our determination of a 90 K temperature increase would completely distort the reaction kinetics of PYP and make it very difficult to isolate the structure of any intermediate states. This, therefore, shows the importance of understanding the thermodynamics of a system before attempting to resolve the states crystallographically. Future experiments using a stepped exposure time/fluence as well the possibility of shuttering the X-ray beam whilst moving between well would help isolate the origin of temperature increase to the silicon absorption or crystal absorption.

Many major international synchrotron facilities are currently undergoing/planning/have upgrades which can deliver beams with lower emittance that will focus to smaller X-ray spot sizes [299, 300, 317, 318]. The smaller spot sizes will effectively increase the X-ray brilliance by 1-2 orders of magnitude [51]. With higher brilliance time-resolved synchrotron sources will be able to work at higher time resolutions, however, higher X-ray fluxes will also leave less time for thermal diffusion meaning the effective heating could be even higher. Thermal diffusion rates may ultimately limit the maximum time-resolution of future crystallography measurements at synchrotron facilities. For slower (ms) reactions, increasing detector frame rates could take advantage of the observed temperature to initiate reactions. Although this would require a more in-depth understanding of the rates of heating and would likely depend on many offer factors

6 | Ultrafast IRTA with Optical Control Reveals rsKiiro Isomerisation Mechanism

X-ray crystallography alone is an extremely powerful tool for understanding the dynamics and conformational landscape of proteins but it cannot definitively resolve protonation and electronic states of amino acids directly. Vibrational spectroscopy, in particular, the interpretation of difference spectra is very sensitive to changes in hydrogen bonding, protonation state and conformational rearrangement. Ultrafast spectroscopic studies of Dronpa/Dronpa-2 have proposed competing mechanisms for the trans/cis photoswitching reaction. This chapter extends the kinetic studies of rsKiiro to ultrafast vibrational spectroscopy. Initially, steady-state Fourier-transform infrared (FTIR) difference spectra allow mode assignments of a trans-neutral and cis-cationic species for the off and on states respectively. Ultrafast 400 nm-pump IR-probe spectroscopy of the off-state assigns the first step in isomerisation to the movement of the arginine 66 residue followed by excited-state isomerisation of the chromophore which occurs within 66 ps. A 515 nm dump pulse, 380 fs after the pump, was then introduced to induced stimulated emission of the excited state. The dump reduced the S_1 population by at least 40%. and suppressed the final photoproduct generation at delays up to 1 ms. Overall this chapter provides further evidence of excited state isomerisation in reversibly switchable fluorescent protein (RSFPs) and increases our understanding of photoswitching and stimulated emission processes in photoactive proteins.

6.1 Introduction

As introduced in chapter 3, the exact mechanism of switching in RSFPs is not fully understood. Van Thor and co-workers interpret transient infrared difference spectra as evidence for excited state isomerisation whilst similar experiments by Meech and co-workers conclude that ground-state isomerisation is responsible for the chromophore switching. In chapter 4 we concluded that difference density signals

could not extract a photoproduct structure from the 100 ps pump-probe delays. This was attributed to a lack of order in the photoproduct which does not support either an excited state or the ground state isomerisation mechanism. Cryo-trapping and cryo-spectroscopy of rsKiiro in chapter 3 suggested that the trans/cis reaction occupies a flattened cis-state intermediate as one of the final steps before full relaxation to the cis-ground state. This was taken as evidence of an excited state isomerisation based on thermodynamic arguments for ground state barrier heights that would have to be surpassed at cryogenic temperatures.

Vibrational mode assignment in GFP-like proteins is based on the work of Esposito *et al.* [319] who performed Hartree-Fock and density functional theory (DFT) calculations of the 4′-hydroxybenzylidene-2,3-dimethyl-imidazolinone (HBDI) isolated chromophore analogue [319]. These calculations were confirmed experimentally with an isotopic substitution study by He *et al.* [320]. Van Thor and co-workers then made specific assignments to GFP [321] and later to Dronpa [5]. Mode assignments in this chapter are based on harmonic frequency calculations of the isolated HBDI chromophore at the DFT B3LYP/6-311+g(d,p) level and interpretation of steady-state IR spectra of Dronpa which were both made by Warren et al. [5]. A summary of the harmonic frequency calculations is seen in table A.15. Overall, Warren et al. made calculations for $^{1}H_2O/^{2}H_2O$ substitutions of the chromophore in neutral, anionic, cationic and zwitterionic forms in the cis and trans conformations. Restrained calculations were made with the phenol and imidazolinone planes frozen at 15°, which is the angle observed in off state crystal structures of Dronpa [322]. These calculations provided unscaled frequencies as a basis for relative shifts under isotope exchanges and relative intensities of chromophore modes which are used as a basis for assignment throughout this chapter. Furthermore, assignments are also supported by the work of van Thor *et al.* in resolving FTIR steady-state spectra of GFP [323].

A summary of the lifetimes and rate constants associated with published visible and IR transient absorption spectroscopy studies of proteins in the Anthozoa family is shown in table 6.1. Most electronic spectroscopy studies of this class of proteins have shown a biphasic decay of the S_1 excited state with a fast (sub 10 ps) and slower component (10-100 ps) followed by slower decays and a µs deprotonation step [212, 233, 235]. Vibrational mid-IR studies have only been undertaken on Dronpa and the M165T Dronpa mutant (known also as Dronpa-2). With the exception of the earliest study by Warren *et al.*, which only collected data up to 100 ps, the vibrational spectroscopy results broadly agree with electronic rate constants being a biphasic evolution/decay of the excited state followed by slower/long-lived ground state evolution. Only the study of Laptenok *et al.* [7] could access long enough pump-probe delays to observe the final product formation. However, their conclusion of a ground state isomerisation is against both against electronic spectroscopy and TR-SFX results [212, 235].

In this chapter ultrafast transient and steady-state vibrational spectra are assigned in the context of

Electronic Spectroscopy			
Sample	Sequentially Fitted Rate Constants [ps]	Mechanism	Reference
Dronpa	2.1 → 10 → 64 → µs	ESI	Yadav *et al.* [233]
Dronpa-2	2.6 → 15 → 160 → 280 →µs	ESI	Yadav *et al.* [233]
rsEGFP-2	0.9 → 3.7 →µs	ESI	Coquelle *et al.* [212]
IrisFP	2.02 → 15.3 →µs	ESI	Colletier *et al.* [232]
rsKiiro	5.78 → 60.65 →µs	ESI	C. Hutchison (Unpublished)
Vibrational Spectroscopy			
Dronpa	9 → Long-lived	ESI	Warren *et al.*[5]
Dronpa-2	0.6 → 14 → 17 → Long-lived	ESI	Kaucikas *et al.*[8]
Dronpa-2	2.3 → 21.6 → 500 → Long-lived	GSI	Lukacs *et al.*[6]
Dronpa-2	16 → 596 → 91 ns → 4.8 µs → 156 µs	GSI	Laptenok *et al.*[7]

Table 6.1: Summary of published rate constants for different RSFPs and the assigned mechanism of either excited state isomerisation (ESI) or ground state isomerisation (GSI).

the time-resolved serial crystallography and cryo-trapping crystallography measurements undertaken in chapters 3 and 4. The ultrafast switching mechanism of rsKiiro is identified using pump-probe and pump-dump-probe spectroscopy which shows the creation of a photoproduct can be quenched by dumping of the excited state. This provides strong evidence that the isomerisation reaction occurs in the excited state.

6.2 Methods

6.2.1 Infrared Spectroscopy: Sample Preparation

Non-deuterated samples were prepared by concentrating the protein sample to around 95 mg/ml in a protein low salt buffer of 50 mM Tris-HCl and 100 mM NaCl at pH 8 (or 50 mM MES-HCl and 100 mM NaCl at pH 5). A Vivispin 2 with membrane size of 30 kDa was used for concentration at 4000 g. 15 µl of concentrated protein was then placed in a Harrick cell between two CaF_2 UV polished-windows (1 mm and 2 mm) using a 6 µm Teflon spacer, giving a sample OD of around 0.7 at 490 nm.

To prepare deuterated samples, the same buffers were prepared using deuterated water at either pD 8.4 or pD 5.9. Around 500 µl of non-deuterated protein sample was diluted to 2 ml and reconcentrated to its original volume in a Vivispin 2 concentrator with membrane size of 30 kDa at 4000 g. This was repeated three times to minimise the water content to below 0.2 %. Final sample concentration was 60 mg/ml and was placed in a Harrick cell between two CaF_2 UV polished-windows (1 mm and 2 mm)

using a 25 µm Teflon spacer, giving a sample OD of around 0.8 at 490 nm.

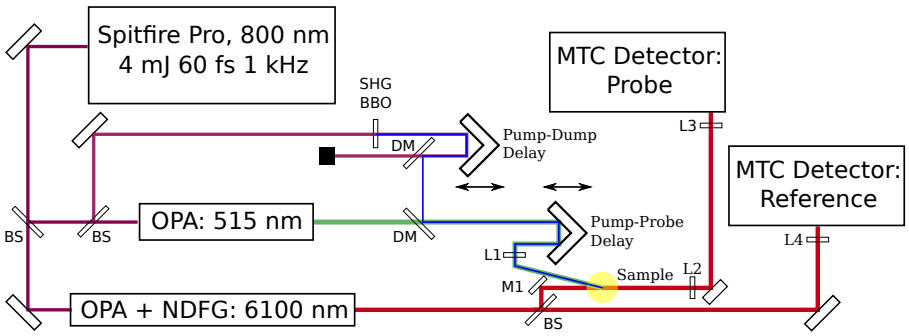
6.2.2 Fourier-Transform Infrared Spectroscopy

Fourier-Transform Infrared (FTIR) spectra were measured using a Bruker ISF 66/S spectrometer equipped with a liquid nitrogen-cooled mercury cadmium telluride-A MCT-A detector. The spectrometer was set to a speed of 40 kHz using a filter 5 kHz, an under-sampling ratio of 2, resolution of 4 cm cm^{-1}, the symmetric interferogram, 80 gain range radius, Laser centre of 15800.8 cm^{-1}, with 100 background scans and 100 sample scans. Dry air was used to purge the spectrometer for 1 hour in between sample replacement and before taking the background spectra. Illumination of the sample (prepared as stated in section 6.2.1) was done through the side port of the spectrometer. A 500 mW 405 nm (100 mW in 25 mm diameter spot at the sample) or 100 mW 488 nm (20 mW in 25 mm diameter spot at the sample) were aligned and defocused using a misaligned 1:1 telescope. The sample was illuminated for around 30 s to allow for full photoswitching. Difference spectra were calculated directly from the unilluminated - illuminated spectra.

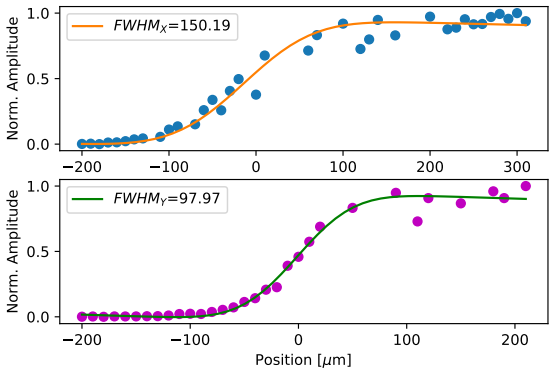
6.2.3 IR Transient Absorption Spectroscopy

The 4 W, 70 fs, 800 nm, 1 kHz output of a Ti:Sapphire regenerative amplifier (Spitfire Pro, Spectra Physics) was split equally into three paths (figure 6.1). Two were used to drive two optical parametric amplifiers (OPA) (TOPAS-C, Light Conversion). The first, was used in second-harmonic-signal mode to generate a 150 fs 515 nm pulse and the second OPA was equipped with a non-collinear difference frequency generator (NDFG) which was set to generate 6.1 µm 150 fs pulses from the difference of the OPA's signal and idler beams. In the third arm, a 100 µm BBO cut for the second harmonic generation (SHG) of 800 nm to 400 nm was used as a pump beam. A delay stage and retroreflector were used to compensate for the path through the visible OPA before the 400 nm and 515 nm beams were combined using collinearly using a two long pass dichroic mirrors (DM in figure 6.1 (a)) which also filtered out any residual 800 nm from the fundamental beam.

The infrared beam was split into a reference and probe. The probe beam was focused onto the sample plane using an off-axis parabolic mirror to a spot size of around 150×100 µm (figure 6.1 (b)). The beam was recollimated using a lens and focused onto the spectrometer (Triax190, Horiba) with a liquid nitrogen cooled 128 pixel MTC array detector (Infrared associates) where data was acquired using an FPAS acquisition system (Infrared Systems Development Corp.). The reference beam was also directed onto a separate spectrometer/detector/data acquisition system of the same specification. The MTC array pixels were separated by 50 µm 'dead space' and were each 0.2 mm in width and 0.5 mm in height. The pixel uniformity was specified better than 15%. The spectrometer slit size was set to 0.5 mm onto a 9000 line/mm grating giving a final spectral resolution of around 4 cm^{-1}. A coarse calibration of



(a) Schematic for IRTA



(b) Fitted knife edge scans of the IR focus

Figure 6.1: (a): A simplified schematic of the IR transient absorption setup described in the main text. Three beamsplitters (BS) split the output of the Spitfire Pro equally and are used to generate the 400 nm pump, using a 100 µm second harmonic generation barium borate (SHG BBO) crystal, the 515 nm dump pulse using a visible OPA, and a 6100 nm beam using an OPA and NDFG. The beams are all focused non-colinearly onto the sample and the IR is directed into the probe detector. **(b)**: Knife edge scans demonstrating the FWHM focus size of 150×100 µm for IR beam.

pixel number to wavelength was performed using a polystyrene film and a fine calibration was made using iterative fitting of water vapour absorption. After fitting, spectra from the IRTA setup and FTIR instrument were in very close agreement (within ∼2cm^{-1}). There was one dead-pixel (pixel 81) close to the centre of the array in the probe channel at a wavelength of approximately 1655 cm^{-1}.

A custom-built sample translator held a Harrick cell with 1 and 2 mm (front and back) CaF_2 UV polished windows. Around 20 µl of deuterated protein sample at a concentration around 50 mg/ml (prepared as described in 6.2.1) was sealed between the windows using a 25 µm spacer. The optical density at 400 nm was around 0.2-0.3 and OD which equated to around 1 OD between 1000-2000 cm^{-1}. The sample translator moved the cell in Lissajous patterns at an average speed of around 50 mm ms^{-1}. An unfocused 100 mW 488 nm CW diode laser of beamsize ∼3×5 mm to preconvert/reconvert the sample to the off state. Data where the pre-conversion laser was switched off, was obtained as a control. Difference signals were seen to slowly vanish with multiple scans, but spectral features seemed identical

so it was concluded the pre-conversion laser had no artificial effect on the kinetics.

An optical chopper wheel was used to cut the repetition rate of the visible pump/pump-dump to 500 Hz. The relative timing of the visible pulses to the IR was adjusted using a delay stage which allowed a maximum delay of 2.2 ns. The polarization of the visible beam was set to the 'magic angle' of 54.7° relative to the IR probe. A 150 mm focal length lens was used to focus the visible beams onto the sample plane. The focus size was measured using a camera. The FWHM diameter of the 400 nm and 515 nm pulses were ∼320×240 and ∼150×135 µm with 1.39 and 1.19 µJ energy respectively. The choice of beam size was a trade-off between maximising and maintaining spatial overlap (of the pump, probe and dump) whilst maintaining sample translation (minimising pumping of the next shot). Data were acquired in a pump on/pump off scheme where every other shot was probe only. Outlier rejection using the mean pump on/pump off signal and the reference spectra was used to discard anomalous signals due to dust or bubbles in the sample.

Pump-probe delays between -40 ps and 2 ns were collected in non-sequential/random order to remove any effects of sample degradation. 5000 spectra were averaged for each time point per scan and between 5-10 scans were collected for each condition. Scans were averaged together using a custom-written Matlab script that applied outlier rejection again. Outlier rejection involved the removal of measurements more than three scaled median absolute deviations from the median. Outliers were not included in the averaging. This was necessary as occasionally the data acquisition system would lock-on to an outlying value and reject accurate data. The noise level of measurements was characterised to 0.3×10^{-4} OD (at 1 σ) level which low and allowed mOD level signals to be detected. Sample imperfections and laser stability were the leading causes of noise in the measurements. For global analysis, negative time points from -40 to -30 ps were averaged and subtracted from the whole dataset to remove any long-lived (millisecond) signals caused by overlap with the previously previously pumped sample due to sample translation speed not being fast enough to measure on fresh sample with each shot. The spectroscopy analysis suite, known as the "Ultrafast Toolbox" was used to analyse the midIR response to the optical pump and dump pulses [186]. Fitting of the instrument response function at t_0 for scans of 400 nm only and 515 nm only on the GaAs were used to correct for dispersion of the probe and to tune the delay between the 400 and 515 nm pulses which was set to 380 fs for these measurements. A sequential model using coupled differential equations was used to retrieve distinct evolution-associated difference spectra (EADS) and rate constants for each intermediate state.

6.3 Results

6.3.1 Steady State FTIR Spectra

The steady-state FTIR difference spectra of rsKiiro allow identification of the protonation state and conformational form of the chromophore in the non-fluorescent and fluorescent states. Additionally, signals belonging to the arginine 66 residue, which forms a hydrogen bond with the imidazoline oxygen in the cis state, can be identified.

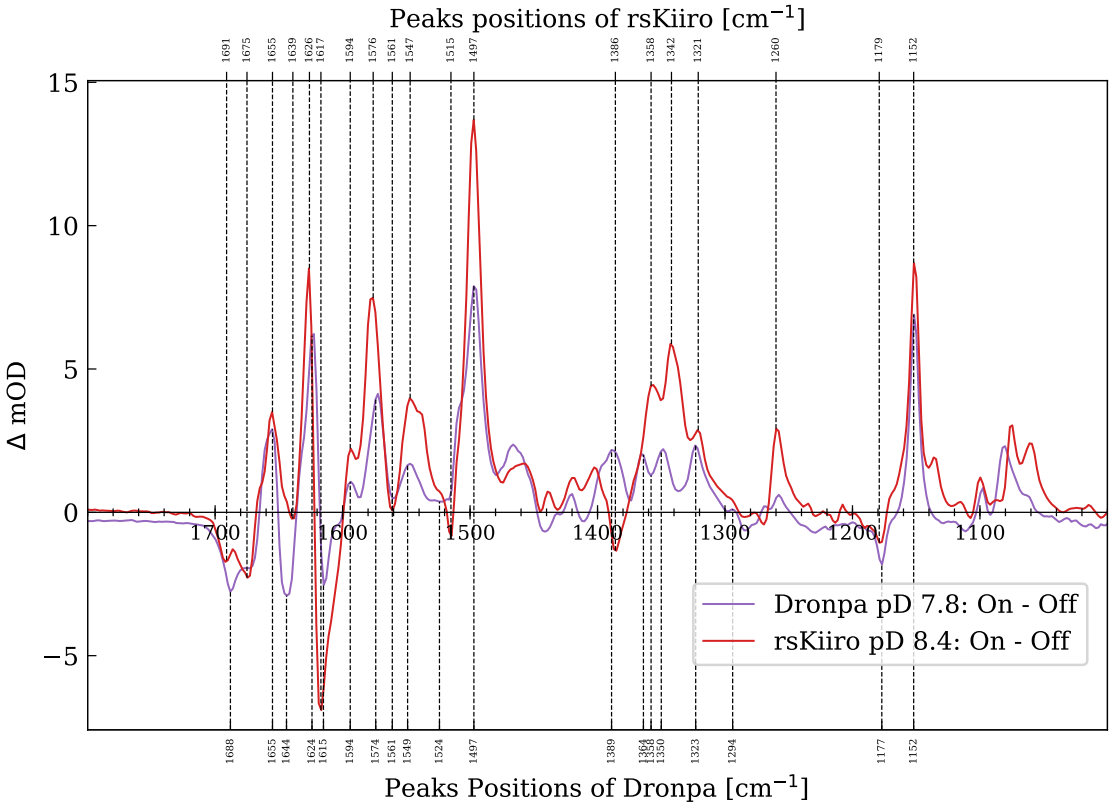


Figure 6.2: Steady state FTIR difference spectra for rsKiiro and Dronpa [5]. Shown between 1800-1000 cm^{-1} at pD 8.4 and 7.8 for rsKiiro and Dronpa respectively. Photoconversion of rsKiiro with 490 nm light is referred to in the legend as the 'off' state. Peaks discussed in the text are labelled on the top and bottom axes as indicated by the dashed lines.

Figure 6.2 shows difference spectra between the dark (on) state and the photoconverted (off) state for rsKiiro and Dronpa at pD 8.4 and 7.8 respectively. It is immediately clear Dronpa and rsKiiro share very similar spectral features in terms of peak locations and relative amplitudes. The peaks at frequency positions 1690 and 1655 cm^{-1} correspond to the well characterised C=O chromophore and are assigned to the trans neutral and cis anion as determined in similar proteins [5, 8, 321, 323, 324]. Alternative protonation states of the imidazolinone, such as a cis-zwitterion or trans cation would cause significant blue shifts to this band (table A.15) and as a result are not assigned. Positive peaks in the difference spectra are therefore due to the fluorescent cis anionic state and negative peaks in the trans

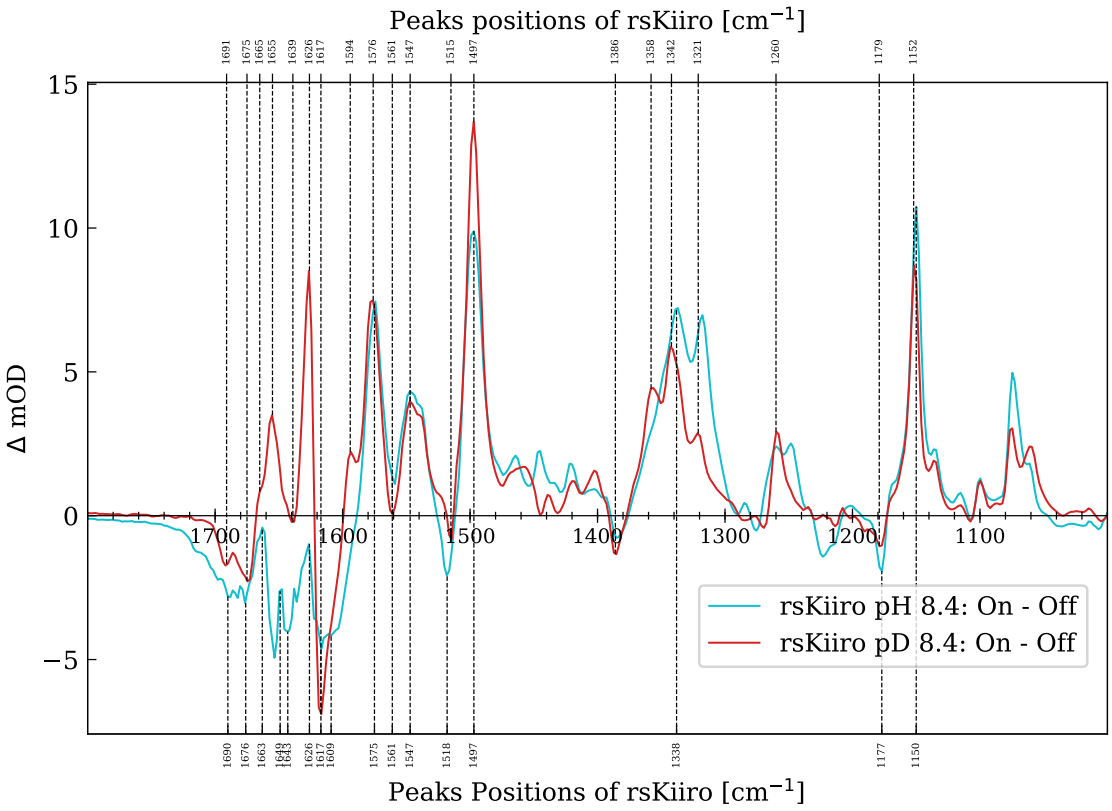


Figure 6.3: Steady state FTIR difference spectra for rsKiiro shown between 1800-1000 cm^{-1} at pD/pH 8.4. Photoconversion with 490 nm is referred to as the off state in the legend. Peaks discussed in the text are labelled on the top and bottom axes as indicated by the dashed lines.

neutral state.

In the deuterated buffer, well-resolved peaks with positive amplitude are seen at 1626, 1576, 1547, 1497 and 1152 cm^{-1} these are assigned to the ν(C=C), Phenol-1, ν(C=N/C=C), Phenol-3 and Phenol stretches, as shown in table 6.2. In ^{1}H buffer, spectra show notable shifts to the C=O to 1663 cm^{-1}, a small shift of the Phenol-1 and Phenol frequencies to 1575 and 1497 cm^{-1} and no shifts to the C=C, C=N/C=C or Phenol-3 peaks. Well-resolved negative peaks in the pD 8.4 spectra are seen at 1690, 1675, 1617, 1561, 1515 and 1179 cm^{-1}, assigned to ν(C=O), ν(C=O), ν(C=C), Phenol-1, ν(C=N/C=C) and Phenol stretches. Isotope substitution, seen in the pH 8.4 spectra show small shifts of both ν(C=O) band, to 1690/1676 cm^{-1}, the ν(C=C) to 1643 cm^{-1}, the Phenol-3 to 1518 cm^{-1} and the Phenol mode to 1177 cm^{-1}. The Phenol-1 band is not observed to shift under isotope exchange.

The most noticeable difference between the Dronpa and rsKiiro spectra are features in the region between 1300-1400 cm^{-1}. In Dronpa, a pattern of peaks at 1389, 1364, 1338, 1350 and 1323 cm^{-1} are observed, rsKiiro has a prominent peak at 1342 cm^{-1} and two smaller peaks at 1358 and 1321 cm^{-1} with an enhanced negative peak at 1386 cm^{-1}. Similar changes were reported by Kaucikas *et al.* for

Dominant Component	Dronpa pD 7.8 On	rsKiiro pD 8.4 On	rsKiiro pH 8.4 On
C=O	1665	1655	1663
Arg 66 asymmetric ($CN_3H_5^+$)	1655	1655	1649
C=C	1624	1626	1626
Arg 66 symmetric ($CN_3H_5^+$)	1594	1594	1609
Phenol-1	1574	1576	1575
Phenol-2	-	-	-
C=N/C=C	1549	1547	1547
Phenol-3	1497	1497	1497
$^{13}C_5$	1350	-	-
Phenol	1152	1152	1149

Table 6.2: rsKiiro and Dronpa [5] on-state FTIR assignments for ^{2}H and ^{1}H buffers. Dronpa assignments made by Warren *et al.*. All values are in units of cm^{-1}.

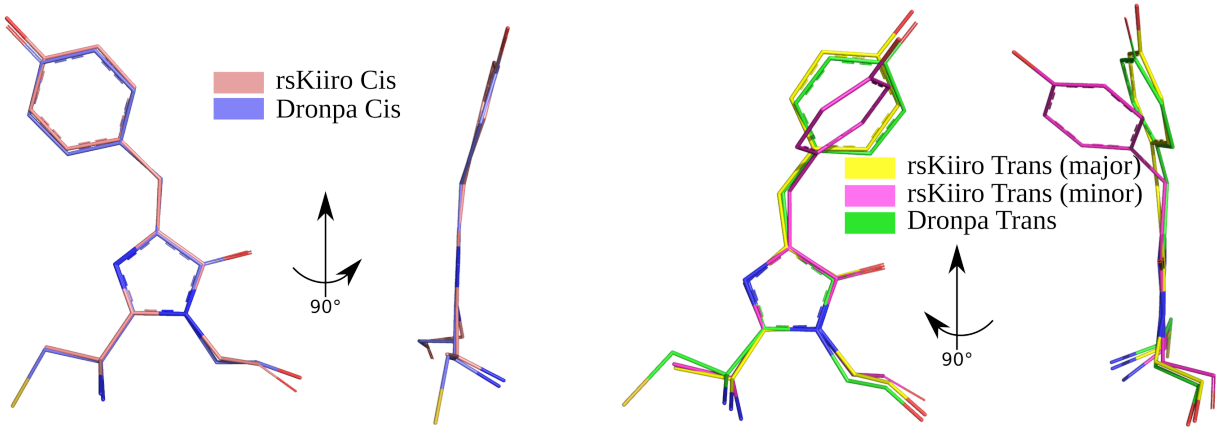
the M159T mutant of Dronpa [8]. In wild-type Dronpa the band at 1350 cm^{-1} is equivalent to the a 1371 cm^{-1} peak seen in anionic HBDI [311]. He *et al.* found in HBDI this peak redshifts by 15 cm^{-1} on isotope substitution for a ^{13}C at position 5 (the carbon in the imidazolinone ring which forms a bond to the methyl bridge - seen in figure 2.3) whilst substitution of other adjacent atoms with ^{15}N or ^{13}C caused no notable shift in mode frequency [320]. It was, suggested that this peak is delocalised and is affected by shifting/bending of the chromophore skeleton. As the crystal structure of rsKiiro (figure 6.4) shows a slight change in the C5 position compared to Dronpa, assignment of a 5-^{13}C sensitive band is therefore made at 1338/1338 cm^{-1} in the $^1H_2O/^2H_2O$ spectra respectively. Finally, bands at 1665/1649 cm^{-1} and 1594/1609 cm^{-1} for the $^1H_2O/^2H_2O$ spectra are due to the $\nu_{asym}(CN_3H_5^+)$ and $\nu_{sym}(CN_3H_5^+)$ of the Arg 66 which form a hydrogen bond with the cis anionic state. The ^1H_2O peak is notably weaker and difficult to resolve compared to the ^2H_2O peak.

6.3.2 Visible Pump Mid-IR Transient Absorption Spectroscopy

Time-resolved difference spectra were taken in the range of -40 ps - 2 ns (and 1 ms) between 1800-1500 cm^{-1}. The instrument response function was fitted with a FWHM of 190 fs. The pump-probe data was fitted with two sequentially related evolution associated difference spectra (EADS) with constants 5.79, 60.65 and a further long-lived component (figure 6.5). These first two decay constants closely correspond to the first two electronic excited state components of 7.67 and 49.99 ps of rsKiiro (C. Hutchison, unpublished) and other published studies on similar RSFPs (table 6.1). The pump-dump experiment shows around a 60% reduction in bleach intensity (figure 6.6) which compares closely to the visible pump-dump experiment with similar energy densities which showed around a 50 % dumping of

Dominant Component	Dronpa pD 7.8 Off	rsKiiro pD 8.4 Off	rsKiiro pH 8.4 Off
C=O	1688	1691/1675	1690/1676
Arg 66 asymmetric ($CN_3H_5^+$)	-	-	-
C=C	1644	1639	1643
Arg 66 symmetric ($CN_3H_5^+$)	-	-	-
Phenol-1	1615	1617	1617
Phenol-2	1600	-	-
C=N/C=C	1561	1561	1561
Phenol-3	1524	1515	1518
$^{13}C_5$	1358	-	-
Phenol	1177	1179	1177

Table 6.3: rsKiiro and Dronpa [5] off-state FTIR assignments for ^{2}H and ^{1}H buffers. Dronpa assignments made by Warren *et al.*. All values are in units of cm^{-1}.



(a) Cis **(b)** Trans

Figure 6.4: Chromophore alignment from the crystal structures of JvT75 at pH 5.5, pH 8.4 and Dronpa. The Dronpa structures were obtained from the protein data bank with IDs: 2POX (trans) and 2IOV (cis).

the excited state (C. Hutchison, unpublished).

Pump-Probe

The quickest rate constant follows the first phase of the S_1 decay (figure 6.5). Ground state bleaches (GSBs) corresponding to the trans neutral are observed at 1689, 1644 and 1513 cm^{-1} for the C=O, C=C and Phenol-3 bands, as assigned above. The GSB at 1632 cm^{-1} is unassigned from the steady-state analysis but possibly arises from phenol and C=N/C=C modes. Induced absorption is seen at 1668 and 1596 cm^{-1} and is assigned to rearrangement of the Arg66 asymmetric and symmetric $CN_3H_5^+$ modes respectively, suggesting rearrangement of this residue/hydrogen bonding network occurs as one of the first steps in isomerisation. The 1570 and 1457 cm^{-1} are unassigned but could be caused by hot vibrational states formed from the intramolecular vibrational redistribution (IVR) of higher frequency modes.

The 60 ps EADS closely agrees with the second phase of the electronic S_1 decay and shows similar spectral features to the first decay spectra but there is a new GSB at 1618 cm^{-1} and excited-state absorption (ESAs) peak at 1632 cm^{-1}. The GSB clearly corresponds to the trans-neutral phenol-1 mode whilst the ESA is assigned to a twisted chromophore (discussed below). Also of note is the very weak ESA around 1530 cm^{-1} and the lack of any ESA features which could correspond to phenol modes in the FTIR spectra.

The final component was fitted as a long-lived ($\tau > 2ns$) component which is interpreted as a flattened ground-state species. The bleaches at 1680, 1638 and 1618 cm^{-1} correspond to the C=O, C=C and Phenol-1 modes respectively. Excited-state absorptions are seen at 1650, 1630 and 1596 cm^{-1} and are assigned to the C=O, C=C and Arg 66 asymmetric $CN_3H_5^+$ modes of a cis neutral state. The absence of a strong phenol-3 and any dominant modes between 1500-1600 cm^{-1} should be noted.

Pump-Dump-Probe

The dump pulse delay was 380 fs when measured by cross-correlation on a GaAs semiconductor. The wavelength (515 nm) was not resonant with any ground-state of visible ESA cross sections meaning a double difference spectra ($\Delta\Delta A = \Delta A_{PDP} - \Delta A_{PP}$) were used to interpret the dump interaction. This double difference spectrum was globally fitted at time points after the 515 nm pulse arrival and the associated coherent artefact had decayed (t>500 fs).

Looking at the $\Delta\Delta A$ (PDP-PP) spectra (figure 6.5) and PDP ΔA_{PDP} decays (figure 6.6), the dump pulse clearly causes a large decrease in amplitude of the strongest GSB at 1644 cm^{-1}, with a decrease of around 40% in amplitude compared to the pump-probe (PP) data. The pump-dump-probe (PDP) and the dump-only traces (figure 6.6) show a large coherent artefact which is attributed to the optical Kerr effect at the 515 nm pulse arrival time. For comparison to the PP data a double-difference spectrogram

was prepared with Pump Dump - Dump Only differences (figure A.10 (d) and A.11 (d)). This shows the clear arrival and reduction of GSB caused by stimulated emission pumping of the dump pulse.

The first EADS component was fitted with a 40 fs component. This is shorter than the fitted IRF so is attributed to a combination of the optical Kerr effect and IVR/vibrational cooling of high frequency excited vibrational modes. It is not plotted in figure 6.5. The next component of the global fit, with a rate constant of 95 ps shows almost identical, but inverted, spectral features as the second pump-probe EADS component and so follows the assignment that the excited state has reduced significantly in population due to the dump pulse (figure 6.5).

The long-lived ground state component with shows very similar GSB frequencies to the final component in the PP data. However, the peaks, attributed to ground-state cis-neutral formation in the PP data is not seen in the PDP spectra (figure A.10 (b) and (d)). Only the 1668 and 1593 cm^{-1} peaks, attributed to the symmetric/asymmetric Arg($CN_3H_5^+$) are seen. Although the global fit shows a low average residual value of 1.41×10^{-5}, the final component does not account for the slight change in spectra between 1.5 and 2 ns. As seen in figure 6.5, the 2 ns spectrum shows two product bands at approximately the positions of the chromophore C=O and C=C stretches as will be discussed.

6.4 Discussion

The FTIR and multi pulse IR-TA measurements have allowed the definitive assignment of the on and off states and the photoswitching mechanism, in rsKiiro which has not been possible with crystallographic measurements (chapter 3 and 4). Overall, FTIR difference spectra have confirmed the cis anion and trans neutral for the fluorescent and non-fluorescent species. These are the same states found in Dronpa [5] which is expected due to rsKiiro's structural and photophysical similarities. IRTA measurements have then assigned alteration to the Arg66 as one of the first steps in the photoswitching reaction followed by excited-state isomerisation of the HBDI chromophore to a cis-neutral state.

Steady-State Assignments

The assignment of a cis anionic state is supported firstly by the nature and position of the C=O band as similar peaks are seen in Dronpa [5] and GFP [323]. No isotope shift is expected in the anionic cis state as the site of isotope exchange, the phenolic oxygen should be deprotonated. This is confirmed for the C=C, Phenol-3 and C=N/C=C bands which do not shift in frequency position. However, small shifts are seen at 1576/1575 and 1152/1149 cm^{-1} for Phenol-1 and Phenol bands in $^1H_2O/^2H_2O$ respectively. The shift in phenol-1 peak is small and possibly caused by the presence of the peak at 1594 cm^{-1}. The shift in Phenol peak is likely caused by strong ^2H_2O absorption around 1200 cm^{-1} skewing the sensitivity of the spectrometer and artificially shifting the peak.

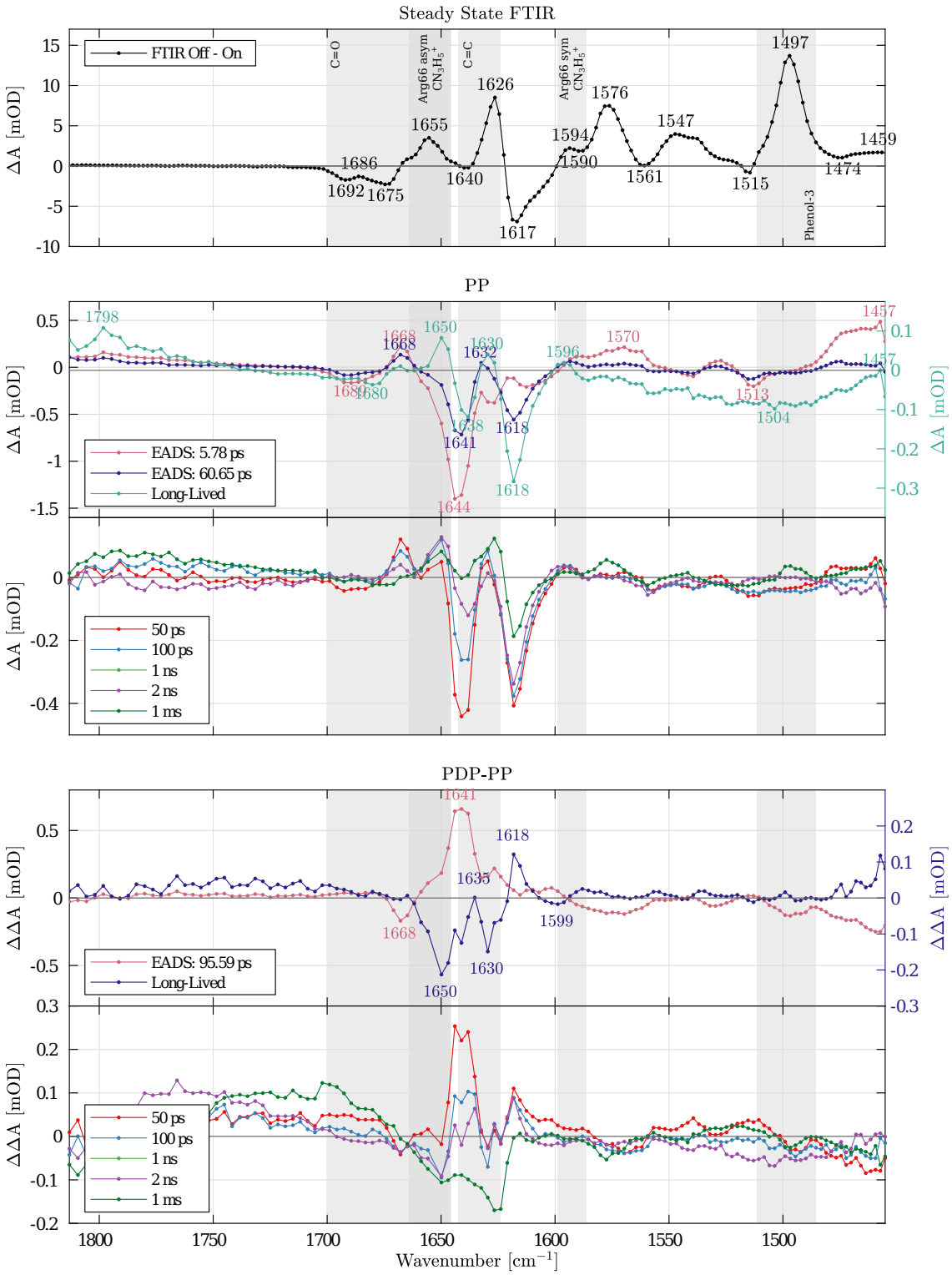


Figure 6.5: Top: FTIR On - Off steady-state difference spectra at pD 8.4. **Middle & Bottom:** Evolution associated difference spectra (EADS) fitted using a sequential model. The long-lived time points are plotted on the right-hand axis for both the pump probe and pump dump condition. The pump-dump data is fitted from 390 fs onward. The raw spectra for pump-probe delays of 100 ps, 2 ns and 1 ms (-40 ps) are plotted on the bottom axis for the pump probe and pump dump probe of each panel. The dump-delay remained constant at 380 fs.

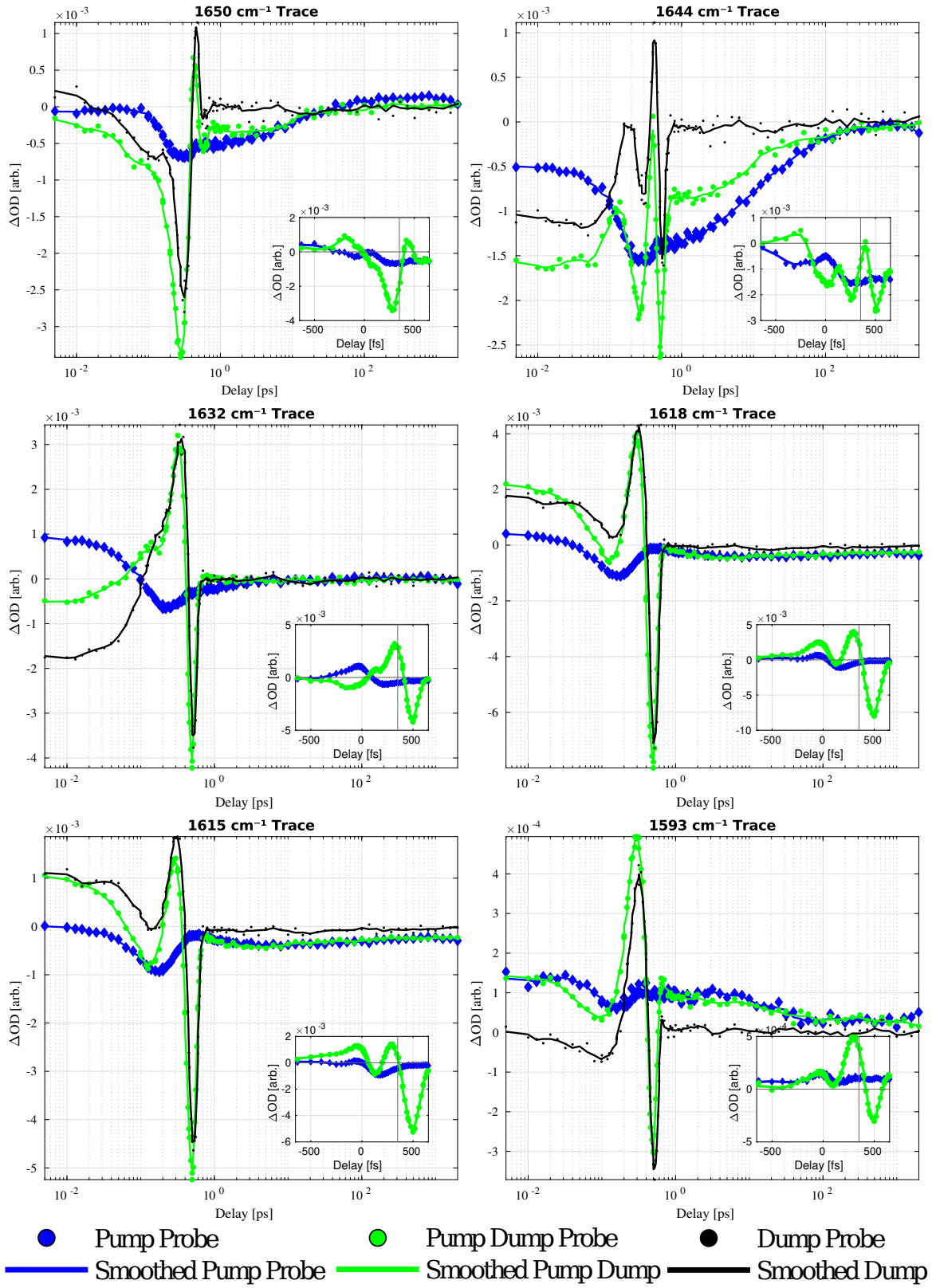


Figure 6.6: Traces for 1650, 1644, 1632, 1618, 1615 and 1593 cm^{-1} as a function of pump probe delay: blue shows pump on - probe only, green shows pump dump (380 fs) - probe only then black shows dump only (380 fs) - probe only. All plotted data had the -40 ps (1 ms) subtracted from subsequent time points and are plotted relative to the PP delay. The PD delay was kept constant at 380 fs.

The C=O shows a large $^2H_2O/^1H_2O$ shift from 1655 to 1663 cm^{-1}. In the isolated chromophore, this would make an anionic assignment unfeasible. However, the protein environment allows for hydrogen bonding and stabilisation of the chromophore which means some bands are not localised only to changes in the HBDI chromophore. The shift in ν(C=O), therefore, implies this stretch is not only due to the chromophore so other amino acids with vibrational absorption frequencies close to 1655/1663 cm^{-1} can be considered. Absorption in this region are typically due to Gln(C=O), Asn(C=O) and symmetric/asymmetric Arg($CN_3H_5^+$) [325]. As the $^2H_2O/^1H_2O$ shifts are indicative of an arginine stretch and as both the symmetric and asymmetric peaks can be assigned the spectroscopy data alone suggests a Arg($CN_3H_5^+$) perturbs the C=O frequency. As detailed in chapter 3 of this thesis, the crystal structure of rsKiiro known to exceptionally high resolution. Figure 6.7 shows the chromophore environment and interaction to aid interpretation of the FTIR spectra. Clearly, a hydrogen bond can be seen between the Arg66 and C=O of the cis chromophore which is not present in the trans structure. This leads to the most likely assignment to be the Arg66 which shifts on isotopic exchange. In Dronpa, this assignment overlaps exactly with the C=O stretch. In rsKiiro a small shoulder can be seen at 1665 cm^{-1} which is therefore assigned to the Arg66 stretch. This overlapping assignment at 1655 cm^{-1} will have a larger relative amplitude than expected for the isolated HBDI C=O peak which explains the enhanced strength of the ^2H_2O C=O stretch over the off state, ^1H_2O peaks and justifies the disagreement with harmonic frequency amplitudes. Comparing the published cis crystal structure of Dronpa [326] and rsKiiro also shows that the hydrogen bond length of Arg66 is almost identical: 2.76 Å in Dronpa and 2.78 Å (figure 6.7) in rsKiiro which suggests the stretch frequencies should be similar as is seen in figure 6.2.

The trans-neutral (negative) rsKiiro bands are notably weaker in intensity than the corresponding cis anion peaks in. This is consistent with the protonation state assignment to a neutral species as reduced oscillator strengths are also seen in trans-neutral Dronpa [5] and the cis-neutral GFP [321]. The shift in frequency between the cis and trans C=O state is large (36/20 cm^{-1}) but expected as shown by the harmonic frequency calculation which shows a difference of 41 cm^{-1}. A trans-anion is calculated to have a C=O frequency of 1690 cm^{-1} which is very close to the observed 1690 cm^{-1} stretch. However, as previously discussed, the weak negative peaks are characteristic of a neutral species and no evidence for an anionic trans-state is seen in the visible spectra [234].

In this work, the double bleach feature at 1691 and 1675 cm^{-1} are both assigned to a C=O stretch. Lukacs *et al.* [6] assign a similar double bleach in M159T Dronpa also to the C=O but Kaucikas et al. [327] disregarded this based on isotope exchange sensitivity. In rsKiiro, the isotope shifts are minimal, 1 cm^{-1} for both peaks, this is attributed to the exchange of the hydrogen-bonded water molecule seen in the crystal structure of rsKiiro (figure 6.7) which is not apparent in Dronpa [322]. Evidence for a double

C=O assignment is also provided by the crystal structure of rsKiiro where a double conformation of the trans-state is seen. The two conformers show a torsional angle of the phenol-imidazoline of 60° and 24° respectively. In comparison, the off-state crystal structure of Dronpa shows an angle of 22°. These observations can immediately explain the difference in C=C and C=O frequencies between rsKiiro and Dronpa. Theoretical calculations by Kaucikas *et al.* [327] suggest that changes in the twisting angle over methyl bridge alters the bond strength which will shift the C=O and C=C accordingly. The splitting of the C=O band is therefore tentatively assigned to the additional minor conformation of the trans-state in the crystal structure. This also explains some of the modest differences in the off-state peaks between wild-type Dronpa and rsKiiro. In particular, the more pronounced shoulder in the negative peaks for the C=C, Phenol-1 and C=N/C=C could be due to occupancy of the minor state found in the crystallography. It is important to note that the crystal structure was obtained at 100 K and therefore could be a cryogenic artefact. Room temperature crystallography measurements of the trans-state have not shown the same multiple conformations, although as they are lower resolution, they are not as able to resolve alternative conformations if the occupancy is low. For these reasons, we can confidently assign the 1688 cm^{-1} bleach to a ν(C=O) stretch of dominant trans-species, as assigned in wild-type and M159T Dronpa [5, 327] and tentatively assign the 1675 cm^{-1} band to the ν(C=O) minor conformation stretch.

The cis anionic C=C is assigned at 1626 cm^{-1} which is significantly lower than the harmonic frequency calculations which indicate a much stronger bond at 1687 cm^{-1}. Studies have suggested this mode has contributions from the torsional angle of the phenol and imidazolinone rings whereby the twisted structures will weaken and downshift the stretch frequency [5, 320, 327, 328]. As the calculations assume a completely planar chromophore geometry which is not seen in the crystal structures, the downshift of the rsKiiro assignment must be due to weakening of the C=C by a twisted geometry. The same reasoning also explains why the rsKiiro C=C is upshifted by 2 cm^{-1} compared to Dronpa. The crystal structures measure twisting angles of 8° and 5° for Dronpa and rsKiiro respectively, meaning rsKiiro creates a more planar and stronger bond over the methyl bridge, explaining the shift in FTIR difference spectrum.

Mid-IR Transient Absorption

The first two sequential EADS of the pump probe data, with lifetimes 5.78 and 60.65 ps, match the first two decay spectra of the electronic S_1 state of rsKiiro with constants 7.67 and 49.99 ps (table 6.1). Considering the double-well potential confirmed with cryo-spectroscopy in chapter 3 we can assign this biphasic behaviour to simultaneous decay and evolution through these two wells. Compared to Dronpa-2 the excited state kinetics are slower in rsKiiro which is possibly due to the difference in chromophore geometry (figure 6.4) and pH/electrostatic environment [330].

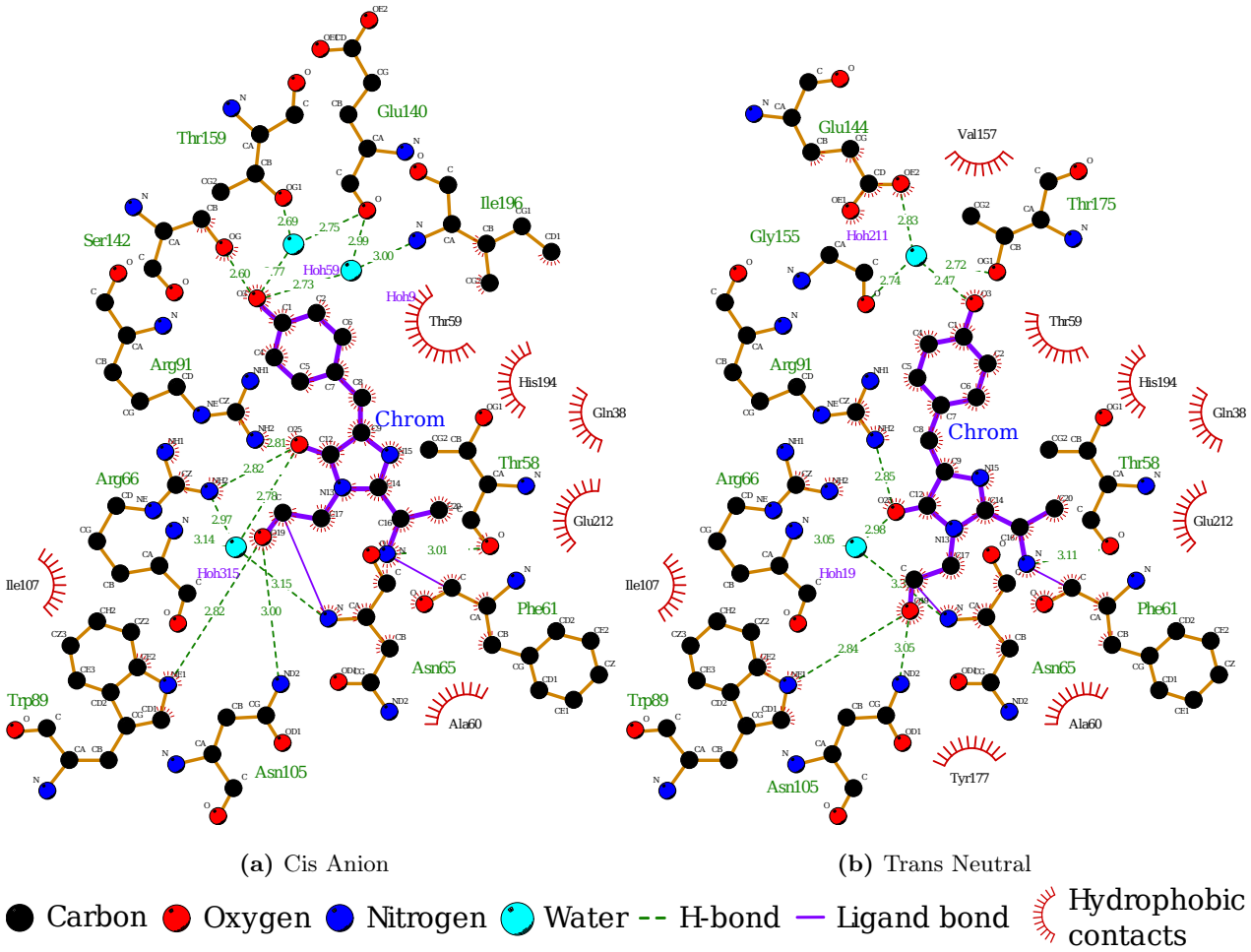


Figure 6.7: (a-d): LigPlot+ [329] plots of crystal structures of rsKiiro in the cis and trans states obtained using cryo-MX. All residues that directly interact with the chromophore are shown. All distances are reported in Å.

The initial assignment of movement of the Arg66 is founded in the simultaneous observation of bands at 1668 and 1596 cm^{-1} which evolve at the same rate and occur at almost exactly the same positions as seen in the steady-state spectra. Given the asymmetric/symmetric modes of the Arg $CN_3H_5^+$ stretch, this assignment is clear. The next EADS in the PP data is assigned to a twisted or a partially isomerised chromophore. The emergence of the 1632 cm^{-1} ν(C=C) peak is indicative of a change in the methyl bridge dihedral angle. This is also supported by the emergence of a further GSB at 1618 cm^{-1} which originates from the phenol-1 stretch of the chromophore. For assignment to a cis-neutral state, a C=O peak should be resolvable. There is a shoulder that corresponds to the 1650 cm^{-1} C=O band which appears to be partially obscured by the C=C GSB at 1641 cm^{-1} however, the assignment to a cis-neutral based on this peak would not be definitive. Instead, it is argued that the chromophore must be at a twisted angle probably close to a cis confirmation as indicated by the ESA assigned to C=C but the exact state is unclear.

A long-lived component then emerges with an ESA band at 1650 cm^{-1} which is assigned to a C=O

stretch. This, combined with the C=C cis-state stretch is strong evidence for the formation of a flattened neutral-cis state probably similar to the cryo-trapped cis state resolved in chapter 3. Kaucikas *et al.* showed through DFT calculations that the C=C and C=O modes of the HBDI chromophore are very sensitive to out-of-plane distortions [219]. Flatter chromophore geometries are expected to show a larger C=C force constant and a weaker C=O bond, resulting in a respective upshift and downshift of the C=C and C=O vibrational modes. However, harmonic frequency calculations of the assigned cis-neutral state predict significant up-shifts in these modes. Although a significant up-shift is not observed a cis-neutral state is still assigned. We argue that the hydrogen bond formed with Arg66 (figure 6.7) reduces the frequency shift between the anionic/neutral forms meaning and instead the flattened geometry dominates the mode behaviour, which overall leads to a downshift in the C=O stretching frequency. This is also supported by work in chapter 7 where we confidently assign a cis-neutral state chromophore in a rsKiiro S142A mutant. In this mutant, a small upshift in the C=O and C=C frequency is observed which is attributed to the hydrogen bond stabilisation by the Arg66. Here the same interaction must account for the lack of cis-neutral upshift in the same two groups due to the neutral species whilst the planar geometry causes a downshift in C=O and upshift in C=C.

Assignment to a cis anion state would imply an excited-state proton transfer mechanism (ESPT) which is not supported by other studies on similar proteins [5–8, 212, 233, 235] or our observations of a µs time constant for the formation of the cis anion 500 nm peak (C. Hutchison, unpublished). Furthermore, the phenol-3 mode (at 1497 cm^{-1} in the steady-state FTIR) is one of the strongest absorption features of the anion but absorbs very weakly in the neutral state (chapter 7). The final 2 ns time point does not show a phenol-3 or phenol-1 peak but the 1 ms spectra does. This is strong evidence that the protonation step takes place between 2 ns - 1 ms. Furthermore, the C=C stretch shifts to the steady-state value of 1626 cm^{-1} indicating the final product has been formed. However, the C=O stretch still seems upshifted, this is possibly due to an unfortunately placed dead-pixel at the 1655 cm^{-1} meaning the shift to the steady-state frequency is not observed.

Overall, the pump-probe measurements of the off-state support an isomerisation of the chromophore to a cis neutral and rearrangement of the Arg66 residue is complete within 65 ps. We do not find any evidence to suggest that ground state isomerisation occurs after the first 100 ps. Further support for this is provided by the measured trans/cis ground-state isomerisation barrier of 91 kJ mol^{-1} (chapter 5). To overcome this barrier a hot ground-state would be required. If a hot-ground state were present in high concentrations hot-bands would be observed at red-shifted frequencies compared to the fundamental transitions due to the anharmonic potential of vibrational transitions [331]. In the PP and PDP spectra the only sign of vibrational cooling is the narrowing and slight blue-shift of the GSB at 1644 cm^{-1} with increasing pump-probe delays.

The pump-dump data shows a definite suppression of the GSB signals by around 50% which is particularly evident at 1644 cm^{-1}. This confirms that the dump pulse is causing stimulated emission from the S_1 excited state. The final 1 ms and 2 ns spectra (figure 6.5) show much lower photoproduct yields as indicated by the much weaker positive difference spectra features. This provides further support for the isomerisation reaction occurring as an excited state process. To compare the photoproduct yields between the PP and PDP data, the ESA amplitude of the phenol-3 bands were compared in the 1 ms spectrum. This band was chosen as it shows clear signals in both the final spectra and is not influenced by any obvious GSB signals which would decrease the amplitude. Considering this band alone, the dump reduces the photoproduct amplitude by around 3 times. However, the amplitude of these signals is small and close to the noise level of the measurement. With the more intense accumulation of the negative time-point, a more accurate impact of the dump pulse on photoproduct yields could be assessed. It should also be noted that the 1 ms spectra is derived from the -40 ps measurement and so depends on the overlap of the previous pump spot with the next pump area. As the 400 nm pulse was slightly large than the 515 nm it is possible this measurement in the dump data set does not include fully dumped population. Overall, consideration of the raw and globally-fitted EADS shows almost no product formation. This is strong evidence that the isomerisation occurs in the excited state and is prevented with the dumping of the excited state population.

6.5 Conclusions

In this chapter steady steady state FTIR difference spectra of rsKiiro have been assigned and shown to resemble the difference spectra of Dronpa. Ultrafast transient absorption measurements revealed the excited state isomerisation of the rsKiiro chromophore occurs with rearrangement of the Arg66 followed by isomerisation within 66 ps. The presence of a dump pulse was then shown to cause stimulated emission and prevent product formation. Future work should concentrate on extracting branching ratios through scanning of the dump time and measurement of vibrational dephasing times through evaluation of the perturbed free induction decay rates. Unfortunately, this was not possible in this work due to the large optical Kerr effect observed at time-zero. Additionally, further time points should be collected beyond 2 ns to study the ground state evolution of the protein to the final cis neutral state.

7 | Neutral to Cation Switching of a GFP-like Protein Resolved Using Dynamic Crystallography and FTIR Spectroscopy

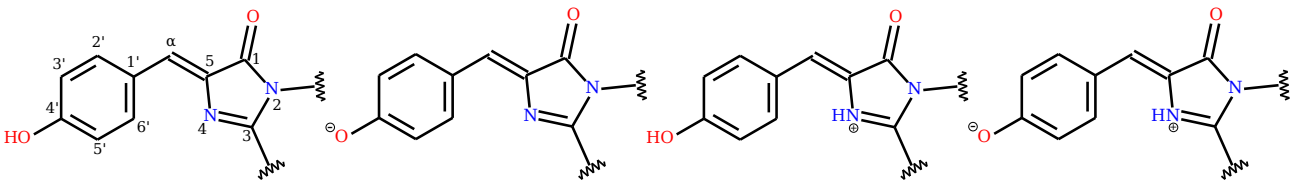
In this chapter, the photoswitching states of a mutant of rsKiiro, known as JvT75 (rsKiiro M159T), is investigated. Time-resolved room-temperature serial synchrotron crystallography, UV/Vis and steady-state FTIR difference are used to assign the conformational and protonation states of the chromophore at neutral and low pHs. At pH 8.4 JvT75 undergoes the same switching behaviour as seen in rsKiiro: 500 nm illumination drives an anionic-cis/trans-neutral reaction with associated fluorescence switching off. Subsequent illumination with 400 nm light then results in the well-studied back-reaction from trans-neutral to cis-anionic. At pH 5.5 protonation of the chromophore leads to a large blue shift in the ground state electronic spectrum and a loss of fluorescence. Illumination with 400 nm light of the low pH species is found to drive a cis-neutral/trans-cationic reaction with high yields which is reversed under 450 nm illumination. This pH dependence of isomerisation mechanism is unreported and seems to be a new class of reaction in reversibly switchable proteins. The work in this chapter extends our understanding of the acid/base equilibrium in fluorescent proteins and presents a well-suited system for studying ultrafast photoswitching in both trans/cis cis/trans directions within a single protein due to high inter-conversion yields at neutral/low pH respectively.

Note to the reader: The serial crystallography experiments in this chapter were led and devised by this author. The beamtime was a collaborative experiment with Diamond Light Source. In particular, Pierre Aller, Agata Butyn, Danny Axford, Robin Owen, Selina Storm, Sam Horrell and Nic Devenish were all instrumental in ensuring the success of these experiments. This work is being prepared for publication as:

> Baxter, J. M. *et al.* Neutral to Cation Switching of GFP-like Protein Resolved Using Dynamic Crystallography and FTIR Spectroscopy *(in preparation).*

7.1 Introduction

The chromophore in GFP-like proteins has two titratable sites, the hydroxyl group at the phenolic site and imidazolinone ring nitrogen at the heterocyclic site. Protonation/deprotonation of these sites leads to the generation of cationic, zwitterionic, neutral and anionic forms of the chromophore, as shown in figure 7.1. The pK_a for protonation of the isolated chromophore, known as 4′-hydroxybenzylidene-2,3-dimethyl-imidazolinone (HBDI) is 2.5 for the cationic species and is 8.5 for deprotonation, generating the anionic state [332, 333]. As mentioned in chapter 2, the strongly fluorescent on state of most RSFPs corresponds to an anionic cis state of the HBDI chromophore and the off, to a neutral trans [199]. There are some notable exceptions in the eqFP578 [334] and qFP611 families [335] which are red FPs and have a strongly fluorescent trans state chromophore. However, the protonation state of the trans-state is unclear and it seems the fluorescence in these proteins originates from a $\pi - \pi$ stacking interaction between the chromophore phenol ring and a nearby histidine [334].



(a) Neutral **(b)** Anionic **(c)** Cationic **(d)** Zwitterionic

Figure 7.1: Possible protonation states of the cis chromophore.

In the drive to understand the fluorescent nature of GFP and the photophysics of RSFPs the isolated HBDI molecule has been the subject of extensive spectroscopic and theoretical studies. It is well known that the reducing flexibility of the HBDI molecule around the methyl bridge promotes fluorescence emission over radiationless decay from the excited state [330, 336]. Numerous theoretical studies have made high-level quantum chemistry calculations of the excited-state potential energy surfaces of the chromophore as summarised by Meech [337]. The key findings are that excitation of the chromophore decreases the bond order over the methyl bridge. This allows an almost barrierless rotation of the double bond between the phenol ring and imidazolinone heterocycle. This twisting motion increases the ground state energy until the S_1 and S_0 surfaces cross through a conical intersection at an angle around 90° [230, 338, 339]. Therefore, experimental observations in solution at room temperature show almost no fluorescence: the HBDI molecule is free to twist and relax down to a 'hot' ground state in a radiationless transition. Fluorescence is only observed in solution with cooling below the glass transition temperature [340, 341] or when increased friction is introduced with viscous solvents [336]. Svendsen *et al.* showed experimentally that in the gas phase, at low temperatures (100 K) the anionic cis HBDI is fluorescent as the excited state population is trapped by a barrier which reduces the rate of internal conversion through the conical intersection [342]. As this barrier extends the S_1

lifetime increased fluorescence yields are observed. These theoretical and experimental studies suggest that fluorescence rates will be highest when HBDI is rigidly confined to planar conformations. In proteins, the HBDI chromophore is stabilised and constrained by hydrogen bonds from surrounding residues within the chromophore environment. A study of HBDI in multiple FPs (Dronpa, rsFastLime, asFP-A143S, mTFP0.7 and IrisFP) has correlated increased planarity of the cis chromophore with the highest fluorescent yields, in agreement with experiments on isolated HBDI [343]. Furthermore, the increased rigidity of the protein matrix surrounding the chromophore has also been associated with greater fluorescence yields - presumably due to reduced ability for twisting-related relaxation processes [337, 344]. It is therefore reasonable to assume that characteristics of the isolated HBDI can be used as a basis to explain the photophysical behaviour in RSFPs.

The first efficient RSFP, Dronpa, has been subject to the largest number of experiments investigating its switching mechanism (as reviewed in previous chapters). Steady-state FTIR spectra assigned the ground, on-state, of Dronpa to the cis-anionic and the on-state to be a neutral trans [5]. Initially, mechanisms of off to on switching proposed an ESPT pathway[213]. However, ultrafast IR transient absorption measurements showed strong evidence that deprotonation was a ground state process in the trans-cis and cis-trans reactions [5–8]. Furthermore, Warren *et al.* and Laptenok *et al.* suggested detailed mechanisms for the steps of isomerisation of the trans-cis reaction (as discussed in chapter 6). Visible-TA measurements by Yadav *et al.* [233] confirmed µs rate constants for protonation/deprotonation observed by Warren *et al.* and Laptenok *et al.* [5, 7] and suggested the chromophore is in an acid-base equilibrium of the anionic and neutral forms. Higashino *et al.* studied Dronpa using Raman spectroscopy. They varied the pH and determined a pK_a of 5.2 for titration of the chromophore and showed it maintained a cis geometry at low pH [345]. Overall, current spectroscopic studies have identified the protonation of the chromophore in the ground-state products of Dronpa. The trans-cis isomerisation mechanism has been the subject of multiple ultrafast studies which have suggested mechanisms for switching in this direction (chapter 6). However, the cis-trans mechanism remains much poorly characterised, primarily as low yields make spectroscopic measurements challenging.

Although Dronpa is the most studied RSFP, GFP is the most widely studied FP, due to its ubiquitous use throughout the life sciences since the 1990s. The evidence for the acid/base equilibrium in Dronpa (described by Higashino *et al.* [345]) is founded in observations in wild-type GFP (avGFP): Fluorescence yields of avGFP are relatively insensitive to changes in pH [196, 346] whilst an S65T mutant showed greatly reduced fluorescence below its pK_a of 6.0 [347, 348]. In avGFP, the Ser 65 donates a hydrogen bond to nearby Glu 222 which promotes the deprotonation of the Glu 222 carboxylate group [349]. The close proximity of this negatively charged Glu 222 to the chromophore then prevents the chromophore and glutamic acid from both being anionic due to electrostatic repulsive forces. Photoacidity of

chromophore excited state results in ESPT which creates the cis anionic chromophore as the phenol proton is shuttled along a series of hydrogen-bonded water molecules to the Glu222 residue [350]. Mutation of S65T (or similarly S65C [197]) withdraws the donation of the hydrogen bond to the Glu 222 which returns it to the neutral state and blocks the ESPT pathway [347]. As a result, the S65T GFP chromophore is in an acid-base equilibrium, mostly occupying the anionic state above pH 6.0 [347, 348]. These assignments were made based on crystal structures of S65T mutant above and below the pK_a [347]. Interestingly, the low and neutral pH structures were very similar in the chromophore structure and differed only in side-chain conformation. In the protein JvT75, studied in this chapter, residue 65 is a cysteine instead of the serine in avGFP. This means JvT75 is likely to follow similar acid/base structural behaviour as the S65T mutant and of Dronpa.

In this chapter, the low and neutral pH behaviour of a mutant of the reversibly switchable fluorescent protein, JvT75 will be investigated. This protein is an S142A mutant of rsKiiro, equivalently an L62A S142A M159T mutant of Skylan-NS which was introduced in chapter 2. Time-resolved room-temperature synchrotron serial crystallography is used to obtain 6 structures of JvT75 at two different pHs and under multiple illumination conditions. These structures are used in conjunction with steady-state FTIR and UV-Vis spectra to identify a new class of photochromic reaction at low pH. The photoactivated isomerisation at pH 8 is found to correspond to a cis-anion/trans-neutral with 400 nm illumination and trans-neutral/cis-anion for illumination with 400 then 500 nm, as is observed in rsKiiro. At pH 5, however, a new class of switching is observed from the neutral-cis to cationic-trans species under 400 nm illumination. To our knowledge, this reaction has not been reported in photochromic proteins. Its discovery improves our understanding of photochromic reactions in proteins and opens up opportunities to understanding the excited state landscape of the cis neutral HBDI chromophore situated in a protein matrix.

7.2 Methods

7.2.1 UV/Vis Steady State Spectra & Thermal Recovery

As explained in detail in section 2.2.1, UV/Vis spectra acquired from 200-1000 nm of the JvT75 protein at around 2 mg/ml concentration at pH 5.5, pH 8.4 and pH 10 in a low salt buffer (50 mM buffering acid, 100 mM NaCl) buffered with 50 mM MES, Tris and HEPES respectively. Samples were converted using either a 405 or 488 nm laser. Thermal recovery measurements were made by pre-converting the sample with a 488 nm laser then measuring spectra every 30 seconds using a 0.5 s exposure time over 8 hours. The exposure time and frequency of measurement were kept to a minimum to reduce the actinic behaviour of the white light probe. Data were fitted and analysed using custom-written MATLAB scripts.

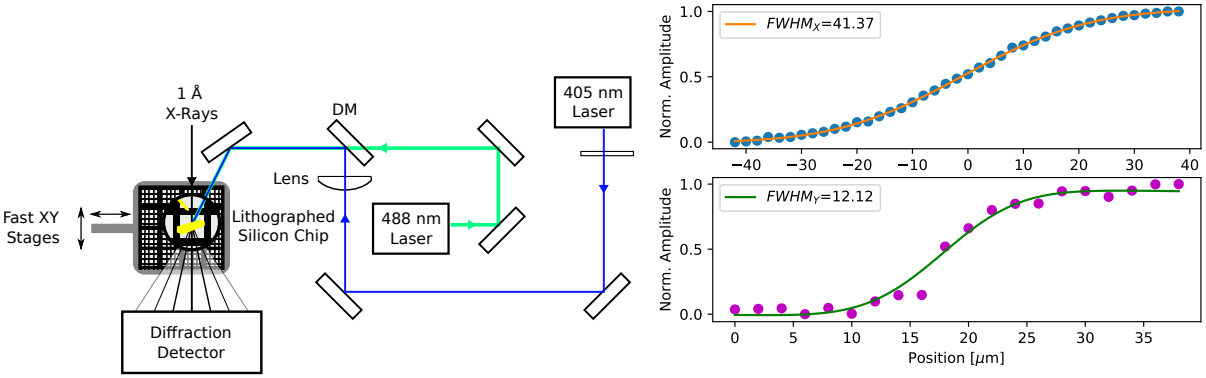
7.2.2 FTIR Difference Spectra

Deuterated samples were prepared at high protein concentrations and placed between two Harrick cell CaF windows as described in detail in 6.2.1. The deuterated buffer was prepared 0.4 pH units below the desired pH to account for the isotope effect in measurement with the pH probe. Spectra were acquired using a Bruker ISF 66/S at 4 cm^{-1} resolution using a liquid nitrogen-cooled MCT-A detector. CW laser illumination with wavelength 405 (500 mW) and 488 nm (100 mW) was directed through a side port of the spectrometer and defocused onto the Harrick cell using a misaligned 1:1 telescope. An LED array focused into a fibre (CoolLED) provided illumination with 435 nm (100 mW) or 460 nm (100 mW) by directing the fibre, unfocused at the sample cell. Illumination times were between 20-30 s to achieve full conversion. Dry air was used to purge the spectrometer for 1-2 hours between opening of the sample box.

7.2.3 Time Resolved Serial Synchrotron Crystallography at Low and Neutral pH

Crystals were prepared in batch conditions as explained in section 2.2.4 at pH 8.4 with PEG3350 concentration of 25% and final protein concentration of 15 $mg\,ml^{-1}$ with a final volume of 500 µl. These conditions formed needle-shaped crystals of dimensions around 2×2×10 µm and density around 10^7-10^8 crystals/ml. Crystals at pH 8.4 were diluted by 8 times before injection. Crystals at pH 5.5 were prepared by centrifuging the slurry at 1 g for 2 minutes followed by removal of 300 µl of mother liquor which was replaced with PEG3350 buffer at pH 5.5. This was repeated 3 times to ensure the pH was close to 5.5. A colour change from bright yellow to dim yellow was immediately observed on exchanging the buffer which indicated the pH was below the pK_a value. Preliminary experiments on larger batch trials showed the final pH was 5.5 ± 0.1 after buffer replacement.

Time-resolved serial synchrotron crystallography (TR-SSX) measurements were made at the Diamond Light Source, i24 beamline. A 100 mW 488 nm unfocused laser with a beam profile around 3×2 mm was directed onto a fixed target serial crystallography chip [127] shown in figure 7.2 (a). Lasers of wavelength 405 and 488 nm laser were combined collinearly using a long pass dichroic mirror (Thorlabs, FEL0450). The 405 nm beam was attenuated with a 1 OD neutral density filter and focused with a lens (300 mm focal length) placed just before the dichroic mirror. This focused the 405 nm beam and left the 488 nm unfocused. TTL modulation was used to set a pulse length of 1 ms for the 405 nm laser. The half-life of the laser rise time was around 600 µs and approximately linear meaning the laser was just below full power by the end of the pulse. The fall time was very short (around 10 µs). The 405 nm laser was focused to a spot size of 12×41 µm as measured by a knife-edge scan (see figure 7.2 (b)) and had a CW power of 21.0 mW at the focus. The preillumination routine consisted of exposing an entire city block to the 488 nm laser for 10 s. The X-ray diffraction data were taken using an X-ray detector exposure time of 10 ms, attenuation of 0 %, wavelength of 0.9801 Å, beam-size of 7×7 µm and



(a) Schematic for fixed target TR-SSX at i24

(b) Fitted knife edge scans of the 405 nm focus

Figure 7.2: (a): Looking down on the beamline setup a 405 and 488 nm laser are combined collinearly on a long-pass dichroic mirror (DM). The 405 nm laser is focused using a 300 mm focal length lens onto the lithographic silicon chip. The chip is translated by 'fast' stages which move between wells at a repetition rate of around 20 Hz. The chip orthogonal to the X-ray, i.e. the plane into the page but is shown here for clarity. Diffraction data are collected by the X-ray detector behind the chip.**(b):** Knife edge scans demonstrating the FWHM focus size of the 405 nm laser.

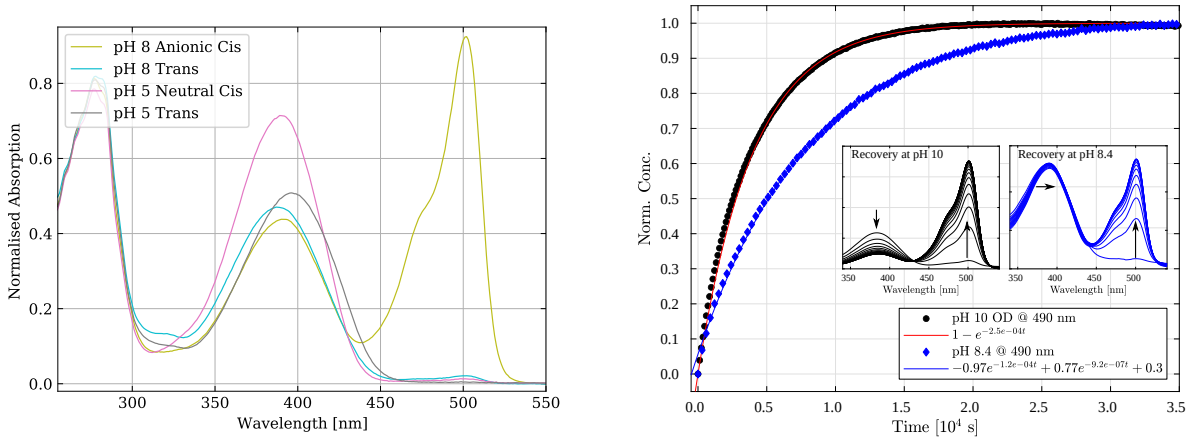
Dectris Pilatus3 6M detector at a distance of 320 mm. Initially both DIALS [163] and CrystFEL [102] using indexing methods XGANDALF [152] and iMOSFLM [284] to index and merge the data. It was found XGANDALF gave the best merging statistics so was used for processing downstream. Detector distance was only coarsely refined in steps of 1 mm as the unit cell distributions and indexing rates did not change significantly with different detector distances. Data from CrystFEL was merged ignoring partialities as detailed in chapter 4.

Structure factors were calculated from intensities using Phenix with the massage option [176]. Q-weighted difference electron density was calculated as detailed in section 3.2.3 and refinement of structures was done using restrained refinement in REFMAC5 [237] and real-space refinement in coot [289]. Omit maps were calculated by setting the chromophore occupancy to 0 and performing '0' rounds of refinement in REFMAC5. Approximate occupancies were calculated as detailed in 5.2.4, through minimisation of crystallographic R-factors. A model containing both cis and trans conformations was generated and the occupancy changed in 1 % intervals from 0 % cis (100 % trans) to 100 % cis (0% Trans) and R-factors calculated after 5 cycles of Rigid body refinement. The minima of R_{Free} and R_{Work} were used to estimate the population of cis/trans conformation in the dataset.

7.3 Results

7.3.1 UV/Vis Steady State Spectra, Fluorescence Measurements & Thermal Recovery

As mentioned in chapter 2, the spectrum of JvT75 dramatically shifts depending on the pH (figure 7.3). **At pH 10**, in the dark/unilluminated the 503 nm band is strongly absorbing and assigned to the cis anionic state as in Dronpa [326]. Illumination with 500 nm generates the 390 nm absorption peak, assigned to a neutral trans which recovers following first-order rate law, as shown in 2.7 (b). **At pH 8.4**, the 490 nm peak dominates although some absorption at 390 nm is seen. This state is also assigned to be cis anionic. Illumination with 500 nm causes a small increase in the intensity of the 390 nm peak and a blue shift of 10 nm to 380 nm. The recovery of this state does not follow a simple 1st order trend, as shown in figure 7.3 (b). **At pH 5.5**, the dark spectrum of the protein changes dramatically. The 503 nm peak is absent and only a 380 nm peak is present which is assigned to a cis neutral state. This absorption profile looks very similar to the pH 8 trans-state but is stronger. Excitation of the cis neutral with 400 nm light causes a bleach in the 380 nm peak and a +10 nm redshift assigned to the formation of a trans conformation. Excitation at 430-450 nm of the trans neutral state at pH 5.5 then reversibly switches back to the original 'dark' state spectra, the cis neutral. The fluorescence emission for the cis neutral and trans states was extremely weak. However, the cis anion state showed appreciable fluorescence with excitation between 450-500 nm. The fluorescence spectra are shown in chapter 2 and exhibit a strokes shift of around 9 nm.



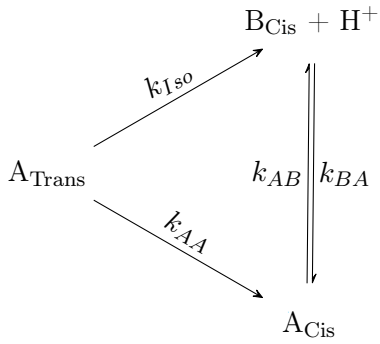
(a) Stationary Spectra

(b) Thermal Recovery

Figure 7.3: **(a)** Stationary UV/Vis spectra of JvT75 in different states. Absorption spectra are plotted on an arbitrary scale, and normalised to each other. **(b)** Thermal recovery rates of the anionic cis peak at pH 8 and 10. The 490 nm peak is seen to recover over time and follows different kinetic rate equations depending on the pH as indicated by the fitted equations in the legend.

A scheme to account for the change in the rate of the thermal recovery is shown below. An initial state

of the neutral trans conformation (A_{Trans}) which can directly isomerise (upon photo-excitation) and deprotonate with rate k_{Iso} to form the anionic cis species (B_{Cis}). An alternative pathway of isomerisation to a neutral cis (A_{Cis}) with rate constant k_{AA} and acid-base equilibrium with forward and back rate constants of k_{AB} and k_{BA} to regenerate the anionic cis state.



The acid base equilibrium constant, K_a can be written in terms of the concentration of the reactants (indicated with square brackets), the buffer pH and the pK_a of JvT75 which was measured in chapter 3 to be 8.17:

$$[A_{Cis}] = \frac{[B_{Cis}][H^+]}{K_a} = \frac{[B_{Cis}]10^{-pH}}{10^{-pK_a}}$$

The rate equations for the scheme suggested above can be written as:

$$\frac{\partial[A_{Trans}]}{\partial t} = -k_{AA}[A_{Trans}] - k_{Iso}[A_{Trans}]$$

$$\frac{\partial[B_{Cis}]}{\partial t} = k_{AB}[A_{Cis}] - k_{BA}[B_{Cis}][H^+] + k_{Iso}[A_{Trans}]$$

$$\frac{\partial[A_{Cis}]}{\partial t} = -k_{AB}[A_{Cis}] + k_{BA}[B_{Cis}][H^+] + k_{AA}[A_{Trans}]$$

$$\text{Substituting: } k_1 = k_{AB}10^{-pH+pK_a} - k_{BA}10^{-pH}$$

$$\frac{\partial}{\partial t}\begin{pmatrix} A_{Trans} \\ A_{Cis} \\ B_{Cis} \end{pmatrix} = \begin{pmatrix} -k_{AA} - k_{Iso} & 0 & 0 \\ k_{AA} & 0 & -k_1 \\ k_{Iso} & 0 & k_1 \end{pmatrix}\begin{pmatrix} A_{Trans} \\ A_{Cis} \\ B_{Cis} \end{pmatrix}$$

Solving for the Eigenvalues gives:

$$\lambda_1 = 0 \text{ (Trivial)} \quad \lambda_2 = k_1 \quad \lambda_3 = -k_{AA} - k_{Iso}$$

With the corresponding Eigenvectors:

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad v_3 = \begin{pmatrix} -(k_1 + k_{AA} + k_{Iso})/k_{Iso} \\ (k_1 + k_{Iso})/k_{Iso} \\ 1 \end{pmatrix}$$

Defining $k_{Obs} = k_{AA} + k_{Iso}$ this gives a solution of the form:

$$[B_{Cis}](t) = \alpha_0 + \alpha_1 \exp(k_1 t) + \alpha_2 \exp(-k_{Obs} t)$$

In this solution, k_1 acts as a sigmoidal function which only becomes significant at pHs close to and below the pK_a. This means in the high pH limit (pH 10) the rates will follow first order rate laws whilst this will not be seen at pH 8.4. Fitting the pH 8.4 thermal recovery gives values of 0.3, 0.77, 9.2e-7 s^{-1}, 0.97, -1.2e-4 s^{-1} for α_0, α_1, k_1, α_2 and k_{Obs}. The pH 10 data was fitted by using the high pH limit of:

$$[B_{Cis}](t) = \alpha_0 + \alpha_1 + \alpha_2 \exp(-k_{Obs} t) = \beta + \alpha_2 \exp(-k_{Obs} t)$$

With coefficients of 0.99, -0.93 and 1.2e-4 s^{-1} for β, α_2 and k_{Obs}.

7.3.2 FTIR Difference Spectra Assignment

Note to the reader: The difference spectra plotted in this section follow the convention set in chapter 6 and by Warren *et al.* [5], Kaucikas *et al.* [8] Lukacs *et al.* [6] and Laptenok *et al.* [7] for probing the 'back' reaction. Therefore, positive peaks in the FTIR difference spectrum refer to the unilluminated state whilst the negative peaks refer to the photoproduct (against usual convention).

The initial approach for the assignment of difference spectra of JvT75 was with direct comparison to the Dronpa stretching frequencies reported by Warren *et al.* [5]. Figure 7.4 shows a comparison of difference spectra of Dronpa and JvT75 in ^2H_2O at pD 7.8 and 8.4 respectively. A majority of spectral features are conserved between the two proteins although differences are seen particularly in the fingerprint region (1280 - 1380 cm^{-1}) and in the carbonyl region (between 1640-1680 cm^{-1}). Overall, the assignments support a neutral trans and anionic cis chromophore in the off and on state respectively at pH/pD 8.4, table 7.1 and 7.2 summarise the assigned peaks. Local minima in the on state ^2H_2O spectrum at 1655, 1630, 1574, 1547, 1499 and 1150 cm^{-1} are assigned to be characteristic of ν(C=O), ν(C=C), Phenol-1, ν(C=N/C=C), Phenol-3 and phenol respectively. Correspondingly, assignment of the off state follows: 1680, 1651, 1616, 1562, 1514 and 1177 cm^{-1} for modes of ν(C=O), ν(C=C), Phenol-1, ν(C=N/C=C), Phenol-3 and Phenol.

Assignments for the pD 5.5 spectra can be made with comparison to the JvT75 pD 8.4 difference spectra seen in figure 7.5. Immediately it is clear that the spectra share similar features although there are clear differences in the intensity of the carbonyl stretching region (1600-1700 cm^{-1}), and a clear decrease in intensity of the phenol-3 peak around (+)1499 cm^{-1} compared to the pD 8.4 spectrum. Confident assignment can be made to the dark state: 1572, 1551, 1509 and 1155 cm^{-1} for the Phenol-1, C=N/C=C, Phenol-3 and Phenol modes. The 400 nm illuminated state (figure 7.6) assignments for this region are clear with peaks at 1611, 1565, 1516, 1176 cm^{-1} for the Phenol-1, C=N/C=C, Phenol-3

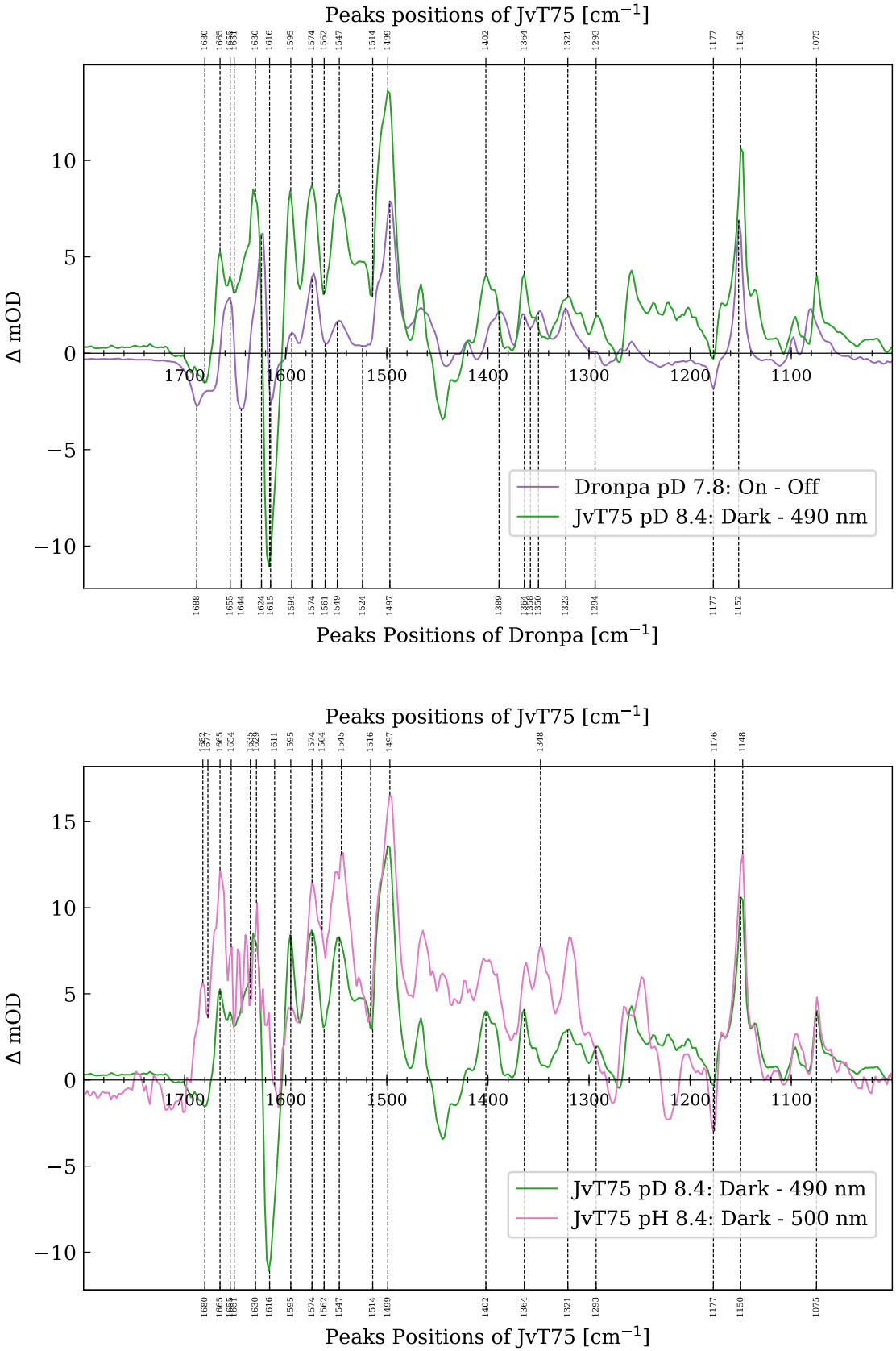


Figure 7.4: Steady-state FTIR difference spectra for JvT75 and the wildtype Dronpa [5]. Spectra for JvT75 are plotted at pH/pD 8.5 , illumination conditions are indicated in the legend. Prominent peaks discussed in the text are labelled on the top and bottom axes as indicated by the dashed lines.

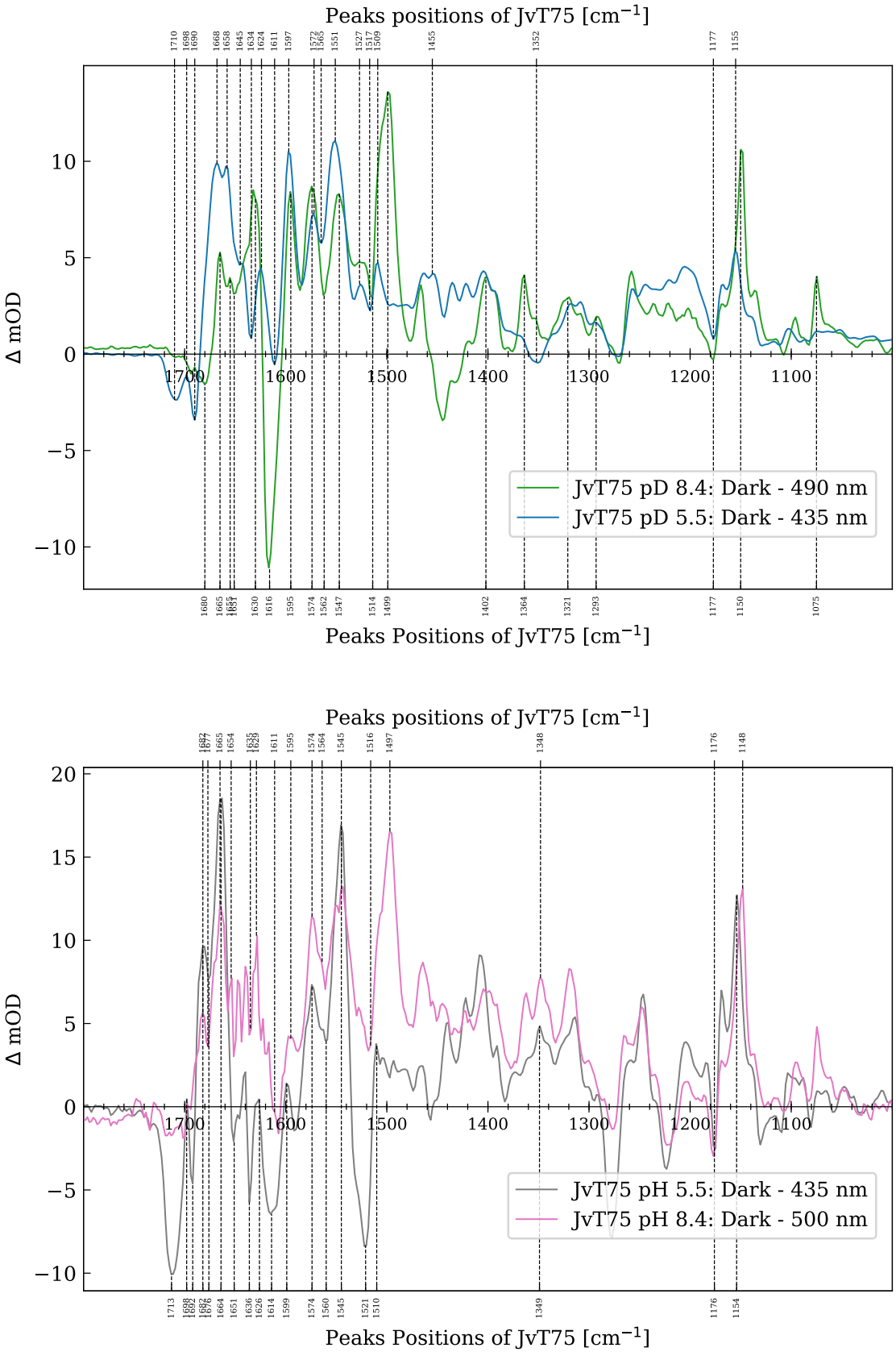


Figure 7.5: Steady-state FTIR difference spectra for JvT75 at pD/pH 8.4 and 5.5, illumination conditions are indicated in the legend. Peaks discussed in the text are labelled on the top and bottom axes as indicated by the dashed lines.

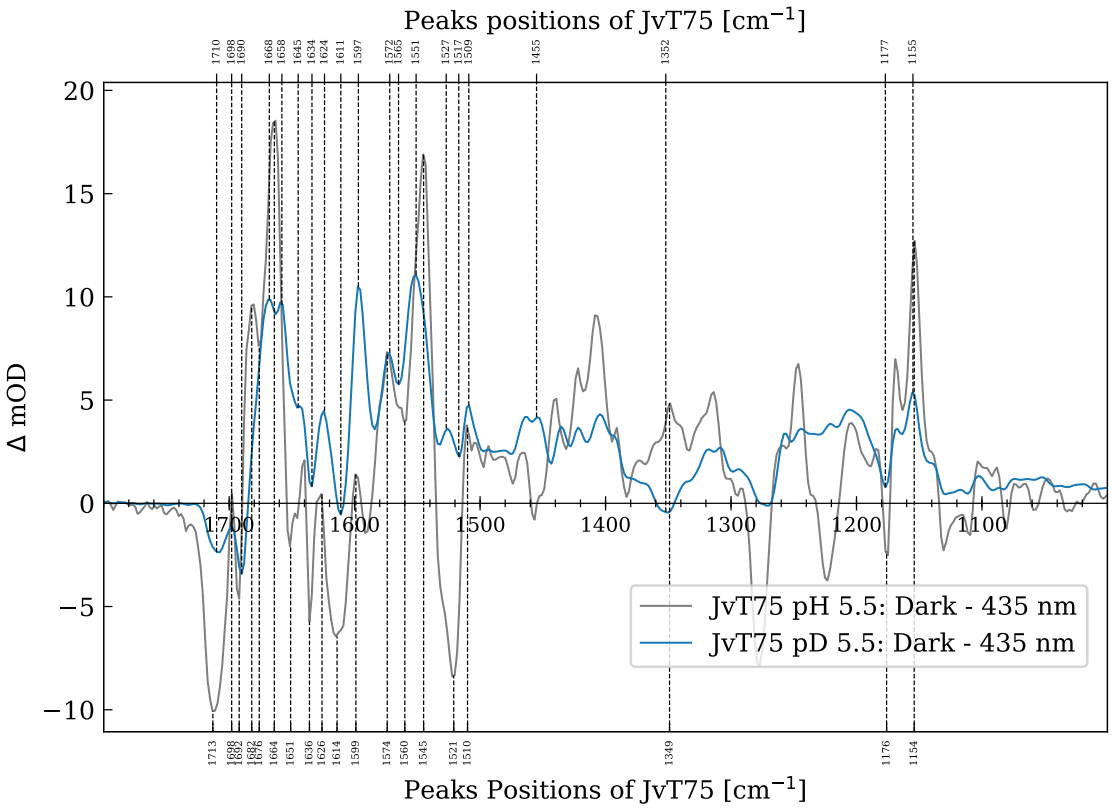


Figure 7.6: Steady-state FTIR difference spectra for JvT75 at pD/pH 5.5, illumination conditions are indicated in the legend. Peaks discussed in the text are labelled on the top and bottom axes as indicated by the dashed lines.

and Phenol modes. These modes are notably blue shifted compared to the pD 8.4 peaks. Larger shifts are seen in the carbonyl stretching region of the chromophore where peaks at 1624 and 1658 cm^{-1} are assigned to the unilluminated state for ν(C=C) and ν(C=O) vibrations. The illuminated (negative) state peak at 1645 cm^{-1} is assigned to the ν(C=C) bands. A double bleach feature similar to the peaks at 1692 and 1710 cm^{-1} (pD 8.4) were assigned to C=O by Lukas *et al.* [6] whilst Warren *et al.* [5] and Kaucikas *et al.* [8] do not make the same assignment based on isotope (^{1}H/^{2}H) exchange shifts which were inconsistent with harmonic frequency calculations [5, 8]. In JvT75 at pD/pH 5.5 (figure 7.6) no ^2H_2O isotope shift is seen for the 1698 cm^{-1} positive peak whilst the negative peaks at 1690 and 1710 cm^{-1} (pD 8.4) shift by +2 and +3 cm^{-1} after ^{1}H substitution. As the isotope shift is small both the 1690 and 1710 cm^{-1} are assigned to the C=O stretch. Overall the assignment in the low pH data suggest a neutral cis conformation of the chromophore in the dark/unilluminated state (positive features). Whilst illumination with 400 nm light is assigned to isomerisation to a cationic trans state.

The peaks at 1668/1682 cm^{-1} in the $^2H_2O/^1H_2O$ spectra (figure 7.6) were tentatively assigned to the Arg 66 $\nu_{asym}(CN_3H_5^+)$ stretch at pD/pH 5.5. At pD/pH 8.4, peaks at 1665/1682 cm^{-1} $^2H_2O/^1H_2O$ peaks are equivalently assigned to the Arg 66 $\nu_{asym}(CN_3H_5^+)$ (figure 7.4). The small shift in the ^2H_2O

peak is potentially caused by a prominent side peak but gives some uncertainty to the assignment. The $\nu_{sym}(CN_3H_5^+)$ is then assigned as 1595/1595/1597/1599 cm^{-1} in the ^2H_2O(pD 8.4)/^1H_2O(pH 8.4)/^2H_2O(pD 5.5)/^1H_2O(pH 5.5) difference spectra.

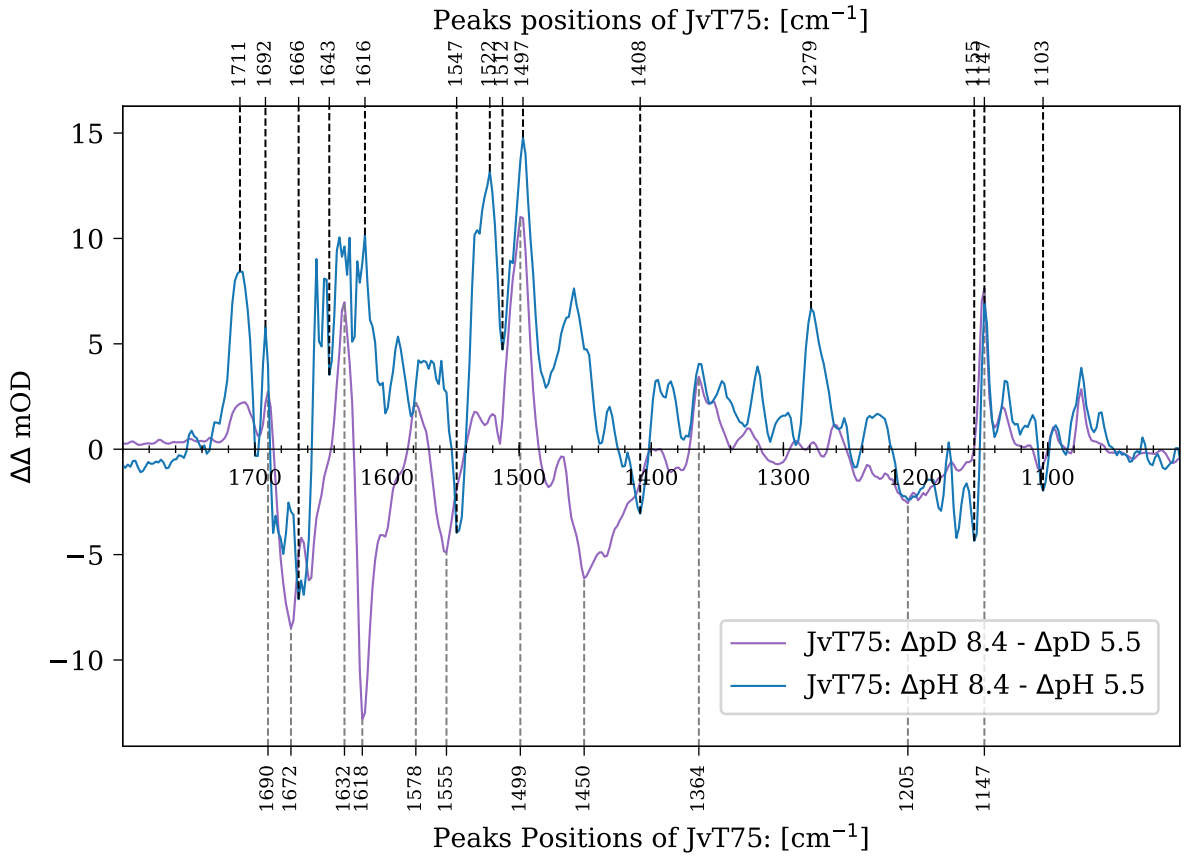


Figure 7.7: Steady state FTIR double difference spectra for JvT75. Shown between 1800-1000 cm^{-1} for pD/pH 8.4 and 5.5. Photoconversion with 500 and 435 nm is referred to in the legend.

Dominant Component	JvT75 pD 8.4	JvT75 pH 8.4	JvT75 pD 5.5	JvT75 pH 5.5
C=O	1655	1665	1658	1664
Arg 66 asym($CN_3H_5^+$)	1665	1682	1668	1682
C=C	1630	1629	1624	1626
Arg 66 sym($CN_3H_5^+$)	1595	1595	1597	1599
Phenol-1	1574	1574	1572	1574
Phenol-2	-	-	-	-
C=N/C=C	1547	1549	1551	1545
Phenol-3	1499	1497	1509	1510
Phenol	1150	1148	1155	1153

Table 7.1: Dark state assignments of JvT75 at pD/pH 5.5 and 8.4. All values in cm^{-1}.

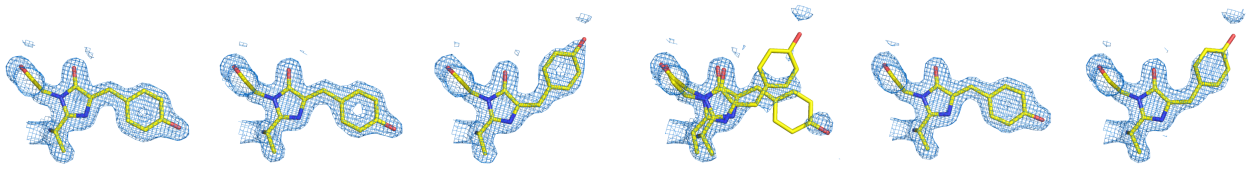
Dominant Component	JvT75 pD 8.4	JvT75 pH 8.4	JvT75 pD 5.5	JvT75 pH 5.5
C=O	1680	1676	1710/1690	1713/1692
C=C	1651	1651	1645	1651
Phenol-1	1616	1607	1611	1614
C=N/C=C	1562	1560	1565	1560
Phenol-3	1514	1520	1516	1521
$^{13}C_5$ Sensitive	N.D.	1375	N.D.	N.D.
Phenol	1177	1176	1176	1177

Table 7.2: Photoconverted state assignments of JvT75 at pD/pH 5.5 and 8.4. All values in cm^{-1}.

7.3.3 Crystallography

In total, 61,864 crystal diffraction images were merged into 6 datasets for conditions: pH 8.4 dark, pH 8.4 500 nm illumination, pH 8.4 400 nm only illumination, pH 8.4 500 followed by 400 nm illumination (pre-conversion followed by a flash), pH 5.5 dark and pH 5.5 400 nm. The limiting resolution was found to be 1.75 Å as determined by a value of 0.5 in Fourier shell correlation [313]. Merged statistics, in table A.18 and A.19 show the refined unit-cell parameters for each condition varied less than 0.9 % in volume. This should minimise the scaling required in the calculation of isomorphous Fourier difference electron density meaning accurate maps can be calculated using a single set of 'dark' phases. Refinement gave coordinates with R_{Work}/R_{Free} values of 0.19-0.20/0.22-0.23 which is around the 50th percentile of all submitted PDB refinement statistics for structures between 1.65-1.85 Å. This indicates that refinement was complete and successful as statistics are appropriate for the applied resolution limits. Merging statistics, such as $CC_{1/2}$ and R_{Split}, were all appropriate for TR-SSX data and maintained similar

values for all conditions.



(a) pH8 Dark **(b)** pH8 400 nm **(c)** pH8 500 nm **(d)** pH8 500/400 nm **(e)** pH5 Dark **(f)** pH5 400 nm

Figure 7.8: Chromophore omitted $2F_o$-F_c maps of serial crystallography data under different illumination conditions. Positive density is shown in blue, plotted at 3σ level.

The $2F_o$-F_c chromophore omit-density, shown in figure 7.8, immediately suggests a cis conformation in the dark state at pH 8.4 and pH 5.5, as is consistent with cryo-MX structures (chapter 2) and TR-SSX measurements (at pH 8.4) of rsKiiro. The pH 8.4 data with 500 nm illumination converts the cis conformation to a trans-state. As is typical in negatively switching FPs, 400 nm light then reconverts the chromophore back towards the cis state as seen in the pH 8.4 500/400 nm map. Additionally, 400 nm illumination only of the pH 8.4 state (with no 500 nm preillumination) causes no change in electron density and the cis conformation is maintained. However, at pH 5.5, 400 nm illumination shows switching towards the trans-state with high yields. Occupancy refinement of this condition is shown in figure A.13 and summarised in table 7.3. Here, the highest level of structural homogeneity in seen in the pH 5.5 Dark state at 96/73 (4/27) % cis (trans) occupancy for the R/R_{Free} respectively. Population transfers calculated from the occupancy refinement suggest the low pH cis-trans isomerisation reaction occurs with higher yields than the neutral pH trans-cis. Looking at the R-factor minima, population transfers of 30 and 17% are determined from these reactions at pH 5.5 and 8.4 respectively.

It was concluded in chapter 5 that R-factor minimisation was less accurate than the difference density minimisation method for occupancy determination. As the structures obtained in this chapter are heterogeneous in population it is challenging to define a 'pure' structure for use other occupancy determination methods. Previously, (chapter 3), we have used cryo-MX to obtain initial coordinate models for these states. However, for JvT75 at low pH, this was not possible as the tris-buffer pH is temperature dependant [351]. This makes protonation during cooling possible meaning it is challenging to capture a low pH 'pure' structure. Therefore, room temperature SSX is one of the only ways to accurately resolve these structures.

Q-weighted difference electron density (DED) maps are shown in figure 7.9, and their 1-dimensional representation is shown in figure 7.10. Both support the assignments made by studying the omit maps, clearly showing the cis-trans difference density signal for the pH 8.4 500 nm and pH 5.5 400 nm data (when compared with the equivalent dark structures). The shape and position of the DED signals

State	Cis Occupancy at R_{Free} min (%)	Cis Occupancy at R min (%)
pH 8.4 Dark	69	81
pH 8.4 500 nm	46	27
pH 8.4 500 then 400 nm	48	44
pH 5.5 Dark	96	73
pH 5.5 400	63	43

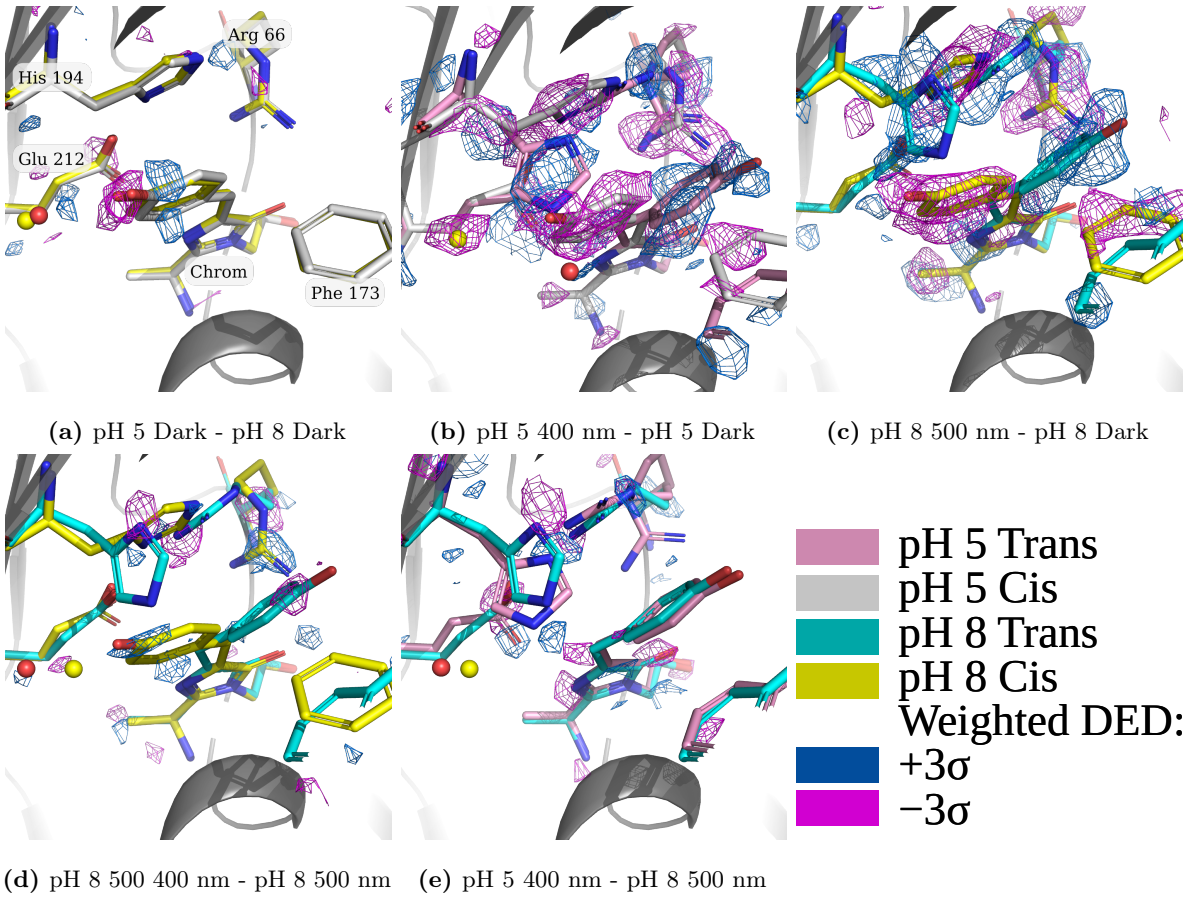
Table 7.3: Occupancy populations of the cis chromophore refined coordinates for JvT75 in different conditions.

suggest that the conformational changes of the protein at neutral and low pH are very similar despite the different illumination schemes. The pH 8.4 500-400 nm differences support the conclusions of the occupancy refinement; it shows weaker signals than the pH 5.5 dark - 400 nm map which implies the yields of photoisomerisation are lower. This is also supported by the DED signal distribution in the 1D difference maps which, are stronger in the pH 5.5 dark - 400 nm than the pH 8.4 500 nm - 500-400 nm (figure 7.10). This increases confidence in the conclusion that the yield of cis-trans photolysis is higher in the pH 5.5 condition compared to the trans-cis at pH 8.4.

DED calculations are extremely sensitive to changes in structure, meaning they can be used to support/confirm atomic movements which are difficult to assign when only using refined coordinates. There are modest differences in the chromophore phenol ring angles between the pH 5.5 and the pH 8.4 cis conformation which are highlighted by the signals in the difference density (figure 7.9 (a)). Water 134, 111 and 128 form part of the hydrogen bonding network to the chromophore phenol at pH 8 as seen in figure 7.11 (a). At low pH water 111 is not seen. This alteration to the hydrogen bonding network is evidence for the cis neutral state as opposed to the cis anion if the the second lone-pair of the oxygen is bonded with the proton instead of donated to the H bond.

Looking at the 1D integrated electron density of the pH8 Dark - pH5 Dark map in figure 7.10, strong signals are seen in the region of residues 142-143, 146-147 and 201. These are all residues that are associated with the crystal contacts between protein molecules. The strongest signals are seen on residue 201, a proline, which is shown in refined coordinates to shift by 0.9 Å. This movement is imparted onto the opposing crystal contacts at residues numbers 142-147 which justifies the strong DED at these residues. In the pH 8 500 nm - pH 5 400 nm similar signals are seen but they are also polluted by differences in cis/trans population so are not as distinguishable in figure 7.10.

Structural changes at neutral and low pHs of the trans-state are highlighted by the difference density in 7.9 (e) and coordinate plots in figure 7.11. Movement of histidine 194 and a slight shift downwards of the chromophore is seen in the low pH DED which is reflected in the extrapolated coordinates.



(a) pH 5 Dark - pH 8 Dark **(b)** pH 5 400 nm - pH 5 Dark **(c)** pH 8 500 nm - pH 8 Dark

(d) pH 8 500 400 nm - pH 8 500 nm **(e)** pH 5 400 nm - pH 8 500 nm

Figure 7.9: Q-weighted difference electron density maps of different states of JvT75 plotted in magenta and blue at 3σ level. Shown in purple, grey, cyan and yellow are the refined coordinates for pH 5.5 trans, pH 5.5 cis, pH 8.4 trans and pH 8.4 cis respectively. The secondary structure of the β barrel is shown in dark grey. Residues around the chromophore (numbers: 173, 194 and 201) associated with isomerisation are also shown.

The negative density on the imidazolinone ring nitrogen is potentially indicative of a protonation or increased delocalisation caused by the formation of a hydrogen bond with the Glu212 residue, only seen in the low pH structure. The absence of this bond at pH 8.4 is contrary to Dronpa where the trans-state chromophore forms a hydrogen bond with the equivalent residue, Glu211 at pH 7.8 [218]. The 1D signals between the low and neutral pH trans states (pH 8.4 500 nm - pH 5.5 400 nm in figure 7.10) are indicative of differences in cis/trans populations *and* changes to the crystal contacts which were also seen in the dark structures at low/neutral pH. Firstly, this indicates the trans occupancy between these conditions is slightly different as is supported by the occupancy refinement results in table 7.3, with 57 and 73 % trans population calculated for pH 5.5 400 nm and pH 8.4 500 nm respectively. Secondly, the signals around residues 142-147 indicate the trans structure shows similar crystal contact rearrangements to those seen in the cis structure although the differences are weaker, perhaps because the phases/scaling are of lower accuracy.

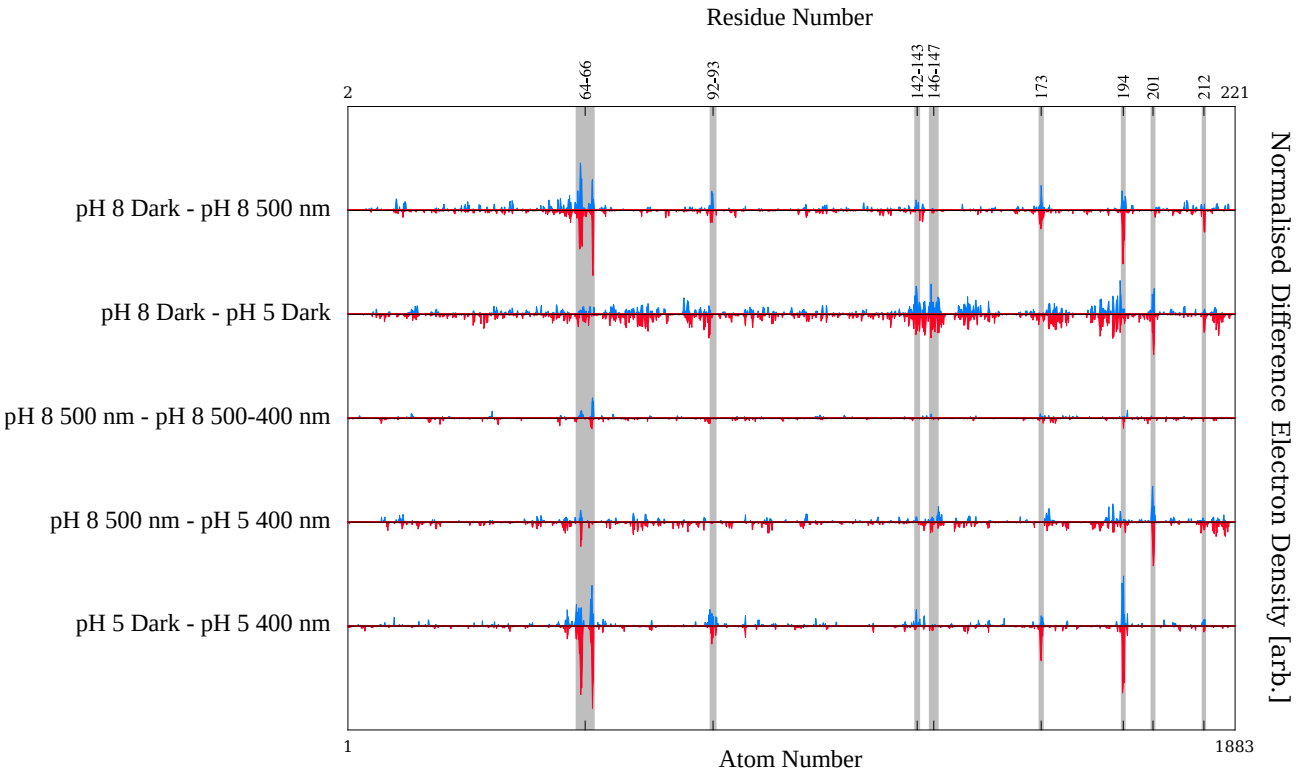
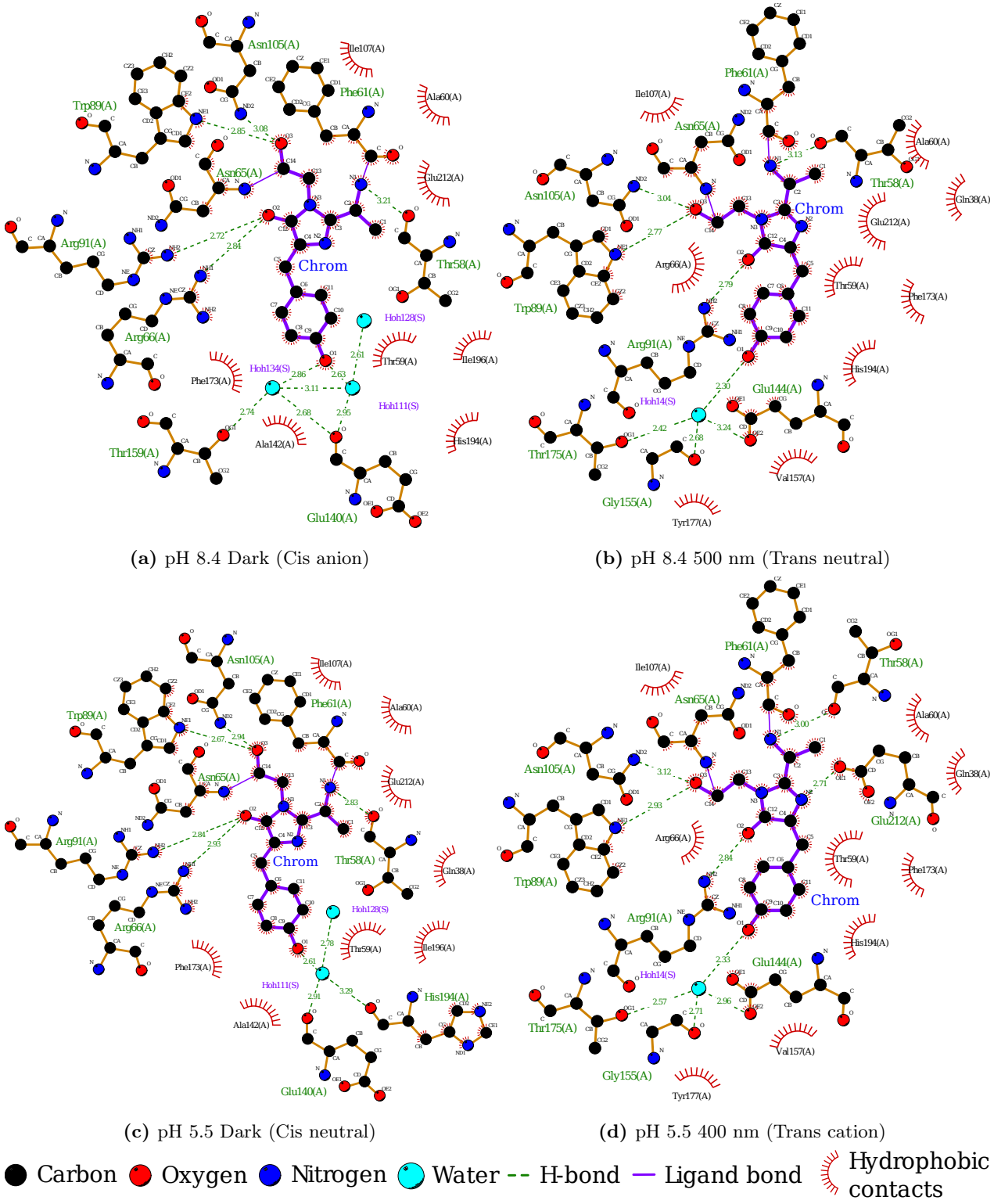


Figure 7.10: 1-Dimensional representation of integrated difference electron density within 2 Å at 3σ level along each atom of the protein chain, all plotted on the same scale (normalised difference electron density). Highlighted on the top axis are residue numbers of interest around the chromophore and crystal contacts as discussed in the main text. Code to produced this figure is based on the published tool by Wickstrand *et al.* [250].



(a) pH 8.4 Dark (Cis anion)

(b) pH 8.4 500 nm (Trans neutral)

(c) pH 5.5 Dark (Cis neutral)

(d) pH 5.5 400 nm (Trans cation)

Figure 7.11: (a-d): LigPlot+ [329] plots showing the hydrogen bonding environment (green lines) around the chromophore, interacting residues, spectator residues and distances (in Å) for different crystal structures.

7.4 Discussion

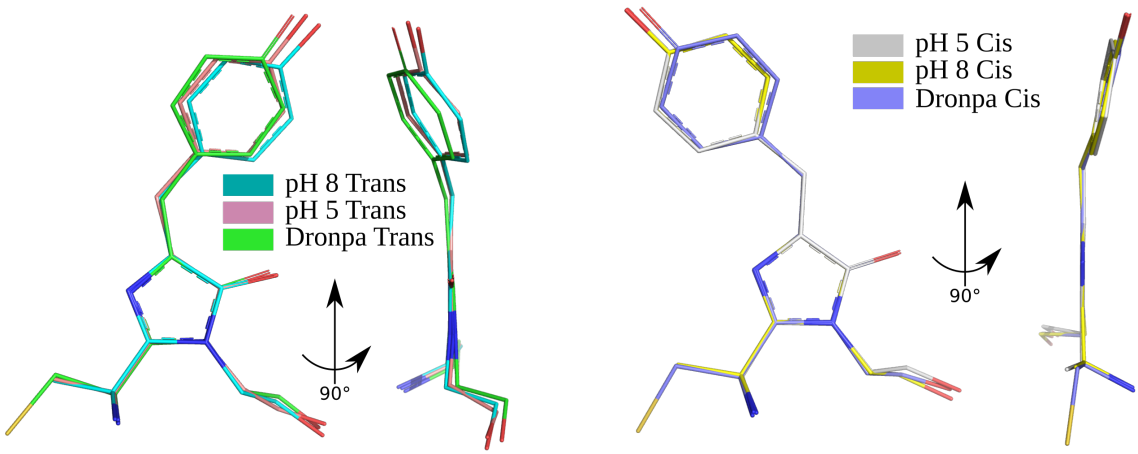
JvT75 shows cis-trans isomerisation in both acidic and neutral pH with similar atomic coordinate modification but distinctively different visible excitation requirements and spectral features. At low pH, this isomerisation occurs reversibly with a change in protonation state to a trans-cation, a new class of reaction not previously observed in reversibly switchable fluorescent proteins (RSFPs). At neutral pH the isomerisation occurs from an anionic cis to a neutral trans which is also reversible. This stark change in protonation state is intriguing and suggests that a protonation change is required for isomerisation, as is the conclusion of a number of theoretical studies [232, 352, 353], although these studies are on different proteins and therefore chromophore environments. The ultrafast switching mechanism of negative switching FPs in the Anthozoa family is debated and early steps of the isomerisation reaction remain unclear from current ultrafast spectroscopy measurements, especially in the forward direction (cis-trans) [5, 7, 8, 233]. The high yields of the cis-trans reaction at pH 5.5 relative to the neutral pH trans-cis reaction are possibly caused by increased internal conversion rates, evidence by the quenching of fluorescence. These characteristics make JvT75 a unique target for gaining insight into excited state motions of the cis-trans reaction which is poorly studied in negative-type switching FPs.

The shifts in visible spectra and the thermal recovery rates modelling both suggest an acid-base equilibrium of the chromophore cis state as suggested in Dronpa [345]. While the protonation state changes are confirmed by the FTIR, the TR-SSX confirms the isomerisation state and shows the reaction is complete after 10 ms. This is expected as the slowest rate constants in similar RSFPs, rsEGFP2, Dronpa, and IrisFP, are assigned to the deprotonation step which is on µs timescales [5, 7, 233, 235]. Confirmation of the visible spectral assignments is seen in the crystal structures. The blue-shifted 380 nm visible absorption is clearly due to a cis conformational as seen in the pH 5.5 dark coordinates whilst the 390 nm peak, caused by 400 nm illumination of the pH 5.5 species is clearly due to the trans-state. The pH 5.5 trans peak is red-shifted compared to the pH 8.4. This is most likely due to a combination of the change in protonation state of the trans conformer and stabilisation caused by the hydrogen bond formed with Glu212 (figure 7.11) which is not seen at pH 8.4.

	Ψ_{Trans}	Φ_{Trans}	Φ_{Cis}
Dronpa	131°	30°	8 °
JvT75 pH 8.4	120°	40°	-3 °
JvT75 pH 5.0	121°	42°	1 °

Table 7.4: Out of plan bending angle, Φ and dihedral angle of the phenol ring torsion, Ψ of the crystallographic chromophore coordinates of Dronpa and JvT75 as seen in figure 7.12.

The protonation states of the pH 8.4 and pH 5.5 dark condition are confirmed by the FTIR difference



(a) Trans **(b)** Cis

Figure 7.12: Chromophore alignment from the crystal structures of JvT75 at pH 5.5, pH 8.4 and Dronpa. The Dronpa structures were obtained from the protein data bank with IDs: 2POX (trans) and 2IOV (cis).

spectra which support an anionic-cis and neutral-cis states respectively. The anionic-cis state assignment is supported by similar visible and IR spectral features in the FPs of Dronpa and GFP [5, 219, 323, 324]. The cis neutral assignment is supported by the large blue shift in electronic spectra and agreement of harmonic frequency calculations (table A.15 [5]) with the difference spectra. Calculations of the HBDI cis neutral chromophore [5] (table A.15) suggest weakened absorption at the phenol-3 cis peak, a stronger C=C band and an upshifted C=O stretch which is seen in the pH/pD 5.5 dark - 400 nm difference spectra. This is strong evidence for the neutral protonation state of the cis chromophore at pH 5.5 as opposed to the cis-anionic at pH 8.4. Furthermore, a change in hydrogen bonding is seen around the chromophore in the cis low pH crystal structure. A donating Hydrogen-bond network through Thr159 is disrupted presumably because the neutral chromophore phenol cannot support the extra hydrogen bond, unlike the anion. A cis-zwitterionic state, with protonation of the imidazolinone nitrogen, is calculated to show strong Phenol-2 absorption as an on-state peak (table A.15). This is not observed in the difference spectra meaning the assignment of a zwitterionic state would be unjustified.

Overall, the assignments of the photo-converted species (negative peaks) at pH 8.4 (500 nm illumination) and pH 5.5 (400 nm illumination) are to a neutral-trans and cationic-trans respectively. At pH 8.4, assignment of the trans cationic state would be unprecedented as both the visible spectra and FTIR difference spectra very closely resemble the trans neutral species in Dronpa [326]. The assignment at pH 5.5 is less clear and will be discussed in here: A slight red-shift of the UV peak from 380 (pH 8.4 trans) to 390 nm (pH 5.5 trans) is potentially characteristic of a cationic trans-species as similar shifts are seen in the electronic spectra of the HBDI chromophore in solution. In HBDI, peaks at 385, 395 and 450 nm are known to correspond to the anionic, cationic and neutral chromophore [333]. Additionally, protonation of the chromophore would most likely be on the nitrogen N4 of the imidazolinone (figure7.1

(c)). Glu212 forms a strong Hydrogen bond (length 2.71 Å) with this nitrogen, as seen in figure 7.11 (d), which suggests a change in the protonation state of the chromophore or the Glu212 residue. Interestingly, a very similar Glu212 donating interaction has been suggested as an intermediate in the FP Kindling by Grigorenko *et al.* [354].

The strongest evidence for a trans cationic assignment at low pH is the agreement of harmonic frequency calculations [5] and observed bands relating to the C=O stretching modes. Calculations predict an up-shift in ν(C=O) of +54 cm^{-1} between neutral and cationic trans-state HBDI molecule. In this study, a shift of +10 cm^{-1} is observed between the neutral and low pD measurements. Some shifts are expected if the chromophore twisting angles differ significantly [8], however, the TR-SSX structures show the trans conformations are very similar (figure 2.3 and table 7.4) meaning any large shifts in chromophore stretching frequencies are unlikely to be caused by changes to the chromophore geometry. We can therefore confidently conclude that considering the restrained and slightly distorted trans geometry of the chromophore (compared to harmonic calculations [5]), the shift of 10 cm^{-1} in ν(C=O) between pD 8.4 to pD 5.5 is due to the formation of a cationic trans-state at low pH. Furthermore, harmonic frequency calculations (table A.15 [5]) suggest the cationic trans should have downshifted ν(C=C), Phenol-1, Phenol-2, ν(C=N/C=C) and Phenol-3 modes. The assignments to these modes all follow this trend except the Phenol-3 stretch which is unexpectantly upshifted compared to the trans neutral assignment.

An alternative assignment for the (−)1710 cm^{-1} peak at pD 5.5 was considered: the only protein stretching modes between 1700-1710 cm^{-1} (in ^2H_2O) is a stretch at 1706 cm^{-1} from a Glu C=O mode [325]. This *could* be interpreted as a signal originating from the Glu 212, indicating deprotonation/proton donation occurs with cis-trans isomerisation of the chromophore. However, several reasons make this assignment unlikely. Firstly, the 1713/1710 cm^{-1} ^2H_2O/^1H_2O shift is too small compared to the expected shift of +6 cm^{-1} given by Bathy *et al.*. Secondly, an additional strong asymmetric (ν_{asym}(COO$^-$))stretch would also be expected close to 1567 cm^{-1} in ^2H_2O and between 1556-1560 cm^{-1} for ^1H_2O. Deprotonation of the illuminated state would imply protonation of the dark state so an upshifted positive Glu(COOH) stretch would also be expected. However, none of these spectral features are seen. Finally, the protonation state of Glu212 (Glu211/Glu222 in Dronpa/GFP) is predicted to be neutral at pH 8 as it is involved in the regulation of the proton transfer mechanism. In avGFP, Glu222 maintains an H bond through serine 65 which stabilises the anionic state glutamic acid which acts as a proton acceptor from the photoacidic chromophore. In Dronpa and GFP S65T, however, there is no donating hydrogen bond to Glu211/222 so the neutral state is maintained. Since JvT75 is most-similar to Dronpa a neutral Glu212 is most likely maintained at pH 8. If the Glu212 is neutral at pH 8 it almost certainly does not undergo deprotonation at pH 5. Therefore, the 1710 cm^{-1} band must

originate from the chromophore and is very similar to the signals assigned in chapter 6 for rsKiiro so is assigned to an additional C=O stretch.

Comparison of the refined chromophore geometry of the off/on states of Dronpa crystals and low pH/neutral pH, cis/trans in JvT75 crystals (figure 7.12) reveals subtle changes to coordinates which also explains the frequency shifts in carbonyl and alkene IR absorption bands in JvT75 compared to Dronpa. The shifts of the cis state spectrum can be attributed to the slight change in C−C_α=C angle and the change in the dihedral angle of the phenol ring torsion (Ψ_{cis}) from 8° to −3°/1° for Dronpa and JvT75 pH 8.4/5.5 as seen in figure 7.12 and table 7.4. The frequency shifts in trans coordinates are larger as the out of plane stretching angle (Φ_{Trans}) and Ψ_{Trans} change by around 10 °compared to Dronpa. Kaucikas *et al.* show that decreased bending of the phenol (towards a planar geometry) blue shifts the C=O and C=C vibrational frequency as does decreasing the Ψ angle. Given the more planar geometry of the JvT75 crystal structure these calculations, therefore, support the blue shift of the C=O and C=C bands in the FTIR spectra of JvT75 at pD/pH 8.4 in comparison to Dronpa. The shifts at low pH, are too large to be explained by a distorted geometry and so supports protonation state assignment of a cation.

The fingerprint region differs between Dronpa and JvT75 at pD 8.4. Both share dominant peaks at 1364 cm^{-1} and around 1321 (1323) cm^{-1} and 1293 (1294) cm^{-1} but a shifted peak is seen at 1402 (1389) cm^{-1} for JvT75 (Dronpa). Similar signals are seen in the fast-switching Dronpa mutant (M159T) were Kaucikas *et al.* attribute this to skeletal deformations of the chromophore which may be associated with an equilibrium of the chromophore in a non-planar geometry [8]. The JvT75 spectra at pD 5.5 show similar changes in the fingerprint region to the pD 8.4 species which are presumably also due to slight deformations of the chromophore. However, a new negative peak at 1352 cm^{-1} is also seen which is unassigned as is the loss of the peak at 1075 cm^{-1} seen in the pD 8.4 data.

The 1-Dimensional DED between the pH 8 dark and pH 5.5 structures shows difference signals throughout the protein structure, as seen in figure 7.13. The pK_a value for each residue in the pH 8.4 dark structure was calculated using the program PROPKA3 [355] which empirically models the pK_a for each titratable residue in the protein. It was proposed that structural differences could be due to changes in the protonation state of particular residues which would mean difference density would correlate with pK_a values. Figure 7.13, shows relative pK_a plotted with 1D difference density signals. It shows there is very little correlation between the calculated pK_a and difference density signal which perhaps indicates that changes in pH affect solvation channels through the structure instead of causing the formation of charged residues.

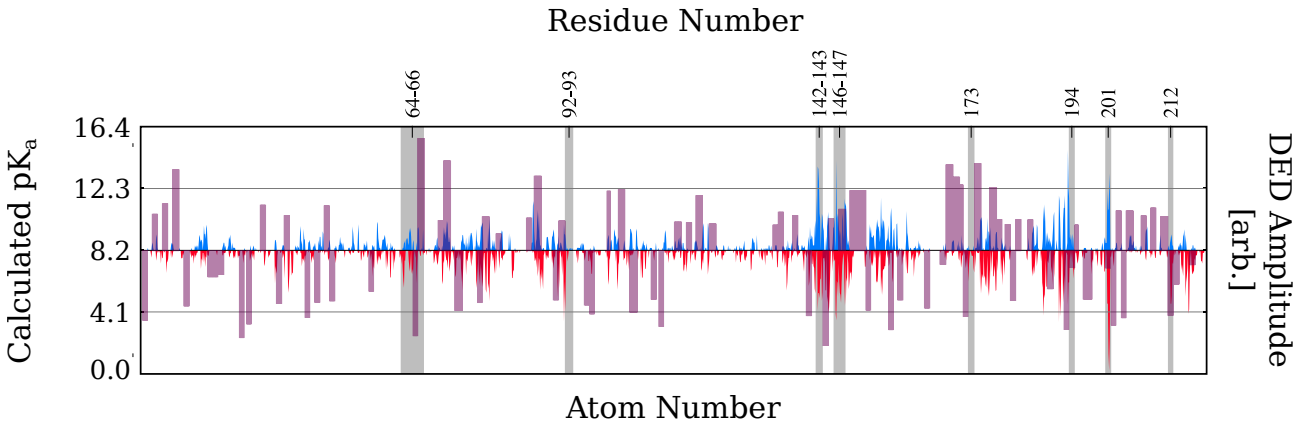


Figure 7.13: Calculated pK_a values for titratable cites using PROPKA3 [355] are plotted on the left axis for each residue in the protein chain (top axis label). Residues without titratable groups do not show a bar in the plot. Plotted on the right axis is the 1 dimensional difference electron density between the pH 8 Dark and pH 5 Dark structures in arbitrary units integrated up to 2 Å and above 3σ level along each atom of the protein chain. Code to generate this figure is based on the published tool by Wickstrand *et al.* [250]. Atom number refers to the PDB file sequential numbering of all atoms in the protein structure.

7.5 Conclusions

Overall, the experiments in this chapter have extensively characterised a new RSFP reaction dynamic and provided ground-state structures, visible and infrared assignments. These high-resolution crystal structures along with spectral assignments will provide a strong foundation for future quantum chemistry calculations of the cis-neutral/trans-cation isomerisation which is currently unstudied. Furthermore, ultrafast spectroscopy using IR-TA complemented by Vis-TA may allow insight into a new excited-state surface of the chromophore and give mechanistic insight into this new reaction type. Further work could also be done to characterise the ground state activation energy barriers. Especially as one of the key findings of chapter 5 was the large magnitude of heating in TR-SSX experiments. Although the spectroscopy and crystallography are extremely complimentary, the yield differences at neutral/low pH could be due to heating over disproportionately sized ground state barriers at the different pHs and reaction coordinates. Hence the determination of these barriers is important.

8 | Conclusions

Overall, this thesis has presented research from a number of crystallography and spectroscopy experiments that have improved out understanding of the switching mechanisms Anthozoa-type RSFPs. In chapter 2, mutagenesis was used to engineer RSFPs with improved switching yields and crystallography. It was found that the L62A mutation increased switching rates across all Eos/Skylan-NS based mutants. In particular, the protein rsKiiro showed the best properties for study with TR-SFX and so was taken forward for further experiments. Recent work by S. Boxer and coworkers [224, 231] concluded that electrostatics heavily influence the chromophore spectrum. It may therefore, be possible to gain further insight effect of each point mutation by calculating the electrostatic effect of amino acid substitutions around the chromophore pocket. Molecular dynamics simulations could estimate the steric effects of mutations which may explain the changes in switching rates between the proteins tested.

Chapter 3 measured the Arrhenius behaviour of rsKiiro and resolved intermediate and steady-state structures on the ground state. The steady-state structures confirmed cis/trans isomerisation with illumination along with rearrangement of three amino acids in the chromophore region. The cryo-trapped structures indicated that the final step in the switching reaction (at low temperatures) was a slight relaxation of the chromophore which supported excited state isomerisation. It would be interesting to extend the spectral region and sensitivity of these measurements the monitor the 400 nm peak, fluorescence, and possibly vibrational modes as function of temperature. This may allow a more in-depth analysis of the ground state surface. It would also be interesting to study other proteins, such as JvT75 or rsEGFP2 [192] to see if similar behaviour is observed. Given changes in pK_a observed in JvT75 we may observe different thermodynamic behaviour.

Having established the steady-state conformation of rsKiiro, chapters 4 and 6 then resolved the ultrafast dynamics in pump-probe and pump-dump-probe schemes spatially, using TR-SFX and spectrally, using vibrational spectroscopy. Interestingly, pump-probe TR-SFX measurements could not resolve an excited-state isomerisation product before 100 µs, presumably because the photoproduct is too disordered to generate detectable light-induced structural differences. The structure resolved was similar to the trans cryo-trapped structure in chapter 3. Pump-probe TA measurements suggested the Arg 66

was one of the initial movements, followed by the chromophore isomerisation which occurred within the electronic excited-state lifetime (∼70 ps). The pump-dump-probe TR-SFX measurements allowed us to assign the observed pump-probe structure to a ground-state structure. This was supported by the pump-dump-probe IRTA measurements which unambiguously show the isomerisation occurs as an excited state reaction with the suppression of photoproduct formation with a dump-pulse at a 380 fs delay to the pump-pulse. There are a number of opportunities to further the work of these chapters. Extending similar multi-pulse excitation schemes to TR-SFX experiments on similar RSFPs or other systems such as retinal-proteins [269, 270, 273] or phytochromes [188] could yield important information for assignment of ultrafast intermediate structures and mechanisms.

Improving sample delivery methods, brighter X-ray sources and better detectors means TR-SSX is a viable and developing method for studying protein dynamics in crystals [28]. In chapter 5, rsKiiro was used to demonstrate precise optical control of photolysis yields, resolved using TR-SSX and to estimate the magnitude of heating by X-rays using the fixed target delivery mechanism. We resolve a temperature increase of around 90 K which could be extremely detrimental for studying temperature dependent reactions. Further work should focus on determining temperature increases as a function of radiation dose rate and using different sample delivery mechanisms. It may be possible to extend the studies by Warren *et al.* [307] and imply temperature changes from the measurement of the stokes shift in ruby microcrystals. The would allow higher accuracy and certainty in the measured of the induced heating effects.

Chapter 7 investigates the protonation states of the protein JvT75 using TR-SSX and steady-state FTIR spectroscopy. We find that above the pK_a of 8.2, reversible photoconversion occurs between a cis-anionic and trans-neutral species. At low pH, however, the photoconversion switches between a cis-neutral/trans-cationic. The observation of this behaviour is not only novel with respect to all other switching behaviours reported for RSFPs but also has implications for understanding the mechanism of numerous other photoswitchable fluorescent proteins. In particular, the assignment of the 390 nm absorption band to a cis-anionic species will allow for more comprehensive interpretation of visible TA spectra of rsEGFP2 and rsKiiro. The JvT75 sample is also an extremely suitable target for pump-probe and multipulse TA measurements on both the cis/trans and trans/cis directions which would be underpinned by the steady state assignment made in this chapter. Additionally, the higher yields of the trans/cis reaction and excellent crystallography are well suited to ultrafast TR-SFX measurements.

Overall, this thesis has achieved the aim of improving the understanding of switching mechanisms in RSFPs. We have presented evidence of new photoswitching behaviour and concluded that RSFP photoisomerisation (in the case of rsKiiro) occurs in the excited state. Furthermore, results from the first two-colour excitation and optical control experiment resolved using TR-SFX have been presented.

These results, along with characterisation of the heating effects of TR-SSX, have also raised important questions on the limitations of time-resolved crystallography for the study of protein dynamics. Hopefully, future work can focus on a protein system with slightly more biological significance. s

Bibliography

[1] Christopher D M Hutchison, Violeta Cordon-preciado, Rhodri M L Morgan, Takanori Nakane, Josie Ferreira, Gabriel Dorlhiac, Alvaro Sanchez-gonzalez, Allan S Johnson, Ann Fitzpatrick, Clyde Fare, Jon P Marangos, Chun Hong, Yoon Id, Mark S Hunter, Daniel P Deponte, S Boutet, Shigeki Owada, Rie Tanaka, Kensuke Tono, So Iwata, and Jasper J Van Thor. X-ray Free Electron Laser Determination of Crystal Structures of Dark and Light States of a Reversibly Photoswitching Fluorescent Protein at Room Temperature. *Int. J. Mol. Sci.*, 18:1918, 2017.

[2] Xi Zhang, Mingshu Zhang, Dong Li, Wenting He, Jianxin Peng, Eric Betzig, and Pingyong Xu. Highly photostable, reversibly photoswitchable fluorescent protein with high contrast ratio for live-cell superresolution microscopy. *Proceedings of the National Academy of Sciences*, 113(37):10364–10369, sep 2016.

[3] M. Andresen, M. C. Wahl, A. C. Stiel, F. Grater, L. V. Schafer, S. Trowitzsch, G. Weber, C. Eggeling, H. Grubmuller, S. W. Hell, and S. Jakobs. Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proceedings of the National Academy of Sciences*, 102(37):13070–13074, 2005.

[4] Xin X Zhou and Michael Z Lin. Photoswitchable fluorescent proteins: ten years of colorful chemistry and exciting applications. *Current Opinion in Chemical Biology*, 17(4):682–690, 2013.

[5] Mark M. Warren, Marius Kaucikas, Ann Fitzpatrick, Paul Champion, J. Timothy Sage, and Jasper J. van Thor. Ground-state proton transfer in the photoswitching reactions of the fluorescent protein Dronpa. *Nature Communications*, 4(1):1461, dec 2013.

[6] Andras Lukacs, Allison Haigney, Richard Brust, Kiri Addison, Michael Towrie, Gregory M. Greetham, Garth A. Jones, Atsushi Miyawaki, Peter J. Tonge, and Stephen R. Meech. Protein photochromism observed by ultrafast vibrational spectroscopy. *Journal of Physical Chemistry B*, 117(40):11954–11959, 2013.

[7] Sergey P. Laptenok, Agnieszka A. Gil, Christopher R. Hall, Andras Lukacs, James N. Iuliano, Garth A. Jones, Gregory M. Greetham, Paul Donaldson, Atsushi Miyawaki, Peter J. Tonge, and

Stephen R. Meech. Infrared spectroscopy reveals multi-step multi-timescale photoactivation in the photoconvertible protein archetype dronpa. *Nature Chemistry*, 10(8):845–852, 2018.

[8] Marius Kaucikas, Martijn Tros, Jasper J Van Thor, Jasper J. van Thor, Jasper J Van Thor, and Jasper J. van Thor. Photoisomerization and Proton Transfer in the Forward and Reverse Photoswitching of the Fast-Switching M159T Mutant of the Dronpa Fluorescent Protein. *The Journal of Physical Chemistry B*, 119(6):2350–2362, feb 2015.

[9] Villy Sundström, Tõnu Pullerits, and Rienk Van Grondelle. Photosynthetic light-harvesting: Reconciling dynamics and structure of purple bacterial LH2 reveals function of photosynthetic unit. *Journal of Physical Chemistry B*, 103(13):2327–2346, 1999.

[10] Tomáč Polívka and Villy Sundström. Ultrafast dynamics of carotenoid excited states-from solution to natural and artificial systems. *Chemical Reviews*, 104(4):2021–2071, 2004.

[11] Jyotishman Dasgupta, Renee R. Frontiera, Keenan C. Taylor, J. Clark Lagarias, and Richard A. Mathies. Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proceedings of the National Academy of Sciences*, 106(6):1784–1789, feb 2009.

[12] R. Mathies, C. Brito Cruz, W. Pollard, and C. Shank. Direct observation of the femtosecond excited-state cis-trans isomerization in bacteriorhodopsin. *Science*, 240(4853):777–779, may 1988.

[13] Arieh Warshel. Bicycle-pedal model for the first step in the vision process. *Nature*, 260(5553):679–683, apr 1976.

[14] G. E. Busch, M. L. Applebury, A. A. Lamola, and P. M. Rentzepis. Formation and decay of prelumirhodopsin at room temperatures. *Proceedings of the National Academy of Sciences of the United States of America*, 69(10):2802–2806, 1972.

[15] Riccardo A.G. Cinelli, Vittorio Pellegrini, Aldo Ferrari, Paolo Faraci, Riccardo Nifosì, Mudit Tyagi, Mauro Giacca, and Fabio Beltram. Green fluorescent proteins as optically controllable elements in bioelectronics. *Applied Physics Letters*, 79(20):3353–3355, 2001.

[16] Christopher D. Bostick, Sabyasachi Mukhopadhyay, Israel Pecht, Mordechai Sheves, David Cahen, and David Lederman. Protein bioelectronics: a review of what we do and do not know. *Reports on Progress in Physics*, 81(2):026601, feb 2018.

[17] Katalin V. Korpany, Pinky Langat, Dong Myeong Kim, Neil Edelman, Daniel R. Cooper, Jay Nadeau, and Amy Szuchmacher Blum. Conductance switching in the photoswitchable protein dronpa. *Journal of the American Chemical Society*, 134(39):16119–16122, 2012.

[18] Villy Sundström, Key Words, Villy Sundström, Key Words, and Villy Sundström. Femtobiology. *Annual Review of Physical Chemistry*, 59(1):53–77, 2008.

[19] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006.

[20] M. G. L. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. SHORT COMMUNICATION. *Journal of Microscopy*, 198(2):82–87, may 2000.

[21] Lothar Schermelleh, Peter M. Carlton, Sebastian Haase, Lin Shao, Lukman Winoto, Peter Kner, Brian Burke, M. Cristina Cardoso, David A. Agard, M. G. L. Gustafsson, Heinrich Leonhardt, and John W. Sedat. Subdiffraction Multicolor Imaging of the Nuclear Periphery with 3D Structured Illumination Microscopy. *Science*, 320(5881):1332–1336, jun 2008.

[22] Samuel T. Hess, Thanu P.K. Girirajan, and Michael D. Mason. Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. *Biophysical Journal*, 91(11):4258–4272, dec 2006.

[23] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, oct 2006.

[24] Alan G. MacDiarmid. "Synthetic metals": A novel role for organic polymers (Nobel lecture). *Angewandte Chemie - International Edition*, 40(14):2581–2590, 2001.

[25] Elke De Zitter, Daniel Thédié, Viola Mönkemöller, Siewert Hugelier, Joël Beaudouin, Virgile Adam, Martin Byrdin, Luc Van Meervelt, Peter Dedecker, and Dominique Bourgeois. Mechanistic investigation of mEos4b reveals a strategy to reduce track interruptions in sptPALM. *Nature Methods*, 16(8):707–710, 2019.

[26] Allen M. Orville. Recent results in time resolved serial femtosecond crystallography at XFELs. *Current Opinion in Structural Biology*, 65:193–208, 2020.

[27] Allen M Orville. Entering an era of dynamic structural biology.... *BMC biology*, 16(1):55, 2018.

[28] Arwen R. Pearson and Pedram Mehrabi. Serial synchrotron crystallography for time-resolved structural biology. *Current Opinion in Structural Biology*, 65:168–174, 2020.

[29] D. M. P. Holland. The Daresbury Synchrotron Radiation Source. *Physica Scripta*, 36(22):23–25, jan 1987.

[30] U W Arndt, R A Crowther, and J F W Mallett. A computer-linked cathode-ray tube microdensitometer for X-ray crystallography. *Journal of Physics E: Scientific Instruments*, 1(5):303, may 1968.

[31] Mariusz Jaskolski, Zbigniew Dauter, and Alexander Wlodawer. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS Journal*, 281(18):3985–4009, 2014.

[32] State of RCSB PDB. Technical report, 2016.

[33] Author J R Helliwell. Synchrotron X-ray crystallography techniques: time-resolved aspects of data collection. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 340(1657):221–232, aug 1992.

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

[35] P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F.-J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, Ph Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H.-D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky, and J. Galayda. First lasing and operation of an ångstrom-wavelength free-electron laser. *Nature Photonics*, 4(9):641–647, sep 2010.

[36] Henry N. Chapman, Petra Fromme, Anton Barty, Thomas A. White, Richard A. Kirian, Andrew Aquila, Mark S. Hunter, Joachim Schulz, Daniel P. DePonte, Uwe Weierstall, R. Bruce Doak, Filipe R. N. C. Maia, Andrew V. Martin, Ilme Schlichting, Lukas Lomb, Nicola Coppola, Robert L. Shoeman, Sascha W. Epp, Robert Hartmann, Daniel Rolles, Artem Rudenko, Lutz Foucar, Nils Kimmel, Georg Weidenspointner, Peter Holl, Mengning Liang, Miriam Barthelmess, Carl Caleman, Sébastien Boutet, Michael J. Bogan, Jacek Krzywinski, Christoph Bostedt, Saša Bajt, Lars Gumprecht, Benedikt Rudek, Benjamin Erk, Carlo Schmidt, André Hömke, Christian Reich, Daniel Pietschner, Lothar Strüder, Günter Hauser, Hubert Gorke, Joachim Ullrich, Sven Herrmann, Gerhard Schaller, Florian Schopper, Heike Soltau, Kai-Uwe Kühnel, Marc Messerschmidt, John D. Bozek, Stefan P. Hau-Riege, Matthias Frank, Christina Y. Hampton, Raymond G. Sierra, Dmitri Starodub, Garth J. Williams, Janos Hajdu, Nicusor Timneanu, M. Marvin Seibert, Jakob Andreasson, Andrea Rocker, Olof Jönsson, Martin Svenda, Stephan Stern, Karol Nass, Robert Andritschke, Claus-Dieter Schröter, Faton Krasniqi, Mario Bott, Kevin E. Schmidt, Xiaoyu Wang,

Ingo Grotjohann, James M. Holton, Thomas R. M. Barends, Richard Neutze, Stefano Marchesini, Raimund Fromme, Sebastian Schorb, Daniela Rupp, Marcus Adolph, Tais Gorkhover, Inger Andersson, Helmut Hirsemann, Guillaume Potdevin, Heinz Graafsma, Björn Nilsson, and John C. H. Spence. Femtosecond X-ray protein nanocrystallography. *Nature*, 470(7332):73–77, 2011.

[37] Zhirong Huang. Brightness and coherence of synchrotron radiation and FELs. *Proceedings of IPAC 2013*, 02-A06:16–20, 2013.

[38] E A Seddon, J A Clarke, D J Dunning, C Masciovecchio, C J Milne, F Parmigiani, D Rugg, J C H Spence, N R Thompson, K Ueda, S M Vinko, J S Wark, and W Wurth. Short-wavelength free-electron laser sources and science: a review. *Reports on Progress in Physics*, 80(11):115901, nov 2017.

[39] Gianluca Geloni, Evgeni Saldin, Evgeni Schneidmiller, and Mikhail Yurkov. Transverse coherence properties of X-ray beams in third-generation synchrotron radiation sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 588(3):463–493, apr 2008.

[40] IOP. Synchrotron light. *Institute of Physics*, 2010.

[41] Edgar Weckert. The potential of future light sources to explore the structure and function of matter. *IUCrJ*, 2:230–245, 2015.

[42] Kwang Je Kim. Angular distribution of undulator power for an arbitrary deflection parameter K. *Nuclear Inst. and Methods in Physics Research, A*, 246(1-3):67–70, may 1986.

[43] Anja Burkhardt, Tim Pakendorf, Bernd Reime, Jan Meyer, Pontus Fischer, Nicolas Stübe, Saravanan Panneerselvam, Olga Lorbeer, Karolina Stachnik, Martin Warmer, Philip Rödig, Dennis Göries, and Alke Meents. Status of the crystallography beamlines at PETRA III. *European Physical Journal Plus*, 131(3), 2016.

[44] A. Meents, M. O. Wiedorn, V. Srajer, R. Henning, I. Sarrou, J. Bergtholdt, M. Barthelmess, P. Y.A. A Reinke, D. Dierksmeyer, A. Tolstikova, S. Schaible, M. Messerschmidt, C. M. Ogata, D. J. Kissick, M. H. Taft, D. J. Manstein, J. Lieske, D. Oberthuer, R. F. Fischetti, and H. N. Chapman. Pink-beam serial crystallography. *Nature Communications*, 8(1), 2017.

[45] P. KIRKPATRICK and A. V. BAEZ. Formation of optical images by X-rays. *Journal of the Optical Society of America*, 38(9):766–774, 1948.

[46] David Flot, Trevor Mairs, Thierry Giraud, Matias Guijarro, Marc Lesourd, Vicente Rey, Denis Van Brussel, Christian Morawe, Christine Borel, Olivier Hignette, Joel Chavanne, Didier Nurizzo,

Sean McSweeney, and Edward Mitchell. The ID23-2 structural biology microfocus beamline at the ESRF. *Journal of Synchrotron Radiation*, 17(1):107–118, 2010.

[47] Robert Hettel. Perspectives and Challenges for Diffraction Limited Storage Ring Light Sources. In *5th International Particle Accelerator Conference*, number 5, pages 0–4, 2014.

[48] Ivan A Vartanyants and Andrej Singer. *Synchrotron Light Sources and Free-Electron Lasers*. Springer International Publishing, 2019.

[49] Donald H Bilderback, Pascal Elleaume, and Edgar Weckert. Review of third and next generation synchrotron light sources. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 38(9):S773–S797, may 2005.

[50] E. C. Longhi, R. Bartolini, R. Fielder, I. Martin, J. C. Schouten, and B. Singh. First year's experience of diamond insertion devices. In *EPAC 2008 - Contributions to the Proceedings*, number January, pages 2285–2287, 2008.

[51] Gerhard Materlik, Trevor Rayment, and David I Stuart. Diamond Light Source: Status and perspectives, 2015.

[52] Alan D. LeGrand, Claude Pradervand, Wilfried Schildkamp, Dominique Bourgeois, and Michael Wulff. An ultrafast mechanical shuttering system for separating single x-ray pulses from a synchrotron. *Review of Scientific Instruments*, 66(2):1721–1721, feb 1995.

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

[54] John M.J. Madey. Stimulated emission of bremsstrahlung in a periodic magnetic field. *Journal of Applied Physics*, 42(5):1906–1913, 1971.

[55] Takashi Tanaka. Current status and future perspectives of accelerator-based x-ray light sources. *Journal of Optics*, 19(9):093001, sep 2017.

[56] M Khorasaninejad, A Zhu, C Roques-Carmes, W T Chen, J Oh, I Mishra, and R C Devlin. Polarization-insensitive metalenses at visible wavelengths. *Work Under Review*, pages 1–15, 2016.

[57] Mike Dunne. LCLS Strategic Facility Development Plan. Technical Report July, SLAC, 2015.

[58] S. V. Milton, E. Gluskin, N. D. Arnold, C. Benson, W. Berg, S. G. Biedron, M. Borland, Y. C. Chae, R. J. Dejus, P. K. Den Hartog, B. Deriy, M. Erdmann, Y. I. Eidelman, M. W. Hahne, Z. Huang, K. J. Kim, J. W. Lewellen, Y. Li, A. H. Lumpkin, O. Makarov, E. R. Moog, A. Nassiri, V. Sajaev, R. Soliday, B. J. Tieman, E. M. Trakhtenberg, G. Travish, I. B. Vasserman, N. A.

Vinokurov, X. J. Wang, G. Wiemerslage, and B. X. Yang. Exponential gain and saturation of a self-amplified spontaneous emission free-electron laser. *Science*, 292(5524):2037–2041, 2001.

[59] I. A. Vartanyants, A. Singer, A. P. Mancuso, O. M. Yefanov, A. Sakdinawat, Y. Liu, E. Bang, G. J. Williams, G. Cadenazzi, B. Abbey, H. Sinn, D. Attwood, K. A. Nugent, E. Weckert, T. Wang, D. Zhu, B. Wu, C. Graves, A. Scherz, J. J. Turner, W. F. Schlotter, M. Messerschmidt, J. Lüning, Y. Acremann, P. Heimann, D. C. Mancini, V. Joshi, J. Krzywinski, R. Soufli, M. Fernandez-Perea, S. Hau-Riege, A. G. Peele, Y. Feng, O. Krupin, S. Moeller, and W. Wurth. Coherence properties of individual femtosecond pulses of an x-ray free-electron laser. *Physical Review Letters*, 107(14):1–5, 2011.

[60] Henry N. Chapman, Carl Caleman, and Nicusor Timneanu. Diffraction before destruction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1647), 2014.

[61] Andrew Aquila, Mark S. Hunter, R. Bruce Doak, Richard A. Kirian, Petra Fromme, Thomas A. White, Jakob Andreasson, David Arnlund, Saša Bajt, Thomas R. M. Barends, Miriam Barthelmess, Michael J. Bogan, Christoph Bostedt, Hervé Bottin, John D. Bozek, Carl Caleman, Nicola Coppola, Jan Davidsson, Daniel P. DePonte, Veit Elser, Sascha W. Epp, Benjamin Erk, Holger Fleckenstein, Lutz Foucar, Matthias Frank, Raimund Fromme, Heinz Graafsma, Ingo Grotjohann, Lars Gumprecht, Janos Hajdu, Christina Y. Hampton, Andreas Hartmann, Robert Hartmann, Stefan Hau-Riege, Günter Hauser, Helmut Hirsemann, Peter Holl, James M. Holton, André Hömke, Linda Johansson, Nils Kimmel, Stephan Kassemeyer, Faton Krasniqi, Kai-Uwe Kühnel, Mengning Liang, Lukas Lomb, Erik Malmerberg, Stefano Marchesini, Andrew V. Martin, Filipe R.N.C. Maia, Marc Messerschmidt, Karol Nass, Christian Reich, Richard Neutze, Daniel Rolles, Benedikt Rudek, Artem Rudenko, Ilme Schlichting, Carlo Schmidt, Kevin E. Schmidt, Joachim Schulz, M. Marvin Seibert, Robert L. Shoeman, Raymond Sierra, Heike Soltau, Dmitri Starodub, Francesco Stellato, Stephan Stern, Lothar Strüder, Nicusor Timneanu, Joachim Ullrich, Xiaoyu Wang, Garth J. Williams, Georg Weidenspointner, Uwe Weierstall, Cornelia Wunderer, Anton Barty, John C. H. Spence, and Henry N. Chapman. Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Optics Express*, 20(3):2706, 2012.

[62] Christopher Milne, Thomas Schietinger, Masamitsu Aiba, Arturo Alarcon, Jürgen Alex, Alexander Anghel, Vladimir Arsov, Carl Beard, Paul Beaud, Simona Bettoni, Markus Bopp, Helge Brands, Manuel Brönnimann, Ingo Brunnenkant, Marco Calvi, Alessandro Citterio, Paolo Craievich, Marta Csatari Divall, Mark Dällenbach, Michael D'Amico, Andreas Dax, Yunpei Deng, Alexander Dietrich, Roberto Dinapoli, Edwin Divall, Sladana Dordevic, Simon Ebner, Christian Erny, Hansrudolf Fitze, Uwe Flechsig, Rolf Follath, Franziska Frei, Florian Gärtner, Romain Ganter, Terence Garvey, Zheqiao Geng, Ishkhan Gorgisyan, Christopher Gough, Andreas Hauff, Christoph

Hauri, Nicole Hiller, Tadej Humar, Stephan Hunziker, Gerhard Ingold, Rasmus Ischebeck, Markus Janousch, Pavle Juranić, Mario Jurcevic, Maik Kaiser, Babak Kalantari, Roger Kalt, Boris Keil, Christoph Kittel, Gregor Knopp, Waldemar Koprek, Henrik Lemke, Thomas Lippuner, Daniel Llorente Sancho, Florian Löhl, Carlos Lopez-Cuenca, Fabian Märki, Fabio Marcellini, Goran Marinkovic, Isabelle Martiel, Ralf Menzel, Aldo Mozzanica, Karol Nass, Gian Orlandi, Cigdem Ozkan Loch, Ezequiel Panepucci, Martin Paraliev, Bruce Patterson, Bill Pedrini, Marco Pedrozzi, Patrick Pollet, Claude Pradervand, Eduard Prat, Peter Radi, Jean-Yves Raguin, Sophie Redford, Jens Rehanek, Julien Réhault, Sven Reiche, Matthias Ringele, Jochen Rittmann, Leonid Rivkin, Albert Romann, Marie Ruat, Christian Ruder, Leonardo Sala, Lionel Schebacher, Thomas Schilcher, Volker Schlott, Thomas Schmidt, Bernd Schmitt, Xintian Shi, Markus Stadler, Lukas Stingelin, Werner Sturzenegger, Jakub Szlachetko, Dhanya Thattil, Daniel Treyer, Alexandre Trisorio, Wolfgang Tron, Seraphin Vetter, Carlo Vicario, Didier Voulot, Meitian Wang, Thierry Zamofing, Christof Zellweger, Riccardo Zennaro, Elke Zimoch, Rafael Abela, Luc Patthey, and Hans-Heinrich Braun. SwissFEL: The Swiss X-ray Free Electron Laser. *Applied Sciences*, 7(7):720, jul 2017.

[63] Makina Yabashi, Hitoshi Tanaka, and Tetsuya Ishikawa. Overview of the SACLA facility. *Journal of Synchrotron Radiation*, 22(3):477–484, may 2015.

[64] Heung Sik Kang, Chang Ki Min, Hoon Heo, Changbum Kim, Haeryong Yang, Gyujin Kim, Inhyuk Nam, Soung Youl Baek, Hyo Jin Choi, Geonyeong Mun, Byoung Ryul Park, Young Jin Suh, Dong Cheol Shin, Jinyul Hu, Juho Hong, Seonghoon Jung, Sangsoo Hee Seonghan Seung Hwan Nam Sunam Kim, Kwang Hoon Woo Kyung Sook Kim, Donghyun Na, Soung Soo Sung Ju Sang Youn Han Park, Yong Jung Park, Jang Hui Han, Young Gyu Jung, Seong Hun Jeong, Heung Soo Hee Seock Hong Gi Lee, Sangbong Sojeong Lee, Sangbong Sojeong Lee, Woul Woo Lee, Bonggi Oh, Hyung Suck Suh, Yong Woon Parc, Soung Soo Sung Ju Sang Youn Han Park, Min Ho Myong Jin Kim, Nam Suk Jung, Young Chan Yongsam Kim, Mong Soo Lee, Bong Ho Lee, Chi Won Sung, Ik Su Mok, Jung Moo Yang, Chae Soon Lee, Hocheol Shin, Jangwoo Ji Hwa Kim, Young Chan Yongsam Kim, Jae Hyuk Lee, Soung Soo Sung Ju Sang Youn Han Park, Jangwoo Ji Hwa Kim, Jaehun Jaeku Jaehyun Park, Intae Eom, Seungyu Rah, Sangsoo Hee Seonghan Seung Hwan Nam Sunam Kim, Ki Hyun Nam, Jaehun Jaeku Jaehyun Park, Jaehun Jaeku Jaehyun Park, Sangsoo Hee Seonghan Seung Hwan Nam Sunam Kim, Soonam Kwon, Soung Soo Sung Ju Sang Youn Han Park, Kwang Hoon Woo Kyung Sook Kim, Hyojung Hyun, Sangsoo Hee Seonghan Seung Hwan Nam Sunam Kim, Sangsoo Hee Seonghan Seung Hwan Nam Sunam Kim, Sun Min Hwang, Min Ho Myong Jin Kim, Chae Yong Lim, Chung Jong Yu, Bong Soo Kim, Tai Hee Kang, Kwang Hoon Woo Kyung Sook Kim, Sangsoo Hee Seonghan Seung Hwan Nam Sunam Kim, Heung Soo Hee Seock Hong Gi Lee, Heung Soo Hee Seock Hong Gi Lee,

Ki Hyeon Park, Tae Yeong Koo, Dong Eon Kim, and In Soo Ko. Hard X-ray free-electron laser with femtosecond-scale timing jitter. *Nature Photonics*, 11(11):708–713, 2017.

[65] H. Weise. Status of the European XFEL. In *Proceedings of the 28th Linear Accelerator Conference, LINAC 2016*, pages 7–11, 2016.

[66] H. M. Berman. The Protein Data Bank. *Nucleic Acids Research*, 28(1):235–242, jan 2000.

[67] Andrea Mozzarelli and Gian Luigi Rossi. Protein function in the crystal. *Annual Review of Biophysics and Biomolecular Structure*, 25:343–365, 1996.

[68] Nathaniel Echols, H. van den Bedem, James S. Fraser, P. Therese Lang, James M. Holton, Tom Alber, Avi J. Samelson, H. van den Bedem, Avi J. Samelson, P. Therese Lang, James M. Holton, Nathaniel Echols, and Tom Alber. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proceedings of the National Academy of Sciences*, 108(39):16247–16252, sep 2011.

[69] A. Rigos, T. Sage, D. Morikis, and L Champion, P.M.and Zhu. Conformational interconversion in protein crystals. *Journal of Molecular Biology*, 224(1):207–215, 2004.

[70] Keith Moffat, Doletha Szebenyi, and Donald Bilderback. X-ray Laue diffraction from protein crystals. *Science*, 223(4643):1423–1425, 1984.

[71] Friedrich Schotte, Jayashree Soman, John S. Olson, Michael Wulff, and Philip A. Anfinrud. Picosecond time-resolved X-ray crystallography: Probing protein function in real time. *Journal of Structural Biology*, 147(3):235–246, 2004.

[72] Annemarie B. Wöhri, Gergely Katona, Linda C. Johansson, Emelie Fritz, Erik Malmerberg, Magnus Andersson, Jonathan Vincent, Mattias Eklund, Marco Cammarata, Michael Wulff, Jan Davidsson, Gerrit Groenhof, and Richard Neutze. Light-induced structural changes in a photosynthetic reaction center caught by laue diffraction. *Science*, 328(5978):630–633, 2010.

[73] Friedrich Schotte, Manho Lim, Timothy A. Jackson, Aleksandr V. Smirnov, Jayashree Soman, John S. Olson, George N. Phillips, Michael Wulff, and Philip A. Anfinrud. Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science*, 300(5627):1944–1947, 2003.

[74] Keith Moffat. Time-resolved biochemical crystallography: A mechanistic perspective. *Chemical Reviews*, 101(6):1569–1581, 2001.

[75] Kyung Hwan Kim, Srinivasan Muniyappan, Key Young Oang, Jong Goo Kim, Shunsuke Nozawa, Tokushi Sato, Shin Ya Koshihara, Robert Henning, Irina Kosheleva, Hosung Ki, Youngmin Kim, Tae Wu Kim, Jeongho Kim, Shin Ichi Adachi, and Hyotcherl Ihee. Direct observation of cooperative

protein structural dynamics of homodimeric hemoglobin from 100 ps to 10 ms with pump-probe X-ray solution scattering. *Journal of the American Chemical Society*, 134(16):7001–7008, 2012.

[76] Hyun Sun Choa, Naranbaatar Dashdorj, Friedrich Schotte, Timothy Graber, Robert Henning, and Philip Anfinrud. Protein structural dynamics in solution unveiled via 100-ps time-resolved x-ray scattering. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16):7281–7286, 2010.

[77] Melissa A. Graewert and Dmitri I. Svergun. Impact and progress in small and wide angle X-ray scattering (SAXS and WAXS). *Current Opinion in Structural Biology*, 23(5):748–754, 2013.

[78] Derren J. Heyes, Samantha J.O. Hardman, Martin N. Pedersen, Joyce Woodhouse, Eugenio De La Mora, Michael Wulff, Martin Weik, Marco Cammarata, Nigel S. Scrutton, and Giorgio Schirò. Light-induced structural changes in a full-length cyanobacterial phytochrome probed by time-resolved X-ray scattering. *Communications Biology*, 2(1):1–8, 2019.

[79] Francesco Stellato, Dominik Oberthür, Mengning Liang, Richard Bean, Cornelius Gati, Oleksandr Yefanov, Anton Barty, Anja Burkhardt, Pontus Fischer, Lorenzo Galli, Richard A Kirian, Jan Meyer, Saravanan Panneerselvam, Chun Hong Yoon, Fedor Chervinskii, Emily Speller, Thomas A White, Christian Betzel, Alke Meents, and Henry N Chapman. Room-temperature macromolecular serial crystallography using synchrotron radiation. *IUCrJ*, 1(4):204–212, jul 2014.

[80] S. C. Harrison, A. J. Olson, C. E. Schutt, F. K. Winkler, and G. Bricogne. Tomato bushy stunt virus at 2.9 resolution. *Nature*, 276(5686):368–373, 1978.

[81] P. Engström, S. Fiedler, and C. Riekel. Microdiffraction instrumentation and experiments on the microfocus beamline at the ESRF. *Review of Scientific Instruments*, 66(2):1348–1350, 1995.

[82] Stephen Cusack, Hassan Belrhali, Andreas Bram, Manfred Burghammer, Anastassis Perrakis, and Christian Riekel. Small is beautiful: protein micro-crystallography. *Nature Structural Biology*, 5(8):634–637, aug 1998.

[83] Janet L. Smith, Robert F. Fischetti, and Masaki Yamamoto. Micro-crystallography comes of age. *Current Opinion in Structural Biology*, 22(5):602–612, 2012.

[84] James M. Holton and Kenneth A. Frankel. The minimum crystal size needed for a complete diffraction data set. *Acta Crystallographica Section D: Biological Crystallography*, 66(4):393–408, 2010.

[85] Kay Diederichs and Meitian Wang. Serial Synchrotron X-Ray Crystallography (SSX). In *Protein Crystallography*, pages 239–272. Springer Protocalls, 1st edition, 2017.

[86] Helena Käck, Katharine J. Gibson, Ylva Lindqvist, and Gunter Schneider. Snapshot of a phosphorylated substrate intermediate by kinetic crystallography. *Proceedings of the National Academy of Sciences of the United States of America*, 95(10):5495–5500, 1998.

[87] George D Calvey, Andrea M Katz, Chris B Schaffer, and Lois Pollack. Mixing injector enables time-resolved crystallography with high hit rate at X-ray free electron lasers. *Structural Dynamics*, 3(5), 2016.

[88] Jose Olmos, Suraj Pandey, Jose M. Martin-Garcia, George Calvey, Andrea Katz, Juraj Knoska, Christopher Kupitz, Mark S. Hunter, Mengning Liang, Dominik Oberthuer, Oleksandr Yefanov, Max Wiedorn, Michael Heyman, Mark Holl, Kanupriya Pande, Anton Barty, Mitchell D. Miller, Stephan Stern, Shatabdi Roy-Chowdhury, Jesse Coe, Nirupa Nagaratnam, James Zook, Jacob Verburgt, Tyler Norwood, Ishwor Poudyal, David Xu, Jason Koglin, Matt Seaberg, Yun Zhao, Saša Bajt, Thomas Grant, Valerio Mariani, Garrett Nelson, Ganesh Subramanian, Euiyoung Bae, Raimund Fromme, Russel Fung, Peter Schwander, Matthias Frank, Thomas White, Uwe Weierstall, Nadia Zatsepin, John Spence, Petra Fromme, Henry N. Chapman, Lois Pollack, Lee Tremblay, Abbas Ourmazd, George N. Phillips, and Marius Schmidt. Enzyme Intermediates Captured "on-the-fly" by Mix-and-Inject Serial Crystallography. *bioRxiv*, pages 1–15, 2017.

[89] U. K. Genick. Structure of a Protein Photocycle Intermediate by Millisecond Time-Resolved Crystallography. *Science*, 275(5305):1471–1475, mar 1997.

[90] Marius Schmidt. Mix and inject: Reaction initiation by diffusion for time-resolved macromolecular crystallography. *Advances in Condensed Matter Physics*, 2013, 2013.

[91] Marius Schmidt. Reaction initiation in enzyme crystals by diffusion of substrate. *Crystals*, 10(2):1–15, 2020.

[92] Keisuke Sakai, Yasuhiro Matsui, Tsutomu Kouyama, Yoshitsugu Shiro, and Shin Ichi Adachi. Optical monitoring of freeze-trapped reaction intermediates in protein crystals: A microspectrophotometer for cryogenic protein crystallography. *Journal of Applied Crystallography*, 35(2):270–273, 2002.

[93] Bertil Halle. Biomolecular cryocrystallography: Structural changes during flash-cooling. *Proceedings of the National Academy of Sciences of the United States of America*, 101(14):4793–4798, 2004.

[94] Douglas H. Juers and Brian W. Matthews. Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions. *Journal of Molecular Biology*, 311(4):851–862, 2001.

[95] D. J. Haas and M. G. Rossmann. Crystallographic studies on lactate dehydrogenase at –75°C. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 26(7):998–1004, jul 1970.

[96] Elspeth F. Garman and Thomas R. Schneider. Macromolecular Cryocrystallography. *Journal of Applied Crystallography*, 30(3):211–237, 1997.

[97] Philip Evans and Airlie McCoy. An introduction to molecular replacement. *Acta Crystallographica Section D Biological Crystallography*, 64(1):1–10, jan 2008.

[98] Thomas A. White. Post-refinement method for snapshot serial crystallography. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1647):0–5, 2014.

[99] Loes M.J. Kroon-Batenburg, Antoine M.M. Schreurs, Raimond B.G. Ravelli, and Piet Gros. Accounting for partiality in serial crystallography using ray-tracing principles. *Acta Crystallographica Section D: Biological Crystallography*, 71:1799–1811, 2015.

[100] Colin Nave. A description of imperfections in protein crystals. *Acta Crystallographica Section D: Biological Crystallography*, 54(5):848–853, 1998.

[101] Richard A Kirian, Xiaoyu Wang, Uwe Weierstall, Kevin E Schmidt, John C. H. Spence, Mark Hunter, Petra Fromme, Thomas White, Henry N Chapman, and James Holton. Femtosecond protein nanocrystallography—data analysis methods. *Optics Express*, 18(6):5713, mar 2010.

[102] Thomas A. White, Richard A. Kirian, Andrew V. Martin, Andrew Aquila, Karol Nass, Anton Barty, and Henry N. Chapman. CrystFEL: A software suite for snapshot serial crystallography. *Journal of Applied Crystallography*, 45(2):335–341, 2012.

[103] Elspeth F. Garman and Martin Weik. X-ray radiation damage to biological macromolecules: further insights. *Journal of Synchrotron Radiation*, 24(1):1–6, jan 2017.

[104] Anton Barty, Carl Caleman, Andrew Aquila, Nicusor Timneanu, Lukas Lomb, Thomas A. White, Jakob Andreasson, David Arnlund, Saša Bajt, Thomas R. M. Barends, Miriam Barthelmess, Michael J. Bogan, Christoph Bostedt, John D. Bozek, Ryan Coffee, Nicola Coppola, Jan Davidsson, Daniel P. DePonte, R. Bruce Doak, Tomas Ekeberg, Veit Elser, Sascha W. Epp, Benjamin Erk, Holger Fleckenstein, Lutz Foucar, Petra Fromme, Heinz Graafsma, Lars Gumprecht, Janos Hajdu, Christina Y. Hampton, Robert Hartmann, Andreas Hartmann, Günter Hauser, Helmut Hirsemann, Peter Holl, Mark S. Hunter, Linda Johansson, Stephan Kassemeyer, Nils Kimmel, Richard A. Kirian, Mengning Liang, Filipe R. N. C. Maia, Erik Malmerberg, Stefano Marchesini, Andrew V. Martin, Karol Nass, Richard Neutze, Christian Reich, Daniel Rolles, Benedikt Rudek, Artem Rudenko, Howard Scott, Ilme Schlichting, Joachim Schulz, M. Marvin Seibert, Robert L.

Shoeman, Raymond G. Sierra, Heike Soltau, John C. H. Spence, Francesco Stellato, Stephan Stern, Lothar Strüder, Joachim Ullrich, X. Wang, Georg Weidenspointner, Uwe Weierstall, Cornelia B. Wunderer, and Henry N. Chapman. Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nature Photonics*, 6(1):35–40, jan 2012.

[105] Muhamed Amin, Ashraf Badawi, and S. S. Obayya. Radiation Damage in XFEL: Case study from the oxygen-evolving complex of Photosystem II. *Scientific Reports*, 6(November):6–11, 2016.

[106] Muhamed Amin, Mikhail Askerka, Victor S. Batista, Gary W. Brudvig, and Marilyn R. Gunner. XFEL Radiation Damage Through the S-State Cycle of the Oxygen-Evolving Complex of Photosystem II. *The Journal of Physical Chemistry B*, page acs.jpcb.7b08371, 2017.

[107] Karol Nass, Lutz Foucar, Thomas R.M. Barends, Elisabeth Hartmann, Sabine Botha, Robert L. Shoeman, R. Bruce Doak, Roberto Alonso-Mori, Andrew Aquila, Saša Bajt, Anton Barty, Richard Bean, Kenneth R. Beyerlein, Maike Bublitz, Nikolaj Drachmann, Jonas Gregersen, H. Olof Jönsson, Wolfgang Kabsch, Stephan Kassemeyer, Jason E. Koglin, Michael Krumrey, Daniel Mattle, Marc Messerschmidt, Poul Nissen, Linda Reinhard, Oleg Sitsel, Dimosthenis Sokaras, Garth J. Williams, Stefan Hau-Riege, Nicusor Timneanu, Carl Caleman, Henry N. Chapman, Sébastien Boutet, and Ilme Schlichting. Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *Journal of Synchrotron Radiation*, 22(2):225–238, 2015.

[108] Karol Nass. Radiation damage in protein crystallography at X-ray free-electron lasers. *Acta Crystallographica Section D Structural Biology*, 75(2):211–218, feb 2019.

[109] J. L. Campbell and Tibor Papp. Widths of the atomic K-N7 levels. *Atomic Data and Nuclear Data Tables*, 77(1):1–56, 2001.

[110] C.G. Darwin. XXXIV. The theory of X-ray reflexion. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 27(158):315–333, feb 1914.

[111] Richard Henderson. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics*, 28(02):171, 1995.

[112] Elisabeth P. Carpenter, Konstantinos Beis, Alexander D. Cameron, and So Iwata. Overcoming the challenges of membrane protein crystallography, oct 2008.

[113] David Phillips. The Biography of George Porter. In *The Life and Scientific Legacy of George Porter*, number 1947, pages 1–304. Imperial College Press, jul 2006.

[114] Alvaro Sanchez-Gonzalez, Allan S. Johnson, Ann Fitzpatrick, Christopher D.M. M. M Hutchison, Clyde Fare, Violeta Cordon-Preciado, Gabriel Dorlhiac, Josie L. Ferreira, Rhodri M. Morgan,

Jon P. Marangos, Shigeki Owada, Takanori Nakane, Rie Tanaka, Kensuke Tono, So Iwata, and Jasper J. van Thor. Coincidence timing of femtosecond optical pulses in an X-ray free electron laser. *Journal of Applied Physics*, 122(20):203105, nov 2017.

[115] Minseok Kim, Chang Ki Min, and Intae Eom. Laser systems for time-resolved experiments at the pohang accelerator laboratory x-ray free-electron laser beamlines. *Journal of Synchrotron Radiation*, 26(3):868–873, 2019.

[116] Franklin D Fuller, Sheraz Gul, Ruchira Chatterjee, E Sethe Burgie, Iris D Young, Hugo Lebrette, Vivek Srinivas, Aaron S Brewster, Tara Michels-Clark, Jonathan A Clinger, Babak Andi, Mohamed Ibrahim, Ernest Pastor, Casper de Lichtenberg, Rana Hussein, Christopher J Pollock, Miao Zhang, Claudiu A Stan, Thomas Kroll, Thomas Fransson, Clemens Weninger, Markus Kubin, Pierre Aller, Louise Lassalle, Philipp Bräuer, Mitchell D Miller, Muhamed Amin, Sergey Koroidov, Christian G Roessler, Marc Allaire, Raymond G Sierra, Peter T Docker, James M Glownia, Silke Nelson, Jason E Koglin, Diling Zhu, Matthieu Chollet, Sanghoon Song, Henrik Lemke, Mengning Liang, Dimosthenis Sokaras, Roberto Alonso-Mori, Athina Zouni, Johannes Messinger, Uwe Bergmann, Amie K Boal, J Martin Bollinger, Carsten Krebs, Martin Högbom, George N Phillips, Richard D Vierstra, Nicholas K Sauter, Allen M Orville, Jan Kern, Vittal K Yachandra, and Junko Yano. Drop-on-demand sample delivery for studying biocatalysts in action at X-ray free-electron lasers. *Nature Methods*, 14(4):443–449, 2017.

[117] Philip Hart, Sébastien Boutet, Gabriella Carini, Mikhail Dubrovin, Brian Duda, David Fritz, Gunther Haller, Ryan Herbst, Sven Herrmann, Chris Kenney, Nadine Kurita, Henrik Lemke, Marc Messerschmidt, Martin Nordby, Jack Pines, Don Schafer, Matt Swift, Matt Weaver, Garth Williams, Diling Zhu, Niels Van Bakel, and John Morse. The CSPAD megapixel x-ray camera at LCLS. *X-Ray Free-Electron Lasers: Beam Diagnostics, Beamline Instrumentation, and Applications*, 8504(October 2012):85040C, 2012.

[118] Mengning Liang, Garth J. Williams, Marc Messerschmidt, M. Marvin Seibert, Paul A. Montanez, Matt Hayes, Despina Milathianaki, Andrew Aquila, Mark S. Hunter, Jason E. Koglin, Donald W. Schafer, Serge Guillet, Armin Busse, Robert Bergan, William Olson, Kay Fox, Nathaniel Stewart, Robin Curtis, Alireza Alan Miahnahri, and Sébastien Boutet. The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *Journal of Synchrotron Radiation*, 22(3):514–519, may 2015.

[119] Uwe Weierstall. Liquid sample delivery techniques for serial femtosecond crystallography. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1647):2–7, 2014.

[120] J. Warner, J C H Spence, U. Weierstall, R. B. Doak, D. Starodub, D. P. DePonte, K. Schmidt,

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

[121] Dominik Oberthuer, Juraj Knoška, Max O. Wiedorn, Kenneth R. Beyerlein, David A. Bushnell, Elena G. Kovaleva, Michael Heymann, Lars Gumprecht, Richard A. Kirian, Anton Barty, Valerio Mariani, Aleksandra Tolstikova, Luigi Adriano, Salah Awel, Miriam Barthelmess, Katerina Dörner, P. Lourdu Xavier, Oleksandr Yefanov, Daniel R. James, Garrett Nelson, Dingjie Wang, George Calvey, Yujie Chen, Andrea Schmidt, Michael Szczepek, Stefan Frielingsdorf, Oliver Lenz, Edward Snell, Philip J. Robinson, Božidar Šarler, Grega Belšak, Marjan Maček, Fabian Wilde, Andrew Aquila, Sébastien Boutet, Mengning Liang, Mark S. Hunter, Patrick Scheerer, John D. Lipscomb, Uwe Weierstall, Roger D. Kornberg, John C. H. Spence, Lois Pollack, Henry N. Chapman, and Saša Bajt. Double-flow focused liquid injector for efficient serial femtosecond crystallography. *Scientific Reports*, 7(January):44628, 2017.

[122] Allen M Orville. CHAPTER 18. Acoustic Methods for On-demand Sample Injection into XFEL Beams. In *Acoustic Methods for On-demand Sample Injection into XFEL Beams*, number 18, pages 348–364. The Royal Society of Chemistry, 2017.

[123] Christian G. Roessler, Rakhi Agarwal, Marc Allaire, Roberto Alonso-Mori, Babak Andi, José F.R. Bachega, Martin Bommer, Aaron S. Brewster, Michael C. Browne, Ruchira Chatterjee, Eunsun Cho, Aina E. Cohen, Matthew Cowan, Sammy Datwani, Victor L. Davidson, Jim Defever, Brent Eaton, Richard Ellson, Yiping Feng, Lucien P. Ghislain, James M. Glownia, Guangye Han, Johan Hattne, Julia Hellmich, Annie Héroux, Mohamed Ibrahim, Jan Kern, Anthony Kuczewski, Henrik T. Lemke, Pinghua Liu, Lars Majlof, William M. McClintock, Stuart Myers, Silke Nelsen, Joe Olechno, Allen M. Orville, Nicholas K. Sauter, Alexei S. Soares, S. Michael Soltis, Heng Song, Richard G. Stearns, Rosalie Tran, Yingssu Tsai, Monarin Uervirojnangkoorn, Carrie M. Wilmot, Vittal Yachandra, Junko Yano, Erik T. Yukl, Diling Zhu, and Athina Zouni. Acoustic Injectors for Drop-On-Demand Serial Femtosecond Crystallography. *Structure*, 24(4):631–640, 2016.

[124] Darren A. Sherrell, Andrew J. Foster, Lee Hudson, Brian Nutter, James O'Hea, Silke Nelson, Olivier Paré-Labrosse, Saeed Oghbaey, R. J Dwayne Miller, and Robin L. Owen. 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*, 22:1372–1378, 2015.

[125] C. Mueller, A. Marx, S. W. Epp, Y. Zhong, A. Kuo, A. R. Balo, J. Soman, F. Schotte, H. T. Lemke, R. L. Owen, E. F. Pai, A. R. Pearson, J. S. Olson, P. A. Anfinrud, O. P. Ernst, and R. J. Dwayne Miller. Fixed target matrix for femtosecond time-resolved and in situ serial micro-crystallography. *Structural Dynamics*, 2(5):054302, sep 2015.

[126] Silke Nelson, Jingshan Ren, Helen M Ginn, Philip Roedig, Karl Harlos, Bernd Reime, Ramona Duman, Martin Warmer, Marcin Sikorsky, Christian David, Daniel S Damiani, David J Rowlands, Roberto Alonso-Mori, Nicholas K Sauter, Aaron S Brewster, Tim Pakendorf, Armin Wagner, Iris D Young, Alke Meents, Geoff Sutton, Elizabeth E Fry, James Kelly, Tara Michels-Clark, Jan Meyer, Thomas S Walter, Ismo Vartiainen, David I Stuart, Pontus Fischer, and Abhay Kotecha. High-speed fixed-target serial virus crystallography. *Nature Methods*, 14(8):805–810, jun 2017.

[127] Robin L. Owen, Danny Axford, Darren A. Sherrell, Anling Kuo, Oliver P. Ernst, Eike C. Schulz, R. J.Dwayne Miller, and Henrike M. Mueller-Werkmeister. Low-dose fixed-target serial synchrotron crystallography. *Acta Crystallographica Section D: Structural Biology*, 73(4):373–378, 2017.

[128] Saeed Oghbaey, Antoine Sarracini, Helen M Ginn, Olivier Pare-Labrosse, Anling Kuo, Alexander Marx, Sascha W. Epp, Darren A. Sherrell, Bryan T. Eger, Yinpeng Zhong, Rolf Loch, Valerio Mariani, Roberto Alonso-Mori, Silke Nelson, Henrik T. Lemke, Robin L. Owen, Arwen R. Pearson, David I. Stuart, Oliver P. Ernst, Henrike M. Mueller-Werkmeister, and R. J. Dwayne Miller. Fixed target combined with spectral mapping: approaching 100% hit rates for serial crystallography. *Acta Crystallographica Section D Structural Biology*, 72(8):944–955, aug 2016.

[129] Eike C Schulz, Pedram Mehrabi, Henrike M Müller-Werkmeister, Friedjof Tellkamp, Ajay Jha, William Stuart, Elke Persch, Raoul De Gasparo, François Diederich, Emil F Pai, and R J Dwayne Miller. The hit-and-return system enables efficient time-resolved serial synchrotron crystallography. *Nature Methods*, 15(11):901–904, nov 2018.

[130] Pedram Mehrabi, Henrike M. Muller-Werkmeister, Jan Philipp Leimkohl, Hendrik Schikora, Jelena Ninkovic, Silvia Krivokuca, Ladislav Andricek, Sascha W. Epp, Darren Sherrell, Robin L. Owen, Arwen R. Pearson, Friedjof Tellkamp, Eike C. Schulz, and R. J. Dwayne Miller. The HARE chip for efficient time-resolved serial synchrotron crystallography. *Journal of Synchrotron Radiation*, 27:360–370, 2020.

[131] J. Schulz, J. Bielecki, R. B. Doak, K. Dörner, R. Graceffa, R. L. Shoeman, M. Sikorski, P. Thute, D. Westphal, and A. P. Mancuso. A versatile liquid-jet setup for the European XFEL. *Journal of Synchrotron Radiation*, 26(2):339–345, 2019.

[132] Marie Luise Grünbein and Gabriela Nass Kovacs. Sample delivery for serial crystallography at free-electron lasers and synchrotrons. *Acta Crystallographica Section D: Structural Biology*, 75:178–191, 2019.

[133] Chitra Rajendran, Florian S. N. Dworkowski, Meitian Wang, and Clemens Schulze-Briese. Radiation damage in room-temperature data acquisition with the PILATUS 6M pixel detector. *Journal of Synchrotron Radiation*, 18(3):318–328, may 2011.

[134] Briony A. Yorke, Godfrey S. Beddard, Robin L. Owen, and Arwen R. Pearson. Time-resolved crystallography using the Hadamard transform. *Nature Methods*, 11(11):1131–1134, 2014.

[135] Tobias Weinert, Petr Skopintsev, Daniel James, Florian Dworkowski, Ezequiel Panepucci, Demet Kekilli, Antonia Furrer, Steffen Brünle, Sandra Mous, Dmitry Ozerov, Przemyslaw Nogly, Meitian Wang, and Jörg Standfuss. Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography. *Science*, 364(6448):61–65, 2019.

[136] Kunio Hirata, Kyoko Shinzawa-Itoh, Naomine Yano, Shuhei Takemura, Koji Kato, Miki Hatanaka, Kazumasa Muramoto, Takako Kawahara, Tomitake Tsukihara, Eiki Yamashita, Kensuke Tono, Go Ueno, Takaaki Hikima, Hironori Murakami, Yuichi Inubushi, Makina Yabashi, Tetsuya Ishikawa, Masaki Yamamoto, Takashi Ogura, Hiroshi Sugimoto, Jian-Ren Shen, Shinya Yoshikawa, and Hideo Ago. Determination of damage-free crystal structure of an X-ray–sensitive protein using an XFEL. *Nature Methods*, 11(7):734–736, jul 2014.

[137] Catherine Dejoie, Lynne B. Mccusker, Christian Baerlocher, Rafael Abela, Bruce D. Patterson, Martin Kunz, and Nobumichi Tamura. Using a non-monochromatic microbeam for serial snapshot crystallography. *Journal of Applied Crystallography*, 46(3):791–794, 2013.

[138] Tobias Weinert, Natacha Olieric, Robert Cheng, Steffen Brünle, Daniel James, Dmitry Ozerov, Dardan Gashi, Laura Vera, May Marsh, Kathrin Jaeger, Florian Dworkowski, Ezequiel Panepucci, Shibom Basu, Petr Skopintsev, Andrew S. Doré, Tian Geng, Robert M. Cooke, Mengning Liang, Andrea E. Prota, Valerie Panneels, Przemyslaw Nogly, Ulrich Ermler, Gebhard Schertler, Michael Hennig, Michel O. Steinmetz, Meitian Wang, and Jörg Standfuss. Serial millisecond crystallography for routine room-temperature structure determination at synchrotrons. *Nature Communications*, 8(1):542, dec 2017.

[139] Jose M. Martin-Garcia, Chelsie E. Conrad, Jesse Coe, Shatabdi Roy-Chowdhury, and Petra Fromme. Serial femtosecond crystallography: A revolution in structural biology. *Archives of Biochemistry and Biophysics*, 602:32–47, jul 2016.

[140] A. Tolstikova, M. Levantino, O. Yefanov, V. Hennicke, P. Fischer, J. Meyer, A. Mozzanica, S. Redford, E. Crosas, N. L. Opara, M. Barthelmess, J. Lieske, D. Oberthuer, E. Wator, I. Mohacsi, M. Wulff, B. Schmitt, H. N. Chapmana, and A. Meents. 1 kHz fixed-target serial crystallography using a multilayer monochromator and an integrating pixel detector. *IUCrJ*, 6:927–937, 2019.

[141] Elspeth F Garman. Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta Crystallographica Section D Biological Crystallography*, 66(4):339–351, apr 2010.

[142] Robin L. Owen, Danny Axford, Joanne E. Nettleship, Raymond J. Owens, James I. Robinson, Ann W. Morgan, Andrew S. Doré, Guillaume Lebon, Christopher G. Tate, Elizabeth E. Fry, Jingshan Ren, David I. Stuart, and Gwyndaf Evans. Outrunning free radicals in room-temperature macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 68(7):810–818, 2012.

[143] T. M. Kuzay, M. Kazmierczak, and B. J. Hsieh. X-ray beam/biomaterial thermal interactions in third-generation synchrotron sources. *Acta Crystallographica Section D: Biological Crystallography*, 57(1):69–81, 2001.

[144] Eugenio de la Mora, Nicolas Coquelle, Charles S Bury, Martin Rosenthal, James M Holton, Ian Carmichael, Elspeth F. Garman, Manfred Burghammer, Jacques-Philippe Colletier, and Martin Weik. Radiation damage and dose limits in serial synchrotron crystallography at cryo- and room temperatures. *Proceedings of the National Academy of Sciences*, 117(8):201821522, feb 2020.

[145] Anton Barty, Richard A. Kirian, Filipe R. N. C. Maia, Max Hantke, Chun Hong Yoon, Thomas A. White, and Henry Chapman. Cheetah : software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *Journal of Applied Crystallography*, 47(3):1118–1131, jun 2014.

[146] Valerio Mariani, Andrew Morgan, Chun Hong Yoon, Thomas J. Lane, Thomas A. White, Christopher O'grady, Manuela Kuhn, Steve Aplin, Jason Koglin, Anton Barty, and Henry N. Chapman. OnDA: Online data analysis and feedback for serial X-ray imaging. *Journal of Applied Crystallography*, 49:1073–1080, 2016.

[147] Hocheol Shin, Seungnam Kim, and Chun Hong Yoon. Data Analysis using Psocake at PAL-XFEL. *Journal of the Korean Physical Society*, 73(1):16–20, 2018.

[148] Lutz Foucar. CFEL-ASG Software Suite (CASS): Usage for free-electron laser experiments with biological focus. *Journal of Applied Crystallography*, 49(4):1336–1346, 2016.

[149] Andrew G. W. Leslie and Harold R. Powell. Processing diffraction data with mosflm. In *Evolving Methods for Macromolecular Crystallography*, pages 41–51. Springer US, 2007.

[150] Harold R. Powell, Owen Johnson, and Andrew G.W. Leslie. Autoindexing diffraction images with iMosflm. *Acta Crystallographica Section D: Biological Crystallography*, 69(7):1195–1203, 2013.

[151] Albert J.M. M Duisenberg. Indexing in single-crystal diffractometry with an obstinate list of reflections. *Journal of Applied Crystallography*, 25(pt 2):92–96, 1992.

[152] Yaroslav Gevorkov, Oleksandr Yefanov, Anton Barty, Thomas A. White, Valerio Mariani, Wolfgang Brehm, Aleksandra Tolstikova, Rolf-Rainer Grigat, and Henry N. Chapman. XGANDALF

– extended gradient descent algorithm for lattice finding. *Acta Crystallographica Section A Foundations and Advances*, 75(5):694–704, sep 2019.

[153] Wolfgang Kabsch. XDS. *Acta Crystallographica Section D Biological Crystallography*, 66(2):125–132, 2010.

[154] Kenneth R. Beyerlein, Thomas A. White, Oleksandr Yefanov, Cornelius Gati, Ivan G. Kazantsev, Nicolai Fog Gade Nielsen, Peter M. Larsen, Henry N. Chapman, and Soren Schmidt. FELIX: An algorithm for indexing multiple crystallites in X-ray free-electron laser snapshot diffraction images. *Journal of Applied Crystallography*, 50:1075–1083, 2017.

[155] Helen Mary Ginn, Philip Roedig, Anling Kuo, Gwyndaf Evans, Nicholas K. Sauter, Oliver Ernst, Alke Meents, Henrike Mueller-Werkmeister, R. J.Dwayne Miller, and David Ian Stuart. TakeTwo: An indexing algorithm suited to still images with known crystal parameters. *Acta Crystallographica Section D: Structural Biology*, 72(8):956–965, 2016.

[156] Yaroslav Gevorkov, Anton Barty, Wolfgang Brehm, Thomas A. White, Aleksandra Tolstikova, Max O. Wiedorn, Alke Meents, Rolf-Rainer Grigat, Henry N. Chapman, and Oleksandr Yefanov. pinkIndexer – a universal indexer for pink-beam X-ray and electron diffraction snapshots. *Acta Crystallographica Section A Foundations and Advances*, 76(2):121–131, mar 2020.

[157] Nicholas K Sauter, Johan Hattne, Ralf W Grosse-Kunstleve, and Nathaniel Echols. New Python-based methods for data processing. *Acta Crystallographica Section D Biological Crystallography*, 69(7):1274–1282, 2013.

[158] Ralf W. Grosse-Kunstleve, Nicholas K. Sauter, Nigel W. Moriarty, and Paul D. Adams. The Computational Crystallography Toolbox: Crystallographic algorithms in a reusable software framework. *Journal of Applied Crystallography*, 35(1):126–136, 2002.

[159] Zepu Zhang, Nicholas K. Sauter, Henry Van Den Bedem, Gyorgy Snell, and Ashley M. Deacon. Automated diffraction image analysis and spot searching for high-throughput crystal screening. *Journal of Applied Crystallography*, 39(1):112–119, 2006.

[160] Johan Hattne, Nathaniel Echols, Rosalie Tran, Jan Kern, Richard J. Gildea, Aaron S. Brewster, Roberto Alonso-Mori, Carina Glöckner, Julia Hellmich, Hartawan Laksmono, Raymond G. Sierra, Benedikt Lassalle-Kaiser, Alyssa Lampe, Guangye Han, Sheraz Gul, Dörte Difiore, Despina Milathianaki, Alan R. Fry, Alan Miahnahri, William E. White, Donald W. Schafer, M. Marvin Seibert, Jason E. Koglin, Dimosthenis Sokaras, Tsu Chien Weng, Jonas Sellberg, Matthew J. Latimer, Pieter Glatzel, Petrus H. Zwart, Ralf W. Grosse-Kunstleve, Michael J. Bogan, Marc Messerschmidt, Garth J. Williams, Sébastien Boutet, Johannes Messinger, Athina Zouni, Junko

Yano, Uwe Bergmann, Vittal K. Yachandra, Paul D. Adams, and Nicholas K. Sauter. Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nature Methods*, 11(5):545–548, 2014.

[161] Oleksandr Yefanov, Valerio Mariani, Cornelius Gati, Thomas A. White, Henry N. Chapman, and Anton Barty. Accurate determination of segmented X-ray detector geometry. *Optics Express*, 23(22):28459, 2015.

[162] David G. Waterman, Graeme Winter, Richard J. Gildea, James M. Parkhurst, Aaron S. Brewster, Nicholas K. Sauter, and Gwyndaf Evans. Diffraction-geometry refinement in the DIALS framework. *Acta Crystallographica Section D: Structural Biology*, 72(4):558–575, 2016.

[163] Aaron S. Brewster, David G. Waterman, James M. Parkhurst, Richard J. Gildea, Iris D. Young, Lee J. O'Riordan, Junko Yano, Graeme Winter, Gwyndaf Evans, and Nicholas K. Sauter. Improving signal strength in serial crystallography with DIALS geometry refinement. *Acta Crystallographica Section D: Structural Biology*, 74(9):877–894, 2018.

[164] Helen Mary Ginn, Aaron S. Brewster, Johan Hattne, Gwyndaf Evans, Armin Wagner, Jonathan M. Grimes, Nicholas K. Sauter, Geoff Sutton, and David Ian Stuart. A revised partiality model and post-refinement algorithm for X-ray free-electron laser data. *Acta Crystallographica Section D: Biological Crystallography*, 71:1400–1410, 2015.

[165] Kay Diederichs and P. Andrew Karplus. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nature Structural Biology*, 4(4):269–275, apr 1997.

[166] P. A. Karplus and K. Diederichs. Linking Crystallographic Model and Data Quality. *Science*, 336(6084):1030–1033, may 2012.

[167] R. Henderson and J. K. Moffat. The difference Fourier technique in protein crystallography: errors and their treatment. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 27(7):1414–1420, jul 1971.

[168] Mark A. Rould and Charles W. Carter. Isomorphous Difference Methods. *Methods in Enzymology*, 374:145–163, 2003.

[169] P. L. Howell and G. D. Smith. Identification of heavy-atom derivatives by normal probability methods. *Journal of Applied Crystallography*, 25(1):81–86, feb 1992.

[170] Maria Cristina Burla, Rocco Caliandro, Carmelo Giacovazzo, and Giampiero Polidori. The difference electron density: A probabilistic reformulation. *Acta Crystallographica Section A: Foundations of Crystallography*, 66(3):347–361, 2010.

[171] A. J.C. C Wilson. Determination of absolute from relative X-ray intensity data. *Nature*, 150(3796):152, aug 1942.

[172] S. French and Krissy Wilson. On the treatment of negative intensity observations. *Acta Crystallographica Section A*, 34(4):517–525, jul 1978.

[173] A. J. C. Wilson. The probability distribution of X-ray intensities. *Acta Crystallographica*, 2(5):318–321, oct 1949.

[174] T. Ursby and D. Bourgeois. Improved Estimation of Structure-Factor Difference Amplitudesfrom Poorly Accurate Data. *Acta Crystallographica Section A Foundations of Crystallography*, 53(5):564–575, sep 1997.

[175] Thomas C. Terwilliger and Joel Berendzen. Bayesian difference refinement. *Acta Crystallographica Section D: Biological Crystallography*, 52(5):1004–1011, 1996.

[176] Paul D. Adams, Pavel V. Afonine, Gábor Bunkóczi, Vincent B. Chen, Ian W. Davis, Nathaniel Echols, Jeffrey J. Headd, Li-Wei Wei Hung, Gary J. Kapral, Ralf W. Grosse-Kunstleve, Airlie J. McCoy, Nigel W. Moriarty, Robert Oeffner, Randy J. Read, David C. Richardson, Jane S. Richardson, Thomas C. Terwilliger, and Peter H. Zwart. PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, 66(2):213–221, feb 2010.

[177] Martyn D Winn, Charles C Ballard, Kevin D Cowtan, Eleanor J Dodson, Paul Emsley, Phil R Evans, Ronan M Keegan, Eugene B Krissinel, Andrew G.W. W Leslie, Airlie McCoy, Stuart J. McNicholas, Garib N Murshudov, Navraj S Pannu, Elizabeth A Potterton, Harold R Powell, Randy J Read, Alexei Vagin, and Keith S Wilson. Overview of the CCP4 suite and current developments, 2011.

[178] Rudi Berera, Rienk van Grondelle, and John T.M. Kennis. Ultrafast transient absorption spectroscopy: Principles and application to photosynthetic systems. *Photosynthesis Research*, 101(2-3):105–118, 2009.

[179] Andreas Barth. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta - Bioenergetics*, 1767(9):1073–1101, 2007.

[180] Shaul Mukamel. *Principles of Nonlinear Optical Spectroscopy By Shaul Mukamel*, volume 118. Oxford University Press, 1996.

[181] Donna Strickland and Gerard Mourou. Compression of amplified chirped optical pulses. *Optics Communications*, 56(3):219–221, 1985.

[182] Bernhard Lang. Photometrics of ultrafast and fast broadband electronic transient absorption spectroscopy: State of the art. *Review of Scientific Instruments*, 89(9), 2018.

[183] Johanna Brazard, Laurie A. Bizimana, and Daniel B. Turner. Accurate convergence of transient-absorption spectra using pulsed lasers. *Review of Scientific Instruments*, 86(5), 2015.

[184] Ivo H.M. Van Stokkum, Delmar S. Larsen, and Rienk Van Grondelle. Global and target analysis of time-resolved spectra. *Biochimica et Biophysica Acta - Bioenergetics*, 1657(2-3):82–104, 2004.

[185] W. D. Hoff, I. H. van Stokkum, H. J. van Ramesdonk, M. E. van Brederode, A. M. Brouwer, J. C. Fitch, T E Meyer, R. van Grondelle, and K. J. Hellingwerf. Measurement and global analysis of the absorbance changes in the photocycle of the photoactive yellow protein from Ectothiorhodospira halophila. *Biophysical Journal*, 67(4):1691–1705, 1994.

[186] Luuk J.G.W. G W van Wilderen, Craig N. Lincoln, and Jasper J. van Thor. Modelling Multi-Pulse Population Dynamics from Ultrafast Spectroscopy. *PLoS ONE*, 6(3):e17373, mar 2011.

[187] Wolfgang Rudolph Jean-Claude Diels. *Ultrashort Laser Pulse Phenomena*. Elsevier, 2006.

[188] Ann E. Fitzpatrick, Craig N. Lincoln, Luuk J.G.W. Van Wilderen, and Jasper J. Van Thor. Pump-dump-probe and pump-repump-probe ultrafast spectroscopy resolves cross section of an early ground state intermediate and stimulated emission in the photoreactions of the Pr ground state of the cyanobacterial phytochrome Cph1. *Journal of Physical Chemistry B*, 116(3):1077–1088, 2012.

[189] Delmar S. Larsen, Ivo H.M. Van Stokkum, Mikas Vengris, Michael A. Van Der Horst, Frank L. De Weerd, Klaas J. Hellingwerf, and Rienk Van Grondelle. Incoherent manipulation of the photoactive yellow protein photocycle with dispersed pump-dump-probe spectroscopy. *Biophysical Journal*, 87(3):1858–1872, 2004.

[190] Devin W. Close, Craig Don Paul, Patricia S. Langan, Matthew C. J. Wilce, Daouda A. K. Traore, Randal Halfmann, Reginaldo C. Rocha, Geoffery S. Waldo, Riley J. Payne, Joseph B. Rucker, Mark Prescott, and Andrew R. M. Bradbury. Thermal green protein, an extremely stable, nonaggregating fluorescent protein created by structure-guided surface engineering. *Proteins: Structure, Function, and Bioinformatics*, 83(7):1225–1237, jul 2015.

[191] Razvan Nutiu and Yingfu Li. Structure-switching signaling aptamers. *Journal of the American Chemical Society*, 125(16):4771–4778, apr 2003.

[192] Tim Grotjohann, Ilaria Testa, Matthias Reuss, Tanja Brakemann, Christian Eggeling, Stefan W. Hell, and Stefan Jakobs. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife*, 2012(1):1–14, 2012.

[193] O. SHIMOMURA, F. H. JOHNSON, and Y. SAIGA. Extraction, purification and properties of aequorin, a bioluminescent. *Journal of cellular and comparative physiology*, 59(165):223–239, 1962.

[194] M Chalfie, Y Tu, G Euskirchen, W. Ward, and D. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805, feb 1994.

[195] Timothy D Craggs. Green fluorescent protein: Structure, folding and chromophore maturation. *Chemical Society Reviews*, 38(10):2865–2875, 2009.

[196] Roger Y. Tsien. The green fluorescent protein. *Annual Review of Biochemistry*, 67:509–544, 1998.

[197] R. Heim, A. B. Cubitt, and R. Y. Tsien. Improved green fluorescence. *Nature*, 373(February):663–664, 1995.

[198] M. Orm, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington. Crystal Structure of the Aequorea victoria Green Fluorescent Protein. *Science*, 273(5280):1392–1395, sep 1996.

[199] Atanu Acharya, Alexey M. Bogdanov, Bella L. Grigorenko, Ksenia B. Bravaya, Alexander V. Nemukhin, Konstantin A. Lukyanov, and Anna I. Krylov. Photoinduced chemistry in fluorescent proteins: Curse or blessing? *Chemical Reviews*, 117(2):758–795, 2017.

[200] Jasper J van Thor, Thomas Gensch, Klaas J Hellingwerf, and Louise N Johnson. Phototransformation of green fluorescent protein with UV and visible light leads to decarboxylation of glutamate 222. *Nature structural biology*, 9(1):37–41, 2002.

[201] George H. Patterson and Jennifer Lippincott-Schwartz. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science*, 297(5588):1873–1877, 2002.

[202] Jörg Wiedenmann, Sergey Ivanchenko, Franz Oswald, Florian Schmitt, Carlheinz Röcker, Anya Salih, Klaus Dieter Spindler, and G. Ulrich Nienhaus. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proceedings of the National Academy of Sciences of the United States of America*, 101(45):15905–15910, 2004.

[203] Dmitriy M. Chudakov, Vladislav V. Verkhusha, Dmitry B. Staroverov, Ekaterina A. Souslova, Sergey Lukyanov, and Konstantin A. Lukyanov. Photoswitchable cyan fluorescent protein for protein tracking. *Nature Biotechnology*, 22(11):1435–1439, 2004.

[204] G. Ulrich Nienhaus, Karin Nienhaus, Angela Hölzle, Sergey Ivanchenko, Fabiana Renzi, Franz Oswald, Michael Wolff, Florian Schmitt, Carlheinz Röcker, Beatrice Vallone, Wolfgang Weidemann, Ralf Heilker, Herbert Nar, and Jörg Wiedenmann. Photoconvertible Fluorescent Protein EosFP: Biophysical Properties and Cell Biology Applications. *Photochemistry and Photobiology*, 82(2):351, 2006.

[205] Yingying Ma, Qiao Sun, Hong Zhang, Liang Peng, Jian Guo Yu, and Sean C Smith. The mechanism of cyclization in chromophore maturation of green fluorescent protein: A theoretical study. *Journal of Physical Chemistry B*, 114(29):9698–9705, jul 2010.

[206] Ryoko Ando, Hideaki Mizuno, and Atsushi Miyawaki. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science*, 306(5700):1370–1373, 2004.

[207] Talley J. Lambert. FPbase: a community-editable fluorescent protein database. *Nature Methods*, 16(4):277–278, 2019.

[208] Eduardo Carrascosa, James N. Bull, Michael S. Scholz, Neville J.A. A Coughlan, Seth Olsen, Uta Wille, and Evan J. Bieske. Reversible Photoisomerization of the Isolated Green Fluorescent Protein Chromophore. *Journal of Physical Chemistry Letters*, 9(10):2647–2651, 2018.

[209] Mats G.L. L Gustafsson. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 102(37):13081–13086, 2005.

[210] Virgile Adam. Phototransformable fluorescent proteins: Which one for which application? *Histochemistry and Cell Biology*, 142(1):19–41, 2014.

[211] Christopher D.M. M Hutchison, Marius Kaucikas, Jason Tenboer, Christopher Kupitz, Keith Moffat, Marius Schmidt, and Jasper J. van Thor. Photocycle populations with femtosecond excitation of crystalline photoactive yellow protein. *Chemical Physics Letters*, 654:63–71, jun 2016.

[212] Nicolas Coquelle, Michel Sliwa, Joyce Woodhouse, Giorgio Schirò, Virgile Adam, Andrew Aquila, Thomas R.M. M Barends, Sébastien Boutet, Martin Byrdin, Sergio Carbajo, Eugenio De La Mora, R. Bruce Doak, Mikolaj Feliks, Franck Fieschi, Lutz Foucar, Virginia Guillon, Mario Hilpert, Mark S Hunter, Stefan Jakobs, Jason E Koglin, Gabriela Kovacsova, Thomas J Lane, Bernard Lévy, Mengning Liang, Karol Nass, Jacqueline Ridard, Joseph S Robinson, Christopher M Roome, Cyril Ruckebusch, Matthew Seaberg, Michel Thepaut, Marco Cammarata, Isabelle Demachy, Martin Field, Robert L Shoeman, Dominique Bourgeois, Jacques Philippe Colletier, Ilme Schlichting, and Martin Weik. Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature Chemistry*, 10(1):31–37, 2018.

[213] Satoshi Habuchi, Peter Dedecker, Jun Ichi Hotta, Cristina Flors, Ryoko Ando, Hideaki Mizuno, Atsushi Miyawaki, and Johan Hofkens. Photo-induced protonation/deprotonation in the GFP-like fluorescent protein Dronpa: Mechanism responsible for the reversible photoswitching. *Photochemical and Photobiological Sciences*, 5(6):567–576, 2006.

[214] A. L. Patterson. A Fourier Series Method for the Determination of the Components of Interatomic Distances in Crystals. *Physical Review*, 46(5):372–376, sep 1934.

[215] Roger Heim and Roger Y. Tsien. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology*, 6(2):178–182, 1996.

[216] Eduard Fron, Cristina Flors, Gerd Schweitzer, Satoshi Habuchi, Hideaki Mizuno, Ryoko Ando, Frans C. De Schryver, Atsushi Miyawaki, and Johan Hofkens. Ultrafast excited-state dynamics of the photoswitchable protein dronpa. *Journal of the American Chemical Society*, 129(16):4870–4871, 2007.

[217] Mita Chattoraj, Brett A. King, Gerold U. Bublitz, and Stevrn G. Boxer. Ultra-fast excited state dynamics in green fluorescent protein: Multiple states and proton transfer. *Proceedings of the National Academy of Sciences of the United States of America*, 93(16):8362–8367, 1996.

[218] Pascal G Wilmann, Kristina Turcic, Jion M Battad, Matthew C.J. J Wilce, Rodney J Devenish, Mark Prescott, and Jamie Rossjohn. The 1.7 Å Crystal Structure of Dronpa: A Photoswitchable Green Fluorescent Protein. *Journal of Molecular Biology*, 364(2):213–224, nov 2006.

[219] Marius Kaucikas, Ann Fitzpatrick, Elana Bryan, Abelone Struve, Robert Henning, Irina Kosheleva, Vukica Srajer, Gerrit Groenhof, and Jasper J. Van Thor. Room temperature crystal structure of the fast switching M159T mutant of the fluorescent protein dronpa. *Proteins: Structure, Function and Bioinformatics*, 83(3):397–402, 2015.

[220] George H. Patterson, Susan M. Knobel, Wallace D. Sharif, Steven R. Kain, and David W. Piston. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical Journal*, 73(5):2782–2790, 1997.

[221] Mingshu Zhang, Zhifei Fu, and Pingyong Xu. Extending the spatiotemporal resolution of super-resolution microscopies using photomodulatable fluorescent proteins. *Journal of Innovative Optical Health Sciences*, 09(03):1630009, 2016.

[222] Fedor V. Subach, Lijuan Zhang, Theodorus W.J. J Gadella, Nadya G. Gurskaya, Konstantin A Lukyanov, and Vladislav V Verkhusha. Red Fluorescent Protein with Reversibly Photoswitchable Absorbance for Photochromic FRET. *Chemistry & Biology*, 17(7):745–755, jul 2010.

[223] Patrick J. Macdonald, Yan Chen, and Joachim D. Mueller. Chromophore maturation and fluorescence fluctuation spectroscopy of fluorescent proteins in a cell-free expression system. *Analytical Biochemistry*, 421(1):291–298, feb 2012.

[224] Matthew G Romei, Chi-yun Lin, Irimpan I Mathews, and Steven G Boxer. Electrostatic control of photoisomerization pathways in proteins. *Science*, 367(6473):76–79, jan 2020.

[225] Jion M. Battad, Daouda A.K. K Traore, Emma Byres, Jamie Rossjohn, Rodney J. Devenish, Seth Olsen, Matthew C.J. J Wilce, and Mark Prescott. A Green Fluorescent Protein Containing a QFG Tri-Peptide Chromophore: Optical Properties and X-Ray Crystal Structure. *PLoS ONE*, 7(10):1–14, 2012.

[226] Hao Chang, Mingshu Zhang, Wei Ji, Juanjuan Chen, Yongdeng Zhang, Bei Liu, Jingze Lu, Junlong Zhang, Pingyong Xu, and Tao Xu. A unique series of reversibly switchable fluorescent proteins with beneficial properties for various applications. *Proceedings of the National Academy of Sciences of the United States of America*, 109(12):4455–4460, 2012.

[227] Virgile Adam, Mickaël Lelimousin, Susan Boehme, Guillaume Desfonds, Karin Nienhaus, Martin J. Field, Joerg Wiedenmann, Sean McSweeney, G. Ulrich Nienhaus, and Dominique Bourgeois. Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47):18343–18348, 2008.

[228] Andre C. Stiel, Simon Trowitzsch, Gert Weber, Martin Andresen, Christian Eggeling, Stefan W. Hell, Stefan Jakobs, and Markus C. Wahl. 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. *Biochemical Journal*, 402(1):35–42, feb 2007.

[229] Ryoko Ando, Cristina Flors, Hideaki Mizuno, Johan Hofkens, and Atsushi Miyawaki. Highlighted generation of fluorescence signals using simultaneous two-color irradiation on Dronpa mutants. *Biophysical Journal*, 92(12):L97–L99, 2007.

[230] María Elena Martin, Fabrizia Negri, and Massimo Olivucci. Origin, Nature, and Fate of the Fluorescent State of the Green Fluorescent Protein Chromophore at the CASPT2//CASSCF Resolution. *Journal of the American Chemical Society*, 126(17):5452–5464, 2004.

[231] Jeffrey Chang, Matthew G. Romei, and Steven G. Boxer. Structural Evidence of Photoisomerization Pathways in Fluorescent Proteins. *Journal of the American Chemical Society*, 141(39):15504–15508, 2019.

[232] Simplex Collection. HHS Public Access. 8(5):583–592, 2016.

[233] Dheerendra Yadav, Fabien Lacombat, Nadia Dozova, Fabrice Rappaport, Pascal Plaza, and Agathe Espagne. Real-time monitoring of chromophore isomerization and deprotonation during the

photoactivation of the fluorescent protein Dronpa. *Journal of Physical Chemistry B*, 119(6):2404–2414, feb 2015.

[234] Xin Li, Lung Wa Chung, Hideaki Mizuno, Atsushi Miyawaki, and Keiji Morokuma. A theoretical study on the nature of on-and off-states of reversibly photoswitching fluorescent protein dronpa: Absorption, emission, protonation, and Raman. *Journal of Physical Chemistry B*, 114(2):1114–1126, 2010.

[235] Jacques Philippe Colletier, Michel Sliwa, François Xavier Gallat, Michihiro Sugahara, Virginia Guillon, Giorgio Schirò, Nicolas Coquelle, Joyce Woodhouse, Laure Roux, Guillaume Gotthard, Antoine Royant, Lucas Martinez Uriarte, Cyril Ruckebusch, Yasumasa Joti, Martin Byrdin, Eiichi Mizohata, Eriko Nango, Tomoyuki Tanaka, Kensuke Tono, Makina Yabashi, Virgile Adam, Marco Cammarata, Ilme Schlichting, Dominique Bourgeois, and Martin Weik. Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *Journal of Physical Chemistry Letters*, 7(5):882–887, 2016.

[236] Joyce Woodhouse, Gabriela Nass Kovacs, Nicolas Coquelle, Lucas M. Uriarte, Virgile Adam, Thomas R.M. M Barends, Martin Byrdin, Eugenio de la Mora, R. Bruce Doak, Mikolaj Feliks, Martin Field, Franck Fieschi, Virginia Guillon, Stefan Jakobs, Yasumasa Joti, Pauline Macheboeuf, Koji Motomura, Karol Nass, Shigeki Owada, Christopher M. Roome, Cyril Ruckebusch, Giorgio Schirò, Robert L. Shoeman, Michel Thepaut, Tadashi Togashi, Kensuke Tono, Makina Yabashi, Marco Cammarata, Lutz Foucar, Dominique Bourgeois, Michel Sliwa, Jacques Philippe Colletier, Ilme Schlichting, and Martin Weik. Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature Communications*, 11(1):1–11, 2020.

[237] Garib N Murshudov, Pavol Skubák, Andrey A Lebedev, Navraj S Pannu, Roberto A Steiner, Robert A Nicholls, Martyn D Winn, Fei Long, and Alexei A Vagin. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):355–367, 2011.

[238] Martyn D. Winn, Charles C. Ballard, Kevin D. Cowtan, Eleanor J. Dodson, Paul Emsley, Phil R. Evans, Ronan M. Keegan, Eugene B. Krissinel, Andrew G. W. Leslie, Airlie McCoy, Stuart J. McNicholas, Garib N. Murshudov, Navraj S. Pannu, Elizabeth A. Potterton, Harold R. Powell, Randy J. Read, Alexei Vagin, and Keith S. Wilson. Overview of the CCP 4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography*, 67(4):235–242, apr 2011.

[239] Marius Schmidt, Sudarshan Rajagopal, Zhong Ren, and Keith Moffat. Application of Singular

Value Decomposition to the Analysis of Time-Resolved Macromolecular X-Ray Data. *Biophysical Journal*, 84(3):2112–2129, mar 2003.

[240] Zhong Ren, Benjamin Perman, Vukica Šrajer, T.-Y. Teng, Claude Pradervand, Dominique Bourgeois, Friederich Schotte, Thomas Ursby, Remco Kort, Michael Wulff, and Keith Moffat. A Molecular Movie at 1.8 ÅResolution Displays the Photocycle of Photoactive Yellow Protein, a Eubacterial Blue-Light Receptor, from Nanoseconds to Seconds †. *Biochemistry*, 40(46):13788–13801, nov 2001.

[241] Suraj Pandey, Richard Bean, Tokushi Sato, Ishwor Poudyal, Johan Bielecki, Jorvani Cruz Villarreal, Oleksandr Yefanov, Valerio Mariani, Thomas A White, Christopher Kupitz, Mark Hunter, Mohamed H Abdellatif, Saša Bajt, Valerii Bondar, Austin Echelmeier, Diandra Doppler, Moritz Emons, Matthias Frank, Raimund Fromme, Yaroslav Gevorkov, Gabriele Giovanetti, Man Jiang, Daihyun Kim, Yoonhee Kim, Henry Kirkwood, Anna Klimovskaia, Juraj Knoska, Faisal H M Koua, Romain Letrun, Stella Lisova, Luis Maia, Victoria Mazalova, Domingo Meza, Thomas Michelat, Abbas Ourmazd, Guido Palmer, Marco Ramilli, Robin Schubert, Peter Schwander, Alessandro Silenzi, Jolanta Sztuk-Dambietz, Alexandra Tolstikova, Henry N Chapman, Alexandra Ros, Anton Barty, Petra Fromme, Adrian P Mancuso, and Marius Schmidt. Time-resolved serial femtosecond crystallography at the European XFEL. *Nature Methods*, 17(1):73–78, jan 2019.

[242] Donald G Truhlar and Amnon Kohen. Convex Arrhenius plots and their interpretation. *Proceedings of the National Academy of Sciences*, 98(3):848–851, jan 2002.

[243] Richard C. Tolman. Statistical mechanics applied to chemical kinetics. *Journal of the American Chemical Society*, 42(12):2506–2528, dec 1920.

[244] V. Massey, B. Curti, and H. Ganther. A temperature-dependent conformational change in D-amino acid oxidase and its effect on catalysis. *Journal of Biological Chemistry*, 241(10):2347–2357, 1966.

[245] Hans Frauenfelder, Guo Chen, Joel Berendzen, Paul W. Fenimore, Helén Jansson, Benjamin H. McMahon, Izabela R. Stroe, Jan Swenson, and Robert D. Young. A unified model of protein dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13):5129–5134, 2009.

[246] Clemens Vonrhein, Claus Flensburg, Peter Keller, Andrew Sharff, Oliver Smart, Wlodek Paciorek, Thomas Womack, and Gérard Bricogne. Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):293–302, 2011.

[247] Wolfgang Kabsch. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D Biological Crystallography*, 66(2):133–144, feb 2010.

[248] Philip R Evans and Garib N Murshudov. How good are my data and what is the resolution? *Acta Crystallographica Section D Biological Crystallography*, 69(7):1204–1214, 2013.

[249] I. J Tickle, C. Flensburg, P. Keller, W. Paciorek, A Sharff, C. Vonrhein, and G Bricogne. STARANISO (Website), 2018.

[250] Cecilia Wickstrand, Gergely Katona, Takanori Nakane, Przemyslaw Nogly, Joerg Standfuss, Eriko Nango, and Richard Neutze. *Structural Dynamics*, (2):024701.

[251] Aline Regis Faro, Philippe Carpentier, Gabriella Jonasson, Guillaume Pompidor, Delphine Arcizet, Isabelle Demachy, and Dominique Bourgeois. Low-temperature chromophore isomerization reveals the photoswitching mechanism of the fluorescent protein Padron. *Journal of the American Chemical Society*, 133(41):16362–16365, 2011.

[252] Ya Ting Kao, Xinxin Zhu, and Wei Min. Protein-flexibility mediated coupling between photoswitching kinetics and surrounding viscosity of a photochromic fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9):3220–3225, 2012.

[253] E. R. Henry, W. A. Eaton, and R. M. Hochstrasser. Molecular dynamics simulations of cooling in laser-excited heme proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 83(23):8982–8986, 1986.

[254] P. Li and P. M. Champion. Investigations of the thermal response of laser-excited biomolecules. *Biophysical Journal*, 66(2):430–436, 1994.

[255] Martin Weik, Raimond B G Ravelli, Gitay Kryger, Sean McSweeney, Maria L Raves, Michal Harel, Piet Gros, Israel Silman, Jan Kroon, and Joel L Sussman. Specific chemical and structural damage to proteins produced by synchrotron radiation. Technical report, 1999.

[256] Alfred Vogel and Vasan Venugopalan. Pulsed laser ablation of soft biological tissues. *Optical-Thermal Response of Laser-Irradiated Tissue*, pages 551–615, 2011.

[257] Delmar S. Larsen, Mikas Vengris, Ivo H.M. Van Stokkum, Michael A. Van Der Horst, Frank L. De Weerd, Klaas J. Hellingwerf, and Rienk Van Grondelle. Photoisomerization and Photoionization of the Photoactive Yellow Protein Chromophore in Solution. *Biophysical Journal*, 86(4):2538–2550, 2004.

[258] Craig N. Lincoln, Ann E. Fitzpatrick, and Jasper J.Van Thor. Photoisomerisation quantum yield and non-linear cross-sections with femtosecond excitation of the photoactive yellow protein. *Physical Chemistry Chemical Physics*, 14(45):15752–15764, 2012.

[259] Kanupriya Pande, Christopher D M Hutchison, Gerrit Groenhof, Andy Aquila, Josef S. Robinson, Jason Tenboer, Shibom Basu, Sébastien Boutet, Daniel P. DePonte, Mengning Liang, Thomas A. White, Nadia A Zatsepin, Oleksandr Yefanov, Dmitry Morozov, Dominik Oberthuer, Cornelius Gati, Ganesh Subramanian, Daniel James, Yun Zhao, Jake Koralek, Jennifer Brayshaw, Christopher Kupitz, Chelsie Conrad, Shatabdi Roy-Chowdhury, Jesse D Coe, Markus Metz, Paulraj Lourdu Xavier, Thomas D Grant, Jason E. Koglin, Gihan Ketawala, Raimund Fromme, Vukica Šrajer, Robert Henning, John C H Spence, Abbas Ourmazd, Peter Schwander, Uwe Weierstall, Matthias Frank, Petra Fromme, Anton Barty, Henry N. Chapman, Keith Moffat, Jasper J van Thor, and Marius Schmidt. Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science (New York, N.Y.)*, 352(6286):725–9, 2016.

[260] Marie Louise Groot, Luuk J.G.W. Van Wilderen, Delmar S. Larsen, Michael A. Van der Horst, Ivo H.M. Van Stokkum, Klaas J. Hellingwerf, and Rienk Van Grondelle. Initial steps of signal generation in photoactive yellow protein revealed with femtosecond mid-infrared spectroscopy. *Biochemistry*, 42(34):10054–10059, 2003.

[261] Sergey Yeremenko, Ivo H.M. M Van Stokkum, Keith Moffat, and Klaas J. Hellingwerf. Influence of the crystalline state on photoinduced dynamics of photoactive yellow protein studied by ultraviolet-visible transient absorption spectroscopy. *Biophysical Journal*, 90(11):4224–4235, 2006.

[262] Christopher D M Hutchison and Jasper J. van Thor. Populations and coherence in femtosecond time resolved X-ray crystallography of the photoactive yellow protein. *International Reviews in Physical Chemistry*, 36(1):117–143, 2017.

[263] Gabriela Nass Kovacs, Jacques-Philippe Colletier, Marie Luise Grünbein, Yang Yang, Alexander Batyuk, Sergio Carbajo, R. Bruce Doak, David Ehrenberg, Lutz Foucar, Raphael Gasper, Alexander Gorel, Mario Hilpert, Marco Kloos, Jason E. Koglin, Jochen Reinstein, Christopher M. Roome, Ramona Schlesinger, Matthew Seaberg, Robert L. Shoeman, Miriam Stricker, Sébastien Boutet, Stefan Haacke, Joachim Heberle, Karsten Heyne, Tatiana Domratcheva, Thomas R. M. Barends, Ilme Schlichting, Till Stensitzki, Alexander Batyuk, Sergio Carbajo, R. Bruce Doak, David Ehrenberg, Lutz Foucar, Raphael Gasper, Alexander Gorel, Mario Hilpert, Marco Kloos, Jason E. Koglin, Jochen Reinstein, Christopher M. Roome, Ramona Schlesinger, Matthew Seaberg, Robert L. Shoeman, Miriam Stricker, Sébastien Boutet, Stefan Haacke, Joachim Heberle, Karsten Heyne, Tatiana Domratcheva, Thomas R. M. Barends, and Ilme Schlichting. Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin. *Nature Communications*, 10(1):3177, dec 2019.

[264] Przemyslaw Nogly, Tobias Weinert, Daniel James, Sergio Carbajo, Dmitry Ozerov, Antonia Furrer, Dardan Gashi, Veniamin Borin, Petr Skopintsev, Kathrin Jaeger, Karol Nass, Petra Båth, Robert

Bosman, Jason Koglin, Matthew Seaberg, Thomas Lane, Demet Kekilli, Steffen Brünle, Tomoyuki Tanaka, Wenting Wu, Christopher Milne, Thomas White, Anton Barty, Uwe Weierstall, Valerie Panneels, Eriko Nango, So Iwata, Mark Hunter, Igor Schapiro, Gebhard Schertler, Richard Neutze, and Jörg Standfuss. Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science*, 361(6398):1–15, 2018.

[265] Cecilia Wickstrand, Przemyslaw Nogly, Eriko Nango, So Iwata, Jörg Standfuss, and Richard Neutze. Bacteriorhodopsin: Structural insights revealed using x-ray lasers and synchrotron radiation. *Annual Review of Biochemistry*, 88:59–83, 2019.

[266] Matteo Levantino, Giorgio Schirò, Henrik Till Lemke, Grazia Cottone, James Michael Glownia, Diling Zhu, Mathieu Chollet, Hyotcherl Ihee, Antonio Cupane, and Marco Cammarata. Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nature Communications*, 6:1–6, 2015.

[267] A. Ansari and A. Szabo. Theory of photoselection by intense light pulses. Influence of reorientational dynamics and chemical kinetics on absorbance measurements. *Biophysical Journal*, 64(3):838–851, 1993.

[268] David Arnlund, Linda C. Johansson, Cecilia Wickstrand, Anton Barty, Garth J. Williams, Erik Malmerberg, Jan Davidsson, Despina Milathianaki, Daniel P. DePonte, Robert L. Shoeman, Dingjie Wang, Daniel James, Gergely Katona, Sebastian Westenhoff, Thomas A. White, Andrew Aquila, Sadia Bari, Peter Berntsen, Mike Bogan, Tim Brandt Van Driel, R. Bruce Doak, Kasper Skov Kjær, Matthias Frank, Raimund Fromme, Ingo Grotjohann, Robert Henning, Mark S. Hunter, Richard A. Kirian, Irina Kosheleva, Christopher Kupitz, Mengning Liang, Andrew V. Martin, Martin Meedom Nielsen, Marc Messerschmidt, M. Marvin Seibert, Jennie Sjöhamn, Francesco Stellato, Uwe Weierstall, Nadia A. Zatsepin, John C.H. Spence, Petra Fromme, Ilme Schlichting, Sébastien Boutet, Gerrit Groenhof, Henry N. Chapman, and Richard Neutze. Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser. *Nature Methods*, 11(9):923–926, aug 2014.

[269] Andrei C. Florean, David Cardoza, James L. White, J. K. Lanyi, Roseanne J. Sension, and Philip H. Bucksbaum. Control of retinal isomerization in bacteriorhodopsin in the high-intensity regime. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27):10896–10900, jul 2009.

[270] Valentyn I Prokhorenko, Alexei Halpin, Philip J.M. Johnson, R. J.Dwayne Miller, and Leonid S Brown. Coherent control of the isomerization of retinal in bacteriorhodopsin in the high intensity regime. *Journal of Chemical Physics*, 134(8):85105, 2011.

[271] M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim, and M. Elstner. Calculating absorption shifts for retinal proteins: Computational challenges. *Journal of Physical Chemistry B*, 109(8):3606–3615, 2005.

[272] C. J. Bardeen, Q. Wang, and C. V. Shank. Selective excitation of vibrational wave packet motion using chirped pulses. *Physical Review Letters*, 75(19):3410–3413, 1995.

[273] Valentyn I. Prokhorenko. Coherent Control of Retinal Isomerization in Bacteriorhodopsin. *Science*, 313(5791):1257–1261, sep 2006.

[274] Joyce Woodhouse and Joyce Woodhouse. Studying a reversibly switchable fluorescent protein by time-resolved crystallography using the X-ray free electron lasers To cite this version : HAL Id : tel-02305244 Etude d ' une protéine fluorescente photo-commutable par cristallographie résolue en te. 2019.

[275] Zhirong Huang and Ingolf Lindau. Free-electron lasers: SACLA hard-X-ray compact FEL. *Nature Photonics*, 6(8):505–506, 2012.

[276] Takashi Kameshima, Shun Ono, Togo Kudo, Kyosuke Ozaki, Yoichi Kirihara, Kazuo Kobayashi, Yuichi Inubushi, Makina Yabashi, Toshio Horigome, Andrew Holland, Karen Holland, David Burt, Hajime Murao, and Takaki Hatsui. Development of an X-ray pixel detector with multi-port charge-coupled device for X-ray free-electron laser experiments. *Review of Scientific Instruments*, 85(3):0–15, 2014.

[277] Robert W Boyd. *Nonlinear Optics, Third Edition*. Academic Press, Inc., USA, 3rd edition, 2008.

[278] Christopher D M Hutchison and Jasper J. van Thor. Optical control, selection and analysis of population dynamics in ultrafast protein X-ray crystallography. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 377(2145):20170474, may 2019.

[279] Rick Trebino, Kenneth W. DeLong, David N. Fittinghoff, John N. Sweetser, Marco A. Krumbügel, Bruce A. Richman, and Daniel J. Kane. Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating. *Review of Scientific Instruments*, 68(9):3277–3295, 1997.

[280] S. Linden, H. Giessen, and J. Kuhl. XFROG - A new method for amplitude and phase characterization of weak ultrashort pulses. *Physica Status Solidi (B) Basic Research*, 206(1):119–124, mar 1998.

[281] Michael Diez, Andreas Galler, Sebastian Schulz, Christina Boemer, Ryan N. Coffee, Nick Hartmann, Rupert Heider, Martin S. Wagner, Wolfram Helml, Tetsuo Katayama, Tokushi Sato,

Takahiro Sato, Makina Yabashi, and Christian Bressler. A self-referenced in-situ arrival time monitor for X-ray free-electron lasers. *Scientific Reports*, 11(1):3562, dec 2021.

[282] Yongbin Xu, Kwang Yeon Hwang, and Ki Hyun Nam. Spectral and structural analysis of large Stokes shift fluorescent protein dKeima570. *Journal of Microbiology*, 56(11):822–827, 2018.

[283] Michael Fleischhauer, Atac Imamoglu, and P. Jonathan Marangos. Electromagnetically induced transparency. *Reviews of Modern Physics*, 77(2):633–673, 2005.

[284] T. Geoff G. Battye, Luke Kontogiannis, Owen Johnson, Harold R. Powell, and Andrew G.W. W Leslie. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):271–281, 2011.

[285] Takanori Nakane, Yasumasa Joti, Kensuke Tono, Makina Yabashi, Eriko Nango, So Iwata, Ryuichiro Ishitani, and Osamu Nureki. Data processing pipeline for serial femtosecond crystallography at SACLA. *Journal of Applied Crystallography*, 49:1035–1041, 2016.

[286] Thomas A White. Processing serial crystallography data with CrystFEL : a step-by-step guide. *Acta Crystallographica Section D Structural Biology*, 75(2):219–233, 2019.

[287] J. KRAUT, L. C. SIEKER, D. F. HIGH, and S. T. FREER. Chymotrypsinogen: a three-dimensional fourier synthesis at 5 angstrom resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 48(8):1417–1424, aug 1962.

[288] P. V. Afonine, R. W. Grosse-Kunstleve, P. D. Adams, and A. Urzhumtsev. Bulk-solvent and overall scaling revisited: Faster calculations, improved results. *Acta Crystallographica Section D: Biological Crystallography*, 69(4):625–634, 2013.

[289] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of Coot. *Acta Crystallographica Section D Biological Crystallography*, 66(4):486–501, apr 2010.

[290] Jasper J. van Thor. Advances and opportunities in ultrafast X-ray crystallography and ultrafast structural optical crystallography of nuclear and electronic protein dynamics. *Structural Dynamics*, 6(5):050901, sep 2019.

[291] Anand T.N. Kumar, Florin Rosca, Allan Widom, and Paul M. Champion. Investigations of amplitude and phase excitation profiles in femtosecond coherence spectroscopy. *Journal of Chemical Physics*, 114(2):701–724, 2001.

[292] Xiong Ye, Dan Ionascu, Florin Gruia, Anchi Yu, Abdelkrim Benabbas, and Paul M Champion. Temperature-dependent heme kinetics with nonexponential binding and barrier relaxation in the absence of protein conformational substates. *Proceedings of the National Academy of Sciences of the United States of America*, 104(37):14682–14687, 2007.

[293] C. J. Bardeen, Q. Wang, and C. V. Shank. Femtosecond chirped pulse excitation of vibrational wave packets in LD690 and bacteriorhodopsin. *Journal of Physical Chemistry A*, 102(17):2759–2766, 1998.

[294] Jason Tenboer, Shibom Basu, Nadia Zatsepin, Kanupriya Pande, Despina Milathianaki, Matthias Frank, Mark Hunter, Sébastien Boutet, Garth J. Williams, Jason E. Koglin, Dominik Oberthuer, Michael Heymann, Christopher Kupitz, Chelsie Conrad, Jesse Coe, Shatabdi Roy-Chowdhury, Uwe Weierstall, Daniel James, Dingjie Wang, Thomas Grant, Anton Barty, Oleksandr Yefanov, Jennifer Scales, Cornelius Gati, Carolin Seuring, Vukica Srajer, Robert Henning, Peter Schwander, Raimund Fromme, Abbas Ourmazd, Keith Moffat, Jasper J. Van Thor, John C. H. Spence, Petra Fromme, Henry N. Chapman, and Marius Schmidt. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science*, 346(6214):1242–1246, dec 2014.

[295] Ali Ebrahim, Tadeo Moreno-Chicano, Martin V. Appleby, Amanda K. Chaplin, John H. Beale, Darren A. Sherrell, Helen M. E. Duyvesteyn, Shigeki Owada, Kensuke Tono, Hiroshi Sugimoto, Richard W. Strange, Jonathan A. R. Worrall, Danny Axford, Robin L. Owen, and Michael A. Hough. Dose-resolved serial synchrotron and XFEL structures of radiation-sensitive metalloproteins. *IUCrJ*, 6(4):543–551, jul 2019.

[296] Cornelius Gati, Gleb Bourenkov, Marco Klinge, Dirk Rehders, Francesco Stellato, Dominik Oberthür, Oleksandr Yefanov, Benjamin P. Sommer, Stefan Mogk, Michael Duszenko, Christian Betzel, Thomas R. Schneider, Henry N. Chapman, and Lars Redecke. Serial crystallography on in vivo grown microcrystals using synchrotron radiation. *IUCrJ*, 1:87–94, 2014.

[297] Elspeth F. Garman and Robin Leslie Owen. Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*, 62(1):32–47, jan 2006.

[298] Douglas H. Juers and Brian W. Matthews. The role of solvent transport in cryo-annealing of macromolecular crystals. *Acta Crystallographica Section D: Biological Crystallography*, 60(3):412–421, 2004.

[299] Christian G. Schroer, Ilya Agapov, Werner Brefeld, Reinhard Brinkmann, Yong Chul Chae, Hung Chun Chao, Mikael Eriksson, Joachim Keil, Xavier Nuel Gavaldà, Ralf Röhlsberger, Oliver H. Seeck, Michael Sprung, Markus Tischer, Rainer Wanzenberg, and Edgar Weckert. PETRA IV: the ultralow-emittance source project at DESY. *Journal of Synchrotron Radiation*, 25(5):1277–1290, 2018.

[300] Pantaleo Raimondi. ESRF-EBS: The Extremely Brilliant Source Project. *Synchrotron Radiation News*, 29(6):8–15, nov 2016.

[301] Filip Leonarski, Aldo Mozzanica, Martin Brückner, Carlos Lopez-Cuenca, Sophie Redford, Leonardo Sala, Andrej Babic, Heinrich Billich, Oliver Bunk, Bernd Schmitt, and Meitian Wang. JUNGFRAU detector for brighter X-ray sources: Solutions for IT and data science challenges in macromolecular crystallography. *Structural Dynamics*, 7(1):1–13, 2020.

[302] Eriko Nango, Antoine Royant, Minoru Kubo, Takanori Nakane, Cecilia Wickstrand, Tetsunari Kimura, Tomoyuki Tanaka, Kensuke Tono, Changyong Song, Rie Tanaka, Toshi Arima, Ayumi Yamashita, Jun Kobayashi, Toshiaki Hosaka, Eiichi Mizohata, Przemyslaw Nogly, Michihiro Sugahara, Daewoong Nam, Takashi Nomura, Tatsuro Shimamura, Dohyun Im, Takaaki Fujiwara, Yasuaki Yamanaka, Byeonghyun Jeon, Tomohiro Nishizawa, Kazumasa Oda, Masahiro Fukuda, Rebecka Andersson, Petra Båth, Robert Dods, Jan Davidsson, Shigeru Matsuoka, Satoshi Kawatake, Michio Murata, Osamu Nureki, Shigeki Owada, Takashi Kameshima, Takaki Hatsui, Yasumasa Joti, Gebhard Schertler, Makina Yabashi, Ana-Nicoleta Bondar, Jörg Standfuss, Richard Neutze, and So Iwata. A three-dimensional movie of structural changes in bacteriorhodopsin. *Science*, 354(6319):1552–1557, 2016.

[303] J. R. Helliwell. Synchrotron X-radiation protein crystallography: Instrumentation, methods and applications. *Reports on Progress in Physics*, 47(11):1403–1497, 1984.

[304] James Nicholson, Colin Nave, Khalid Fayz, Barry Fell, and Elspeth Garman. Modelling heating effects in cryocooled protein crystals. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 467-468(PART II):1380–1383, 2001.

[305] S. Kriminski, M. Kazmierczak, and R. E. Thorne. Heat transfer from protein crystals: implications for flash-cooling and X-ray beam heating. *Acta Crystallographica - Section D Biological Crystallography*, 59(4):697–708, apr 2003.

[306] Ashutosh Mhaisekar, Michael J. Kazmierczak, and Rupak Banerjee. Three-dimensional numerical analysis of convection and conduction cooling of spherical biocrystals with localized heating from synchrotron X-ray beams. *Journal of Synchrotron Radiation*, 12(3):318–328, 2005.

[307] Anna J. Warren, Danny Axford, and Robin L. Owen. Direct measurement of X-ray-induced heating of microcrystals. *Journal of Synchrotron Radiation*, 26(4):991–997, jul 2019.

[308] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J.

van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R.J. J Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, Aditya Vijaykumar, Alessandro Pietro Bardelli, Alex Rothberg, Andreas Hilboll, Andreas Kloeckner, Anthony Scopatz, Antony Lee, Ariel Rokem, C. Nathan Woods, Chad Fulton, Charles Masson, Christian Häggström, Clark Fitzgerald, David A. Nicholson, David R. Hagen, Dmitrii V. Pasechnik, Emanuele Olivetti, Eric Martin, Eric Wieser, Fabrice Silva, Felix Lenders, Florian Wilhelm, G. Young, Gavin A. Price, Gert Ludwig Ingold, Gregory E. Allen, Gregory R. Lee, Hervé Audren, Irvin Probst, Jörg P. Dietrich, Jacob Silterra, James T. Webber, Janko Slavič, Joel Nothman, Johannes Buchner, Johannes Kulick, Johannes L. Schönberger, José Vinícius de Miranda Cardoso, Joscha Reimer, Joseph Harrington, Juan Luis Cano Rodríguez, Juan Nunez-Iglesias, Justin Kuczynski, Kevin Tritz, Martin Thoma, Matthew Newville, Matthias Kümmerer, Maximilian Bolingbroke, Michael Tartre, Mikhail Pak, Nathaniel J. Smith, Nikolai Nowaczyk, Nikolay Shebanov, Oleksandr Pavlyk, Per A. Brodtkorb, Perry Lee, Robert T. McGibbon, Roman Feldbauer, Sam Lewis, Sam Tygier, Scott Sievert, Sebastiano Vigna, Stefan Peterson, Surhud More, Tadeusz Pudlik, Takuya Oshima, Thomas J. Pingel, Thomas P. Robitaille, Thomas Spura, Thouis R. Jones, Tim Cera, Tim Leslie, Tiziano Zito, Tom Krauss, Utkarsh Upadhyay, Yaroslav O. Halchenko, and Yoshiki Vázquez-Baeza. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020.

[309] Marcos A. de Araújo, Rubens Silva, Emerson de Lima, Daniel P. Pereira, and Paulo C. de Oliveira. Measurement of Gaussian laser beam radius using the knife-edge technique: improvement on data analysis. *Applied Optics*, 48(2):393, jan 2009.

[310] Dorothee Liebschner, Pavel V. Afonine, Matthew L. Baker, Gábor Bunkoczi, Vincent B. Chen, Tristan I. Croll, Bradley Hintze, Li Wei Hung, Swati Jain, Airlie J. McCoy, Nigel W. Moriarty, Robert D. Oeffner, Billy K. Poon, Michael G. Prisant, Randy J. Read, Jane S. Richardson, David C. Richardson, Massimo D. Sammito, Oleg V. Sobolev, Duncan H. Stockwell, Thomas C. Terwilliger, Alexandre G. Urzhumtsev, Lizbeth L. Videau, Christopher J. Williams, and Paul D. Adams. Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology*, 75:861–877, 2019.

[311] Xiang He, Alasdair F. Bell, and Peter J. Tonge. Ground state isomerization of a model green fluorescent protein chromophore. *FEBS Letters*, 549(1-3):35–38, 2003.

[312] Hans Frauenfelder, Hermann Hartmann, Fritz Parak, Martin Karplus, John Kuriyan, I. D. Kuntz,

Robert F. Tilton, Gregory A. Petsko, Dagmar Ringe, Michael L. Connolly, and Nelson Max. Thermal Expansion of a Protein. *Biochemistry*, 26(1):254–261, 1987.

[313] Marin van Heel and Michael Schatz. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3):250–262, sep 2005.

[314] Charles S. Bury, Jonathan C. Brooks-Bartlett, Steven P. Walsh, and Elspeth F. Garman. Estimate your dose: RADDOSE-3D. *Protein Science*, 27(1):217–228, jan 2018.

[315] J Hubbell and S Seltzer. X-Ray Mass Attenuation Coefficients. In *NIST Standard Reference Database 126*, chapter 2. 2019.

[316] Marius Schmidt, Vukica Srajer, Robert Henning, Hyotcherl Ihee, Namrta Purwar, Jason Tenboer, and Shailesh Tripathi. Protein energy landscapes determined by five-dimensional crystallography. *Acta Crystallographica Section D Biological Crystallography*, 69(12):2534–2542, dec 2013.

[317] Diamond Light. A 10-Year Vision for Diamond Light Source. Technical Report October, 2015.

[318] MAX-LAB. MAX IV Detailed Design Report on the MAX IV Facility. Technical report, 2010.

[319] A. P. Esposito, P. Schellenberg, W. W. Parson, and P. J. Reid. Vibrational spectroscopy and mode assignments for an analog of the green fluorescent protein chromophore. *Journal of Molecular Structure*, 569(1-3):25–41, 2001.

[320] Xiang He, Alasdair F. Bell, and Peter J. Tonge. Isotopic labeling and normal-mode analysis of a model green fluorescent protein chromophore. *Journal of Physical Chemistry B*, 106(23):6056–6066, 2002.

[321] Jasper J. Van Thor, Georgi Y. Georgiev, Michael Towrie, and J. Timothy Sage. Ultrafast and low barrier motions in the photoreactions of the green fluorescent protein. *Journal of Biological Chemistry*, 280(39):33652–33659, 2005.

[322] Martin Andresen, Andre C. Stiel, Simon Trowitzsch, Gert Weber, Christian Eggeling, Markus C. Wahl, Stefan W. Hell, Stefan Jakobs, Andre C. Stiel, Markus C. Wahl, Gert Weber, Stefan W. Hell, Simon Trowitzsch, Martin Andresen, Andre C. Stiel, Simon Trowitzsch, Gert Weber, Christian Eggeling, Markus C. Wahl, Stefan W. Hell, Stefan Jakobs, Andre C. Stiel, Markus C. Wahl, Gert Weber, Stefan W. Hell, Simon Trowitzsch, Martin Andresen, Andre C. Stiel, Simon Trowitzsch, Gert Weber, Christian Eggeling, Markus C. Wahl, Stefan W. Hell, and Stefan Jakobs. Structural basis for reversible photoswitching in Dronpa. *Proceedings of the National Academy of Sciences*, 104(32):13005–13009, aug 2007.

[323] Jasper J. Van Thor, Antonio J. Pierik, Inge Nugteren-Roodzant, Aihua Xie, and Klaas J. Hellingwerf. Characterization of the photoconversion of green fluorescent protein with FTIR spectroscopy. *Biochemistry*, 37(48):16915–16921, 1998.

[324] Jasper J. van Thor. Photoreactions and dynamics of the green fluorescent protein. *Chemical Society Reviews*, 38(10):2935–2950, 2009.

[325] Andreas Barth and Christian Zscherp. What vibrations tell us about proteins. *Quarterly Reviews of Biophysics*, 35(4):369–430, 2002.

[326] L. H. Andersen, H. Bluhme, S. Boyé, T. J. D. Jørgensen, H. Krogh, I. B. Nielsen, S. Brøndsted Nielsen, and A. Svendsen. Experimental studies of the photophysics of gas-phase fluorescent protein chromophores. *Phys. Chem. Chem. Phys.*, 6(10):2617–2627, 2004.

[327] M. Kaucikas, J. Barber, and J. J. Van Thor. Polarization sensitive ultrafast mid-IR pump probe micro-spectrometer with diffraction limited spatial resolution. *Optics Express*, 21(7):8357, 2013.

[328] Peter Schellenberg, Ethan Johnson, Anthony P. Esposito, Philip J. Reid, and William W. Parson. Resonance raman scattering by the green fluorescent protein and an analogue of its chromophore. *Journal of Physical Chemistry B*, 105(22):5316–5322, 2001.

[329] Roman A. Laskowski and Mark B. Swindells. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling*, 51(10):2778–2786, 2011.

[330] A. Toniolo, S. Olsen, L. Manohar, and T. J. Martinez. Ultrafast excited state dynamics in the green fluorescent protein chromophore. *Femtochemistry and Femtobiology: Ultrafast Events in Molecular Science*, 31(0):425–432, 2004.

[331] P. Hamm, S. M. Ohline, and W. Zinth. Vibrational cooling after ultrafast photoisomerization of azobenzene measured by femtosecond infrared spectroscopy. *Journal of Chemical Physics*, 106(2):519–529, 1997.

[332] Alasdair F. Bell, Xiang He, Rebekka M. Wachter, and Peter J. Tonge. Probing the ground state structure of the green fluorescent protein chromophore using Raman spectroscopy. *Biochemistry*, 39(15):4423–4431, 2000.

[333] Naomi M. Webber and Stephen R. Meech. Electronic spectroscopy and solvatochromism in the chromophore of GFP and the Y66F mutant. *Photochemical and Photobiological Sciences*, 6(9):976–981, 2007.

[334] Nadya V. Pletneva, Vladimir Z. Pletnev, Irina I. Shemiakina, Dmitriy M. Chudakov, Igor Artemyev, Alexander Wlodawer, Zbigniew Dauter, and Sergei Pletnev. Crystallographic study of red fluorescent protein eqFP578 and its far-red variant Katushka reveals opposite ph-induced isomerization of chromophore. *Protein Science*, 20(7):1265–1274, 2011.

[335] Jan Petersen, Pascal G. Wilmann, Travis Beddoe, Aaron J. Oakley, Rodney J. Devenish, Mark Prescott, and Jamie Rossjohn. The 2.0-ÅCrystal Structure of eqFP611, a Far Red Fluorescent Protein from the Sea Anemone Entacmaea quadricolor. *Journal of Biological Chemistry*, 278(45):44626–44631, 2003.

[336] Konstantin L. Litvinenko, Naomi M. Webber, and Stephen R. Meech. Internal conversion in the chromophore of the green fluorescent protein: Temperature dependence and isoviscosity analysis. *Journal of Physical Chemistry A*, 107(15):2616–2623, 2003.

[337] Stephen R. Meech. Excited state reactions in fluorescent proteins. *Chemical Society Reviews*, 38(10):2922–2934, 2009.

[338] I. V. Polyakov, B. L. Grigorenko, E. M. Epifanovsky, A. I. Krylov, and A. V. Nemukhin. Potential energy landscape of the electronic states of the GFP chromophore in different protonation forms: Electronic transition energies and conical intersections. *Journal of Chemical Theory and Computation*, 6(8):2377–2387, 2010.

[339] Seth Olsen and Sean C. Smith. Bond selection in the photoisomerization reaction of anionic green fluorescent protein and kindling fluorescent protein chromophore models. *Journal of the American Chemical Society*, 130(27):8677–8689, 2008.

[340] Rinat Gepshtein, Dan Huppert, and Noam Agmon. Deactivation mechanism of the green fluorescent chromophore. *Journal of Physical Chemistry B*, 110(9):4434–4442, 2006.

[341] P. E. Ivashkin, K. A. Lukyanov, and I. V. Yampolsky. Synthesis of biosynthetic precursors of chromophores of red fluorescent proteins. *Russian Journal of Bioorganic Chemistry*, 37(4):411–420, 2011.

[342] Annette Svendsen, Hjalte V. Kiefer, Henrik B. Pedersen, Anastasia V. Bochenkova, and Lars H. Andersen. Origin of the Intrinsic Fluorescence of the Green Fluorescent Protein. *Journal of the American Chemical Society*, 139(25):8766–8771, 2017.

[343] Tanja Brakemann, Gert Weber, Martin Andresen, Gerrit Groenhof, Andre C. Stiel, Simon Trowitzsch, Christian Eggeling, Helmut Grubmüller, Stefan W. Hell, Markus C. Wahl, and Stefan Jakobs. Molecular basis of the light-driven switching of the photochromic fluorescent protein padron. *Journal of Biological Chemistry*, 285(19):14603–14609, 2010.

[344] Haruki Niwa, Satoshi Inouye, Takashi Hirano, Tatsuki Matsuno, Satoshi Kojima, Masayuki Kubota, Mamoru Ohashi, and Frederick I. Tsuji. Chemical nature of the light emitter of the Aequorea green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24):13617–13622, 1996.

[345] Asuka Higashino, Misao Mizuno, and Yasuhisa Mizutani. Chromophore Structure of Photochromic Fluorescent Protein Dronpa: Acid-Base Equilibrium of Two Cis Configurations. *Journal of Physical Chemistry B*, 120(13):3353–3359, 2016.

[346] William W. Ward, Hugh J. Prentice, Amy F. Roth, Chris W. Cody, and Sue C. Reeves. SPECTRAL PERTURBATIONS OF THE AEQUOREA GREEN-FLUORESCENT PROTEIN. *Photochemistry and Photobiology*, 35(6):803–808, jun 1982.

[347] Marc André Elsliger, Rebekka M. Wachter, George T. Hanson, Karen Kallio, and S. James Remington. Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry*, 38(17):5296–5301, 1999.

[348] Malea Kneen, Javier Farinas, Yuxin Li, and A. S. Verkman. Green fluorescent protein as a noninvasive intracellular pH indicator. *Biophysical Journal*, 74(3):1591–1599, 1998.

[349] Katjuša Brejc, Titia K. Sixma, Paul A. Kitts, Steven R. Kain, Roger Y. Tsien, Mats Ormö, and S. James Remington. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 94(6):2306–2311, 1997.

[350] John T M Kennis, Delmar S Larsen, Ivo H M van Stokkum, Mikas Vengris, Jasper J van Thor, and Rienk van Grondelle. Uncovering the hidden ground state of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 101(52):17988–17993, 2004.

[351] Stathys Papadimitriou, Socratis Loucaides, Victoire Rérolle, Eric P. Achterberg, Andrew G. Dickson, Matthew Mowlem, and Hilary Kennedy. The measurement of pH in saline and hypersaline media at sub-zero temperatures: Characterization of Tris buffers. *Marine Chemistry*, 184:11–20, 2016.

[352] Seth Olsen, Kristina Lamothe, and Todd J. Martínez. Protonic gating of excited-state twisting and charge localization in GFP chromophores: A mechanistic hypothesis for reversible photoswitching. *Journal of the American Chemical Society*, 132(4):1192–1193, 2010.

[353] Lars V. Schäfer, Gerrit Groenhof, Martial Boggio-Pasqua, Michael A. Robb, and Helmut Grubmüller. Chromophore protonation state controls photoswitching of the fluoroprotein asFP595. *PLoS Computational Biology*, 4(3), 2008.

[354] Bella L. Grigorenko, Igor V. Polyakov, Alexander P. Savitsky, and Alexander V. Nemukhin. Unusual emitting states of the kindling fluorescent protein: Appearance of the cationic chromophore in the GFP family. *Journal of Physical Chemistry B*, 117(24):7228–7234, 2013.

[355] Mats H.M. M Olsson, Chresten R. SØndergaard, Michal Rostkowski, and Jan H. Jensen. PROPKA3: Consistent treatment of internal and surface residues in empirical p K a predictions. *Journal of Chemical Theory and Computation*, 7(2):525–537, 2011.

A | Appendix

A.1 Development of RSFPs

	JvT75 Cis		JvT75 Trans	
	Overall	High Res.	Overall	High Res.
High resolution limit (Å)	1.24	1.24	1.26	1.24
Low resolution limit (Å)	53.85	1.26	39.15	1.26
Completeness	100	99.1	94.6	88.5
Multiplicity	6.2	5.7	3.3	2.9
I/sigma	9.4	1.7	8.6	1.8
Rmerge(I)	0.115	1.715	0.069	0.903
Rmeas(I)	0.126	1.89	0.08	1.076
Rpim(I)	0.05	0.785	0.04	0.574
$CC_{1/2}$	0.998	0.524	0.998	0.439
Wilson B factor (Å ^{2})	7.429		9.07	
Total observations	401142	18044	201225	8280
Total unique	65034	3183	61414	2831
Refinement				
R_{Cryst}	0.178		0.171	
R_{Free}	0.2		0.192	
R.M.S Bonds	0.015		0.0157	
R.M.S Angles	1.9		1.852	

Table A.1: Merging and refinement statistics of JvT75 in the cis and trans states.

A.2 Cryo-Kinetics and Crystallography

Condition:	**0**	**1**	**3**	**4**
Autoproc + STARANISO Processing				
Cell (Å)	38.56 73.46 78.03	38.74 73.49 78.12	38.74 73.63 78.02	38.64 73.74 78.14
Low res. lim. (Å)	53.492	53.53	53.551	53.633
High res. lit (Å)	1.159	0.939	1.29	1.459
Rmerge (all I+ & I-)	0.122	0.067	0.089	0.135
Rmeas (all I+ & I-)	0.127	0.073	0.093	0.14
Rpim (all I+ & I-)	0.035	0.03	0.026	0.039
Total observations	791089	769150	591566	452492
Total unique	59821	132999	46087	34513
Mean(I)/sd(I)	12.1	10.2	15.9	10.7
Completeness (ellip.)	94.8	94.6	91.3	94.1
Multiplicity	13.2	5.8	12.8	13.1
CC(1/2)	0.999	0.999	1	0.999
Anom. compl. (ellip.)	94.7	93	91	94
Anom. multi.	6.9	3	6.8	6.9
CC(ano)	-0.105	-0.316	-0.223	-0.0042
Refinement				
R-factor / R-free	0.11 / 0.16	0.13 / 0.15	0.16 / 0.19	0.16 / 0.19
Bonds	0.015	0.034	0.012	0.012
Angles	1.85	2.21	1.77	1.76
B-factors	Anisotropic	Anisotropic	Isotropic	Isotropic

Table A.2: Showing the datasets selected for comparison by making isomorphous difference maps. The top row refers to illumination conditions labelled in figure 3.4

Condition:	**3a**	**3aa**	**4a**
Autoproc + STARANISO Processing			
Cell (Å)	38.57 73.22 77.91	38.51 73.08 77.95	38.67 73.80 78.11
Low res. limit (Å)	53.36	53.314	53.645
High res. limit (Å)	1.209	1.393	1.336
Rmerge (all I+ & I-)	0.076	0.098	0.087
Rmeas (all I+ & I-)	0.079	0.102	0.091
Rpim (all I+ & I-)	0.022	0.028	0.026
Total observations	702323	427758	496420
Total unique	56025	32764	44451
Mean(I)/sd(I)	18.3	16	15.7
Completeness (ellip.)	94.4	94.5	90.6
Multiplicity	12.5	13.1	11.2
CC(1/2)	1	1	0.999
Anom. compl. (ellip.)	94.4	94.2	90.4
Anom. multi.	6.6	6.9	5.9
CC(ano)	-0.214	-0.115	-0.182
Refinement			
R-factor / R-free	0.16 / 0.19	0.16 / 0.20	0.17 / 0.19
Bonds	0.014	0.011	0.014
Angles	1.83	1.68	1.85
B-factors	Isomorphous	Isomorphous	Isomorphous

Table A.3: Showing the datasets selected for comparison by making isomorphous difference maps. The top row refers to illumination conditions labelled in figure 3.4 (a)

A.3 SFX of rsKiiro

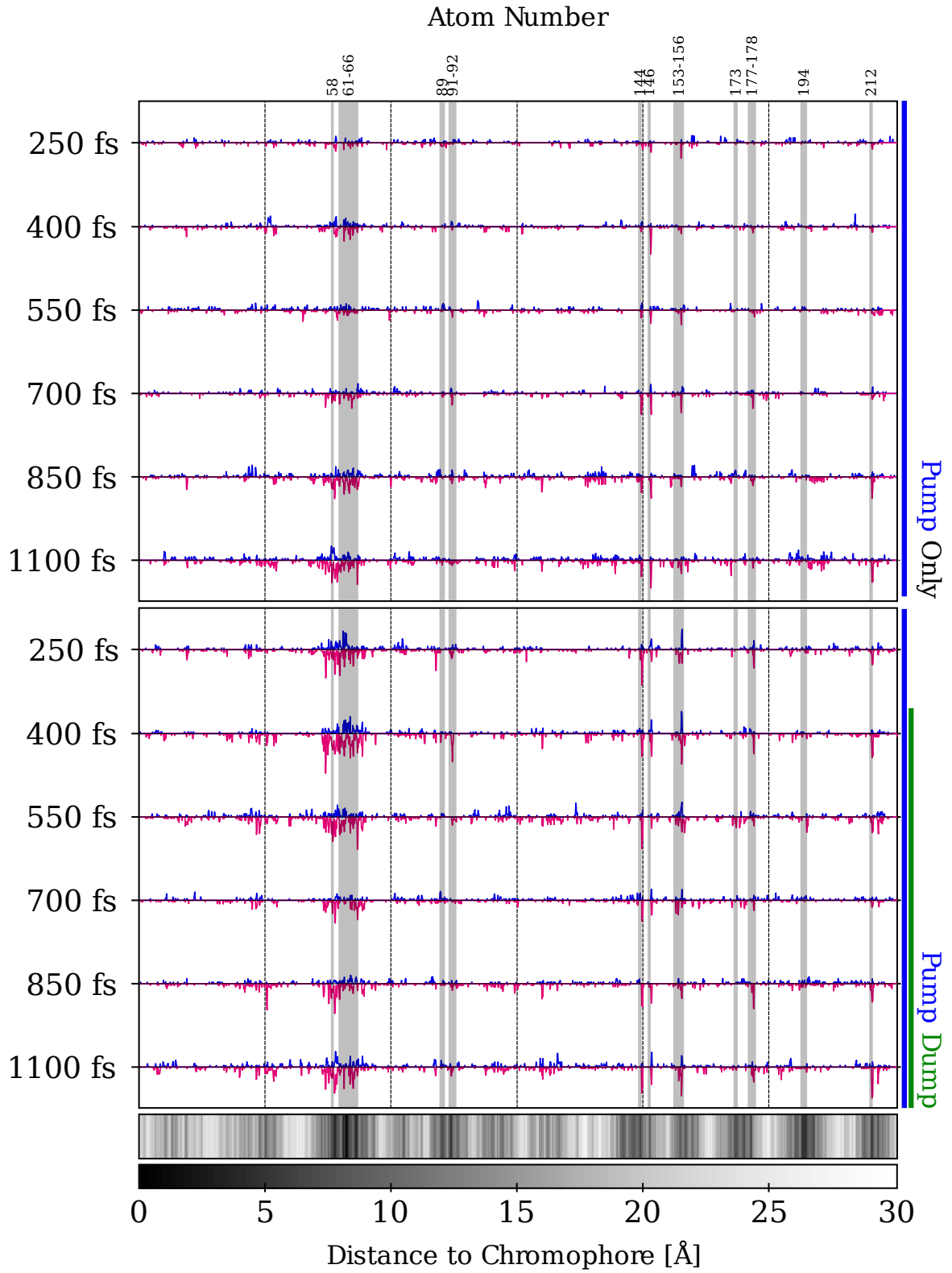


Figure A.1: Integrated electron density within 2 Å of each sequential atom in the protein chain for the Pump only data (top) and pump-dump data (bottom). Plotted for each time bin is the positive (blue) and negative (red) electron density. Shown at the bottom is a colour bar indicating the distance from the centre of the chromophore to the atom number. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

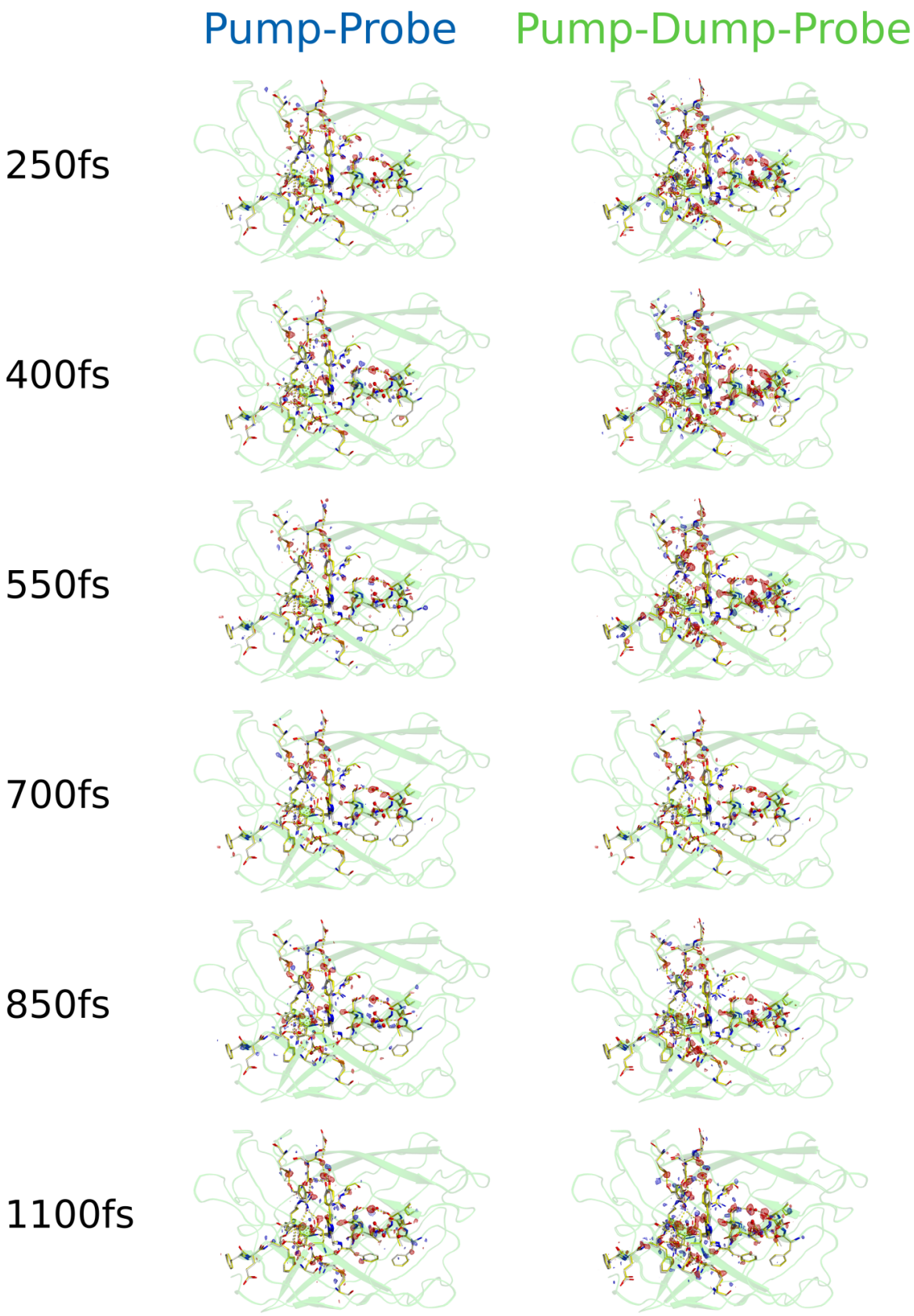


Figure A.2: Q-weighted difference electron density for each bin of the pump probe (400 nm) on the left and of the pump-dump probe (400-515 nm) on the right. The secondary structure (green) of the ground state coordinates (yellow) are shown with electron density contoured at +3σ rms (blue) and -3σ rms (red). Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

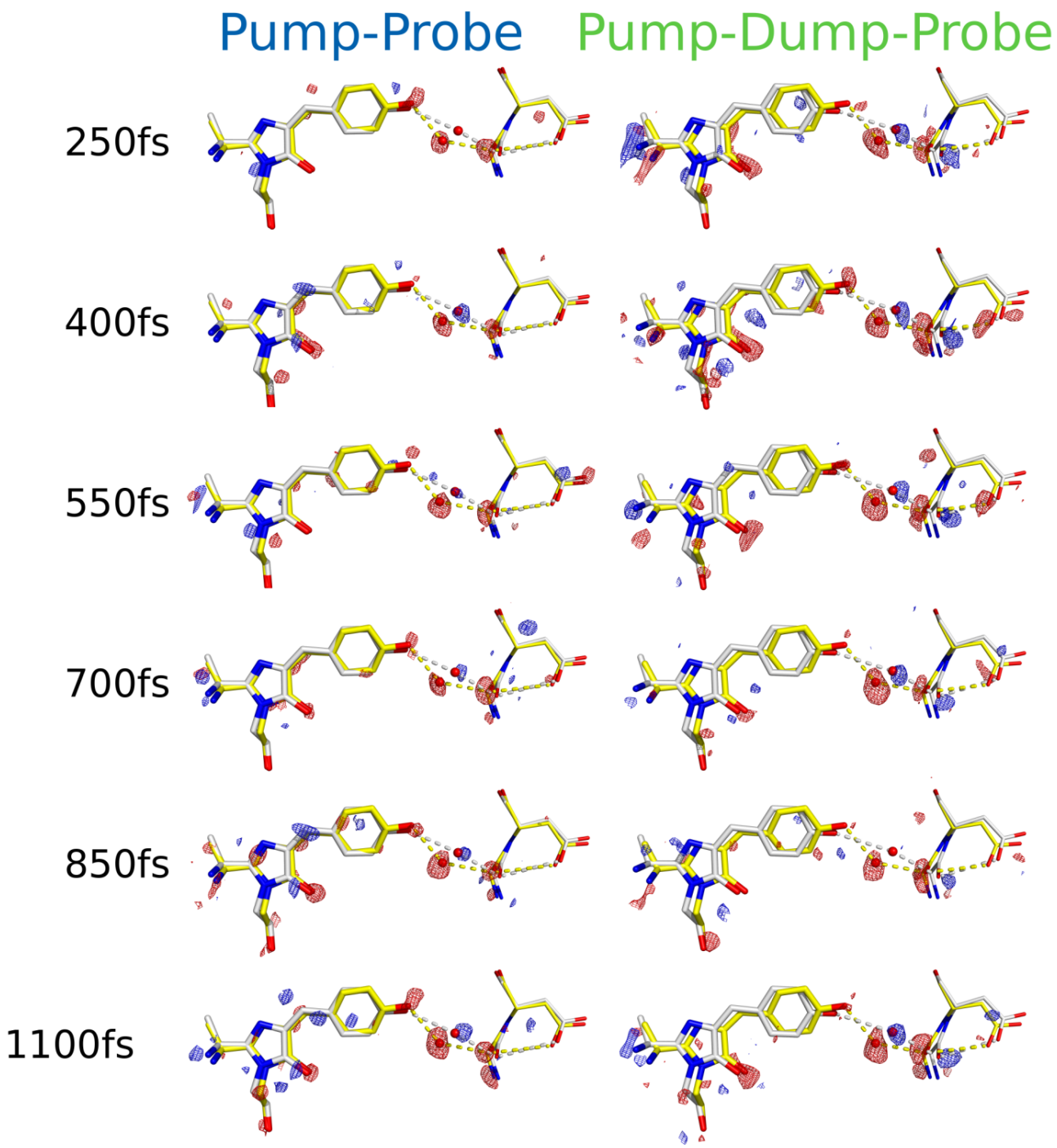
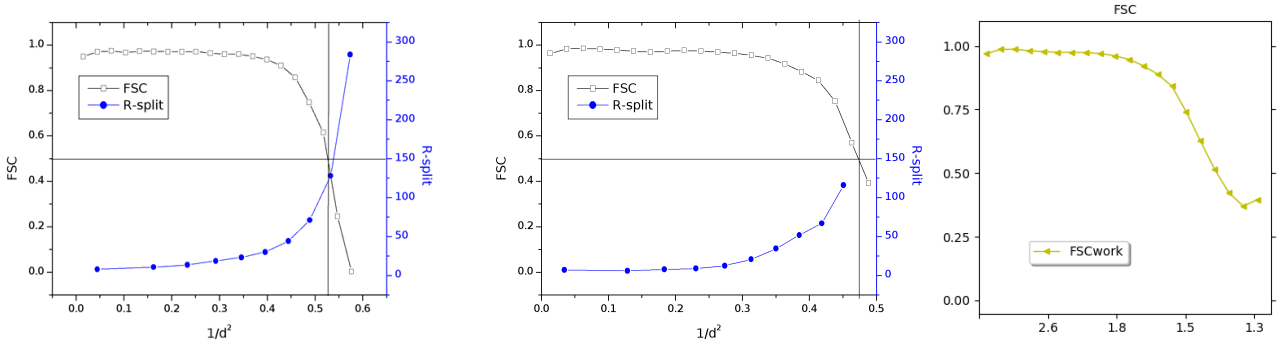


Figure A.3: Difference maps for the chromophore region, hydrogen bonded water and for the femtosecond time bins for PP (Left) and PDP (right) data. The coordinates for ground state (yellow carbons) and the refined coordinates for light data after partiality refinement (grey carbons) are shown with electron density contoured at +3σ rms (blue) and -3σ rms (red). Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

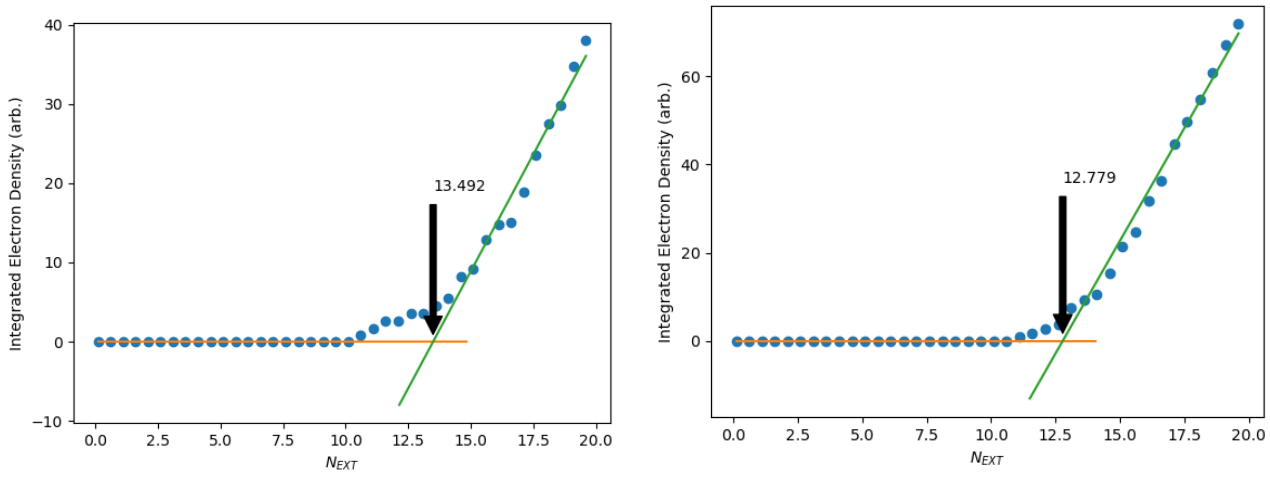
Name	Affiliation
Jasper Van Thor	Imperial College London
Christopher Hutchison	Imperial College London
Marc Morgan	Imperial College London
Rafal Rakowski	Imperial College London
Karim Maghlaoui	Imperial College London
Violeta Cordon Preciado	Imperial College London
Ben Taylor	Imperial College London
Flo Glencross	Imperial College London
Ann Fitzpatrick	Diamond Light Source
Anton Barty	CFEL
Gabriel Dorlhiac	Lawrence Berkeley National Laboratory
Gediminas Galinis	Imperial College London
Jonathan Barnard	Imperial College London
Douglas Garratt	Imperial College London
Josie Ferreira	Imperial College London
Ben Taylor	Imperial College London
Sian Stockton	Imperial College London
Letong Yuan	Imperial College London
Volha (Olga) Chukhutsina	Imperial College London

Table A.4: SFX Beamtime Participants



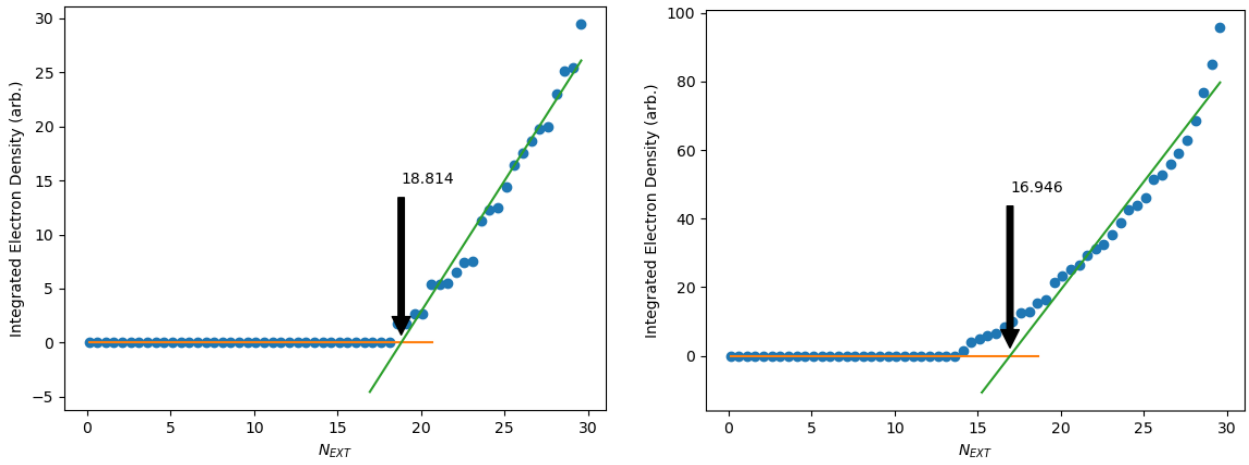
(a) LCLS **(b)** SACLA **(c)** PAL

Figure A.4: The Fourier shell coefficient (FSC) and R-Split for the all merged laser off, dark data. An FSC of 0.5 is chosen for a limiting 1.38, 1.44 and 1.41 Å for the LCLS, SACLA and PAL experiments respectively. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).



(a) Pump-probe **(b)** Pump-dump-probe

Figure A.5: LCLS: Pump-probe 0-1 ps (left) and pump-dump-probe 0-1 ps (right): Extrapolated factor, N_{EXT}, determination through the intersection point of two linear fits to the magnitude of integrated negative electron density, shown as blue circles. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).



(a) Pump-probe **(b)** Pump-dump-probe

Figure A.6: SACLA: Pump-probe 550 fs (left) and pump-dump-probe 550 fs (right): Extrapolated factor, N_{EXT}, determination through the intersection point of two linear fits to the magnitude of integrated negative electron density, shown as blue circles. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	Dark	Laser-On-400nm	Laser-On-400nm-515nm
Indexed Patterns:	35030	47590	39754
Resolution Limits Å:	15.40-1.80 (1.864-1.800)a	15.40-1.80 (1.864-1.800)a	18.01-1.80 (1.864-1.800)a
	15.40-1.50 (1.554-1.500)b	15.40-1.50 (1.554-1.500)b	18.01-1.50 (1.554-1.500)b
	15.40-1.35 (1.398-1.350)c	15.40-1.35 (1.398-1.350)c	18.01-1.35 (1.398-1.350)c
No. Unique reflection Indices:	21941a	21941a	21964a
	37542b	37542b	37565b
	51229c	51229c	51252c
No. Merged Reflections:	8859960 (583183)a	9663635 (454993)a	7042698 (265101)a
	12622674 (790070)b	11562142 (263229)b	7928580 (96156)b
	14971721 (647523)c	11990599 (64419)c	8056732 (13369)c
Completeness (%):	100.00 (100.00)a	100.00 (100.00)a	99.96 (100.00)a
	100.00 (100.00)b	100.00 (100.00)b	99.98 (100.00)b
	100.00 (100.00)c	99.77 (97.73)c	96.67 (68.35)c
Signal to noise:	8.568 (5.24)a	9.793 (5.65)a	8.272 (4.27)a
	6.328 (1.86)b	7.225 (2.36)b	5.790 (1.20)b
	4.843 (0.40)c	5.607 (0.61)c	4.566 (0.88)c
Wilson b factor:	27.23a	13.84a	16.96a
	32.89b	16.13b	19.19b
	35.68c	16.60c	19.87c
R_{Split} (%):	9.27 (18.03)a	8.76 (16.90)a	9.83 (22.54)a
	9.51 (79.05)b	9.83 (40.27)b	11.19 (82.62)b
	9.68 (291.14)c	11.01 (142.12)c	12.19 (350.70)c
CC*	1.00 (0.98)a	1.00 (0.99)a	1.00 (0.98)a
	1.00 (0.61)b	1.00 (0.94)b	1.00 (0.82)b
	1.00 (0.07)c	1.00 (0.62)c	1.00 (-nan)c
$CC_{1/2}$	0.99 (0.94)a	0.99 (0.94)a	0.99 (0.91)a
	0.99 (0.23)b	0.99 (0.79)b	0.99 (0.51)b
	0.99 (0.00)c	0.99 (0.24)c	0.99 (0.05)c
Refinement			
Resolution	1.3	1.5	1.5
No. reflections all/free	54155 / 1872	37523/1814	37510 / 1867
R-free / R-factor	0.180/0.154	0.162 / 0.155	0.182 / 0.178
Bonds	0.0288	0.0598	0.0589
Angles	2.767	2.854	3.014

Table A.5: TR-SSX data merging and refinement crystallography statistics from LCLS data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	400nm-150fs-300fs	400nm-300fs-450fs	400nm-450fs-600fs	400nm-515nm-150fs-300fs
Indexed Patterns:	8851	8937	7082	6965
Resolution Limits Å:	15.38-1.80 (1.864-1.800)a	15.38-1.80 (1.864-1.800)a	15.38-1.80 (1.864-1.800)a	15.40-1.80 (1.864-1.800)a
	15.38-1.50 (1.554-1.500)b	15.38-1.50 (1.554-1.500)b	15.38-1.50 (1.554-1.500)b	15.40-1.50 (1.554-1.500)b
	15.38-1.35 (1.398-1.350)c	15.38-1.35 (1.398-1.350)c	15.38-1.35 (1.398-1.350)c	15.40-1.38 (1.398-1.350)c
No. Unique reflection Indices:	21941a	21941a	21941a	21946a
	37542b	37542b	37542b	37547b
	51229c	51229c	51229c	51234c
No. Merged Reflections:	1805581 (85788)a	1854228 (87736)a	1418718 (66546)a	1232379 (46402)a
	2158555 (48658)b	2221023 (51247)b	1703429 (39980)b	1385663 (16018)b
	2238366 (11858)c	2303118 (11548)c	1767350 (8446)c	1405717 (1427)c
Completeness (%):	100.00 (100.00)a	99.99 (100.00)a	100.00 (100.00)a	99.97 (99.95)a
	99.97 (99.76)b	99.97 (99.76)b	99.91 (99.43)b	98.69 (90.42)b
	95.83 (64.77)c	95.72 (62.66)c	94.08 (52.76)c	84.07 (12.25)c
Signal to noise:	4.366 (2.59)a	4.417 (2.63)a	3.974 (2.43)a	3.641 (1.96)a
	3.261 (1.17)b	3.295 (1.14)b	2.999 (1.14)b	2.751 (2.26)b
	2.601 (-0.83)c	2.754 (0.69)c	2.486 (0.48)c	2.447 (-0.07)c
Wilson b factor:	13.63a	13.71a	13.23a	16.78a
	15.74b	15.92b	15.76b	18.31b
	16.31c	16.74c	16.53c	20.07c
R_{Split} (%):	20.62 (40.80)a	20.52 (39.65)a	23.38 (44.55)a	23.86 (53.95)a
	23.14 (91.06)b	23.02 (88.59)b	25.94 (95.65)b	26.15 (140.71)b
	24.29 (221.54)c	24.15 (217.17)c	26.93 (262.02)c	26.36 (307.19)c
CC*	0.98 (0.91)a	0.98 (0.91)a	0.98 (0.90)a	0.98 (0.88)a
	0.99 (0.74)b	0.98 (0.75)b	0.98 (0.72)b	0.98 (0.57)b
	0.99 (0.38)c	0.98 (0.48)c	0.98 (0.30)c	0.98 (0.67)c
$CC_{1/2}$	0.94 (0.70)a	0.93 (0.72)a	0.92 (0.67)a	0.92 (0.63)a
	0.94 (0.38)b	0.94 (0.40)b	0.92 (0.35)b	0.92 (0.19)b
	0.94 (0.08)c	0.94 (0.13)c	0.92 (0.05)c	0.93 (0.28)c

Table A.6: TR-SSX data merging crystallography statistics from LCLS data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	400nm-515nm-300fs-450fs	400nm-515nm-450fs-600fs	400nm-515nm-600fs-750fs	400nm-515nm-750fs-up
Indexed Patterns:	7445	5549	5013	9846
Resolution Limits Å:	15.40-1.80 (1.864-1.800)a	15.38-1.80 (1.864-1.800)a	15.40-1.80 (1.864-1.800)a	15.40-1.80 (1.864-1.800)a
	15.40-1.50 (1.554-1.500)b	15.38-1.50 (1.554-1.500)b	15.40-1.50 (1.554-1.500)b	15.40-1.50 (1.554-1.500)b
	15.40-1.38 (1.398-1.350)c	15.38-1.40 (1.398-1.350)c	15.40-1.39 (1.398-1.350)c	15.40-1.36 (1.398-1.350)c
No. Unique reflection Indices:	21957a	21958a	21946a	21962a
	37558b	37559b	37547b	37563b
	51245c	51246c	51234c	51250c
No. Merged Reflections:	1321858 (49215)a	983632 (37102)a	881928 (33374)a	1729036 (64348)a
	1485405 (17502)b	1107532 (12031)b	991780 (11593)b	1945978 (24248)b
	1505921 (1068)c	1119446 (430)c	1003575 (601)c	1975033 (1784)c
Completeness (%):	99.93 (99.95)a	99.88 (99.91)a	99.93 (99.81)a	99.91 (100.00)a
	98.96 (92.60)b	97.52 (81.91)b	97.22 (80.65)b	99.56 (96.93)b
	84.22 (9.41)c	79.44 (4.00)c	79.30 (5.47)c	86.82 (15.05)c
Signal to noise:	3.721 (2.01)a	3.317 (1.82)a	3.097 (1.64)a	4.190 (2.24)a
	2.681 (0.75)b	2.371 (0.77)b	2.505 (3.14)b	2.981 (0.74)b
	2.478 (1.92)c	2.216 (2.14)c	2.212 (0.72)c	2.587 (-0.48)c
Wilson b factor:	16.77a	16.44a	16.72a	16.97a
	19.26b	18.79b	18.66b	19.21b
	19.11c	18.48c	20.17c	20.29c
R_{Split} (%):	22.98 (52.12)a	26.59 (61.50)a	28.50 (66.46)a	20.46 (46.97)a
	25.21 (147.71)b	28.71 (164.21)b	30.61 (158.01)b	22.84 (155.02)b
	25.46 (263.64)c	28.80 (-nan)c	30.69 (244.85)c	23.24 (164.72)c
CC*	0.98 (0.87)a	0.97 (0.84)a	0.97 (0.82)a	0.98 (0.91)a
	0.98 (0.69)b	0.98 (0.49)b	0.97 (0.65)b	0.99 (0.57)b
	0.98 (0.32)c	0.98 (-nan)c	0.97 (-nan)c	0.99 (0.87)c
$CC_{1/2}$	0.93 (0.61)a	0.90 (0.54)a	0.89 (0.51)a	0.94 (0.70)a
	0.93 (0.31)b	0.91 (0.14)b	0.90 (0.27)b	0.95 (0.19)b
	0.93 (0.05)c	0.91 (-nan)c	0.90 (-nan)c	0.95 (0.60)c

Table A.7: TR-SSX data merging crystallography statistics from LCLS data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	400nm-515nm-min50fs-150fs	400nm-515nm-neg	400nm-600fs-750fs	400nm-750fs-up
Indexed Patterns: 13591	4936	284	6109	9672
Resolution Limits Å:	15.40-1.80 (1.864-1.800)a	15.11-1.81 (1.864-1.800)a	15.40-1.80 (1.864-1.800)a	15.38-1.80 (1.864-1.800)a
	15.40-1.50 (1.554-1.500)b	15.11-1.61 (1.554-1.500)b	15.40-1.50 (1.554-1.500)b	15.38-1.50 (1.554-1.500)b
	15.40-1.38 (1.398-1.350)c	15.11-1.61 (1.398-1.350)c	15.40-1.35 (1.398-1.350)c	15.38-1.35 (1.398-1.350)c
No. Unique reflection Indices:	21941a	21936a	21941a	21941a
	37542b	37537b	37542b	37542b
	51229c	51224c	51229c	51229c
No. Merged Reflections:	893288 (34506)a	43550 (1078)a	1230614 (57712)a	1943189 (89769)a
	1008790 (12766)b	46240 (129)b	1472507 (32758)b	2301688 (47550)b
	1024201 (1156)c	46327 (6)c	1523940 (6677)c	2372926 (8662)c
Completeness (%):	99.94 (99.86)a	53.18 (20.90)a	99.99 (100.00)a	100.00 (100.00)a
	97.63 (83.83)b	34.26 (1.69)b	99.86 (98.80)b	99.97 (99.81)b
	81.54 (10.20)c	25.19 (0.06)c	92.66 (44.74)c	94.50 (53.59)c
Signal to noise:	3.169 (1.78)a	2.943 (4.17)a	3.691 (2.22)a	4.559 (2.70)a
	2.480 (1.69)b	3.108 (0.65)b	2.761 (0.98)b	3.376 (1.10)b
	2.293 (0.17)c	3.102 (-3.33)c	2.297 (0.52)c	2.748 (0.39)c
Wilson b factor:	16.26a	15.17a	13.82a	14.54a
	18.60b	16.91b	16.18b	16.60b
	19.64c	11.82c	16.80c	17.41c
R_{Split} (%):	27.92 (60.83)a	75.15 (89.99)a	25.04 (46.72)a	19.56 (37.93)a
	30.07 (158.53)b	75.17 (-nan)b	27.87 (109.07)b	21.89 (93.43)b
	30.24 (163.15)c	75.17 (-nan)c	28.70 (251.11)c	22.95 (231.68)c
CC*	0.97 (0.85)a	0.67 (0.31)a	0.98 (0.89)a	0.99 (0.92)a
	0.97 (0.60)b	0.67 (-nan)b	0.98 (0.67)b	0.99 (0.76)b
	0.97 (-nan)c	0.67 (-nan)c	0.98 (0.46)c	0.99 (0.39)c
$CC_{1/2}$	0.90 (0.56)a	0.29 (0.05)a	0.91 (0.66)a	0.94 (0.74)a
	0.90 (0.22)b	0.29 (-nan)b	0.92 (0.29)b	0.95 (0.41)b
	0.90 (0.42)c	0.29 (-nan)c	0.92 (0.12)c	0.95 (0.08)c

Table A.8: TR-SSX data merging crystallography statistics from LCLS data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	400nm-min50fs-150fs	400nm-neg
Indexed Patterns:	6939	674
Resolution Limits Å:	15.38-1.80 (1.864-1.800)a	14.60-1.80 (1.864-1.800)a
	15.38-1.50 (1.554-1.500)b	14.60-1.50 (1.554-1.500)b
	15.38-1.35 (1.398-1.350)c	14.60-1.41 (1.398-1.350)c
No. Unique reflection Indices:	21941a	21940a
	37542b	37541b
	51229c	51228c
No. Merged Reflections:	1411466 (67442)a	139031 (6484)a
	1704989 (42954)b	164828 (3137)b
	1775287 (9868)c	168886 (365)c
Completeness (%):	99.99 (100.00)a	90.00 (77.23)a
	99.94 (99.65)b	74.89 (33.61)b
	94.77 (57.49)c	58.26 (3.46)c
Signal to noise:	3.903 (2.43)a	2.098 (0.80)a
	2.902 (1.14)b	2.221 (1.61)b
	2.696 (4.68)c	1.833 (0.48)c
Wilson b factor:	12.78a	10.30a
	15.25b	12.30b
	16.21c	13.32c
R_{Split} (%):	23.92 (45.92)a	62.60 (94.66)a
	26.70 (95.74)b	63.19 (114.60)b
	27.82 (259.45)c	63.21 (-nan)c
CC*	0.98 (0.86)a	0.80 (0.50)a
	0.98 (0.70)b	0.80 (0.77)b
	0.98 (0.23)c	0.80 (-nan)c
$CC_{1/2}$	0.91 (0.59)a	0.47 (0.14)a
	0.92 (0.33)b	0.48 (0.42)b
	0.92 (0.03)c	0.48 (-nan)c

Table A.9: TR-SSX data merging crystallography statistics from LCLS data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	dark	PDP-100ps	PDP-3ps	PDP
Indexed Patterns:	77581	25204	83187	54743
Resolution Limits Å:	18.84-1.70 (1.761-1.700)a	18.53-1.70 (1.761-1.700)a	18.84-1.70 (1.761-1.700)a	18.84-1.70 (1.761-1.700)a
	18.84-1.65 (1.709-1.650)b	18.53-1.65 (1.709-1.650)b	18.84-1.65 (1.709-1.650)b	18.84-1.65 (1.709-1.650)b
	18.84-1.46 (1.491-1.440)c	18.53-1.55 (1.491-1.440)c	18.84-1.46 (1.491-1.440)c	18.84-1.46 (1.491-1.440)c
No. Unique reflection Indices:	26206a	26206a	26209a	26206a
	28578b	28578b	28581b	28578b
	42704c	42704c	42707c	42704c
No. Merged Reflections:	9886523 (135932)a	3553008 (32192)a	10902871 (150894)a	7427749 (120500)a
	9957691 (88288)b	3567097 (17641)b	10982523 (98384)b	7492574 (80164)b
	10044009 (1397)c	3575821 (4)c	11079889 (1945)c	7576103 (1587)c
Completeness (%):	100.00 (100.00)a	99.90 (99.07)a	99.99 (100.00)a	100.00 (100.00)a
	99.99 (99.93)b	99.28 (93.24)b	99.99 (99.96)b	100.00 (100.00)b
	88.38 (13.59)c	72.85 (0.05)c	89.46 (17.89)c	88.49 (14.60)c
Signal to noise:	12.242 (4.19)a	7.900 (2.86)a	12.716 (4.26)a	10.637 (3.99)a
	11.500 (3.37)b	7.486 (2.62)b	11.936 (3.38)b	10.016 (3.23)b
	9.357 (4.59)c	7.648 (7.35)c	9.669 (4.27)c	8.305 (7.19)c
Wilson b factor:	22.08a	21.95a	22.31a	22.07a
	22.23b	22.08b	22.48b	22.21b
	21.18c	22.54c	21.73c	21.73c
R_{Split} (%):	7.45 (28.00)a	10.91 (43.80)a	7.09 (27.15)a	8.81 (28.97)a
	7.61 (36.70)b	11.08 (56.75)b	7.25 (35.41)b	8.97 (37.75)b
	8.06 (126.90)c	11.12 (-nan)c	7.71 (62.87)c	9.40 (90.12)c
CC*	1.00 (0.96)a	0.99 (0.90)a	1.00 (0.97)a	1.00 (0.96)a
	1.00 (0.94)b	0.99 (0.84)b	1.00 (0.95)b	1.00 (0.93)b
	1.00 (-nan)c	0.99 (-nan)c	1.00 (0.92)c	1.00 (0.79)c
$CC_{1/2}$	0.99 (0.87)a	0.98 (0.68)a	0.99 (0.88)a	0.98 (0.86)a
	0.99 (0.80)b	0.98 (0.55)b	0.99 (0.81)b	0.98 (0.76)b
	0.99 (0.05)c	0.98 (-nan)c	0.99 (0.74)c	0.98 (0.45)c

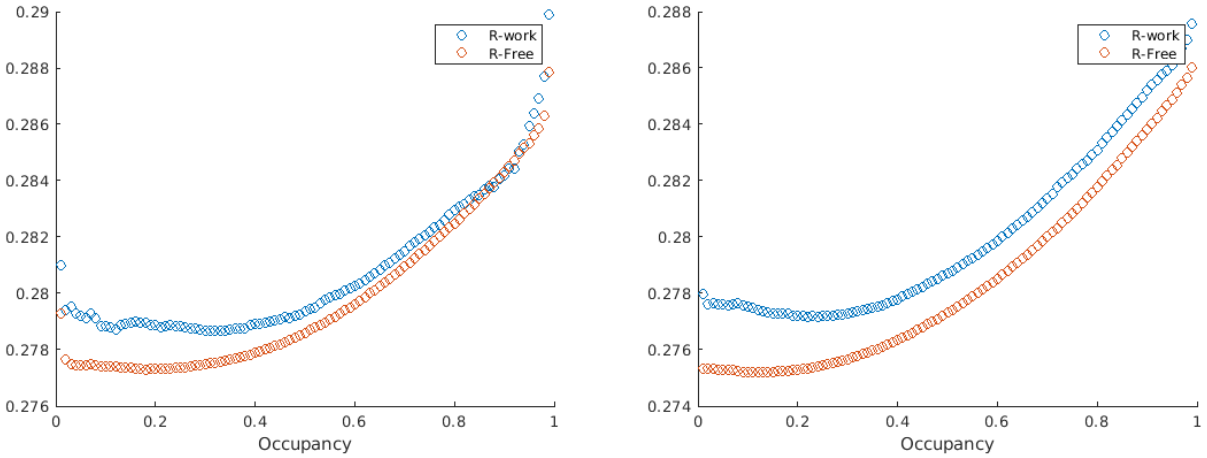
Table A.10: TR-SSX data merging crystallography statistics from SACLA data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	PP	PP-100ps	PP-3ps	PP-min5ps
Indexed Patterns:	18886	18886	20472	17627
Resolution Limits Å:	18.53-1.70 (1.761-1.700)a	18.53-1.70 (1.761-1.700)a	18.53-1.70 (1.761-1.700)a	18.53-1.70 (1.761-1.700)a
	18.53-1.65 (1.709-1.650)b	18.53-1.65 (1.709-1.650)b	18.53-1.65 (1.709-1.650)b	18.53-1.65 (1.709-1.650)b
	18.53-1.50 (1.491-1.440)c	18.53-1.53 (1.491-1.440)c	18.53-1.53 (1.491-1.440)c	18.53-1.53 (1.491-1.440)c
No. Unique reflection Indices:	26206a	26205a	26205a	26206a
	28578b	28577b	28577b	28578b
	42704c	42703c	42703c	42704c
No. Merged Reflections:	4822444 (50790)a	2258197 (18609)a	2327708 (29358)a	1841923 (19475)a
	4847105 (30746)b	2266705 (10743)b	2342594 (18535)b	1850853 (11375)b
	4872185 (182)c	2272675 (25)c	2356435 (47)c	1857695 (18)c
Completeness (%):	99.93 (99.42)a	99.37 (94.80)a	99.86 (98.80)a	99.41 (95.38)a
	99.75 (97.92)b	97.78 (81.99)b	99.33 (94.18)b	97.90 (82.93)b
	79.58 (2.09)c	70.45 (0.29)c	75.54 (0.55)c	71.09 (0.22)c
Signal to noise:	8.796 (3.18)a	6.574 (2.75)a	6.303 (2.59)a	6.017 (3.88)a
	8.295 (2.67)b	6.547 (5.65)b	6.013 (2.57)b	5.823 (4.10)b
	7.482 (1.69)c	6.453 (8.15)c	5.711 (2.14)c	5.836 (0.74)c
Wilson b factor:	22.28a	21.80a	22.25a	22.40a
	22.42b	21.93b	22.31b	22.39b
	21.50c	16.17c	21.84c	21.28c
R_{Split} (%):	10.98 (39.35)a	14.57 (50.17)a	15.12 (48.15)a	16.92 (52.66)a
	11.16 (51.44)b	14.63 (49.95)b	15.26 (49.46)b	16.98 (53.67)b
	11.35 (-nan)c	14.65 (-nan)c	15.32 (-nan)c	16.99 (-nan)c
CC*	0.99 (0.92)a	0.99 (0.89)a	0.98 (0.90)a	0.98 (0.85)a
	0.99 (0.88)b	0.99 (0.88)b	0.98 (0.89)b	0.98 (0.88)b
	0.99 (-nan)c	0.99 (-nan)c	0.98 (-nan)c	0.98 (-nan)c
$CC_{1/2}$	0.97 (0.74)a	0.95 (0.65)a	0.94 (0.68)a	0.92 (0.56)a
	0.97 (0.64)b	0.95 (0.63)b	0.94 (0.67)b	0.92 (0.63)b
	0.97 (-nan)c	0.95 (-nan)c	0.94 (-nan)c	0.92 (-nan)c

Table A.11: TR-SSX data merging crystallography statistics from SACLA data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).

Name:	100us-laser-on	1ps-laser-on	neg-1ps-laser-on
Indexed Patterns:	60993	51310	49465
Resolution Limits Å:	78.69-1.55 (1.605-1.550)a	78.69-1.55 (1.605-1.550)a	78.69-1.55 (1.605-1.550)a
	78.69-1.45 (1.502-1.450)b	78.69-1.45 (1.502-1.450)b	78.69-1.45 (1.502-1.450)b
	78.69-1.35 (1.398-1.350)c	78.69-1.35 (1.398-1.350)c	78.69-1.35 (1.398-1.350)c
No. Unique reflection Indices:	34233a	34233a	34233a
	41638b	41638b	41638b
	51398c	51398c	51398c
No. Merged Reflections:	14106334 (252617)a	11897106 (217770)a	11752976 (247091)a
	14357015 (91162)b	12107818 (75781)b	12005396 (92501)b
	14403050 (10532)c	12149512 (10075)c	12060432 (14662)c
Completeness (%):	100.00 (100.00)a	100.00 (100.00)a	100.00 (100.00)a
	99.98 (99.78)b	99.95 (99.49)b	99.96 (99.59)b
	95.02 (56.33)c	95.09 (57.28)c	95.90 (64.79)c
Signal to noise:	7.601 (1.81)a	6.644 (1.89)a	6.648 (1.95)a
	6.434 (0.76)b	5.682 (0.99)b	5.685 (0.97)b
	5.537 (0.18)c	4.964 (0.87)c	4.905 (0.56)c
Wilson b factor:	25.62a	22.49a	23.39a
	26.28b	22.59b	23.52b
	26.54c	22.10c	22.87c
R_{Split} (%):	8.64 (56.11)a	10.58 (56.19)a	10.18 (53.79)a
	8.96 (146.34)b	11.03 (118.58)b	10.58 (114.57)b
	9.16 (419.81)c	11.31 (237.53)c	10.85 (236.85)c
CC*	1.00 (0.91)a	1.00 (0.90)a	1.00 (0.90)a
	1.00 (0.63)b	1.00 (0.68)b	1.00 (0.69)b
	1.00 (0.06)c	1.00 (0.50)c	1.00 (0.55)c
$CC_{1/2}$	0.99 (0.69)a	0.98 (0.68)a	0.99 (0.68)a
	0.99 (0.25)b	0.98 (0.30)b	0.99 (0.31)b
	0.99 (0.00)c	0.99 (0.15)c	0.99 (0.18)c

Table A.12: TR-SSX data merging crystallography statistics from PAL data. The letters a, b and c refer to three separate resolution limits not the crystallographic axes. Table and caption adapted with permission from Hutchison *et al.* (unpublished).



(a) 100 µs **(b)** negative 1 ps

Figure A.7: R_{Work}/R_{Free} calculation for varied occupancy of refined coordinates. Occupancy refers to the proportion of ground state (trans) coordinates (dark) compared to the product (cis) coordinates. R_{Work} (and R_{Free}) minimise at 18% (21%) and 12% (16%) for the cis product in the 100 µs and negative 1 ps delays respectively. Figure and caption adapted with permission from Hutchison *et al.* (unpublished).

A.4 Quantitative Serial Synchrotron Crystallography of rsKiiro

A.4.1 Steady-State Quantum Yield Measurements of rsKiiro

The yields of both the off to on and on to off reaction were measured using solution of rs-Kiiro. Using the flash photolysis setup in section 5.2.1. The pump was used as a probe simultaneously. Data were fitted using python.

Trans to Cis: 405 nm laser was used a a pump / probe focused to 381x302 µm and power of 64.9 µW

$$\text{Trans} \xrightarrow{k_{Switch}, k_{TH}} \text{Cis}$$

The switching rate constant depends on the power density per an area (P), the proportion of absorbed photons in a sample of absorbance, A, the quantum yield, ϕ and the thermal rate, k_{TH}:

	Mean OD at 500 nm	Yield
Trans to Cis	0.15	5.28± 0.4 %
Cis to Trans	0.14	35.73 ±2.6 %

Table A.13: Measured steady state quantum yields of the Trans to cis and cis to trans reactions in rsKiiro.

$$-\frac{d[Trans]}{dt} = \frac{P}{l}(1 - 10^{-A})\phi + k_{TH}[Trans]$$

For the case where OD « 1 a Taylor expansion can be used to approximate the absorbed fraction as:

$$1 - 10^{-A} \approx Aln(10)$$

Using this approximation and the Beer-Lambert law, the switching rate can be rewritten as:

$$-\frac{d[Trans]}{dt} = P\epsilon_{Trans}\phi ln10[Trans] + k_{TH}[Trans]$$

The rate of thermal recovery of Trans state during the flash-photolysis measurement is much smaller than the rate of photo-induced switching so can be set to zero.

$$-\frac{d[Trans]}{dt} = P\epsilon_{Trans}\phi ln10[Trans]$$

Giving an integrated rate equation of:

$$[Trans] = [Trans]_i exp(-P\epsilon_{Trans}\phi ln10t)$$

Assuming the total concentration of Cis and Trans is constant the fraction of Cis population can be calculated as:

$$N_{Cis} = 1 - [Trans]_i exp(-\epsilon_{405}\phi ln10E)$$

Cis to Trans: 488 nm laser was focused to 48x54 μm with a power of 7.33 μW

$$N_{Trans} = 1 - [Cis]_i exp(-\epsilon_{488}\phi ln10E)$$

A.4.2 rsKiiro Spectra, Absorption Coefficient and Preconversion

The spectra of rsKiiro and Skylan-Ns can be seen in figure A.8. The published maximum absorption coefficient for Skylan-NS is 133,770 $M^{-1}cm^{-1}$ at 499 nm. rsKiiro absorbs most strongly at 485 nm with a calculated absorption coefficient of 98,727 $M^{-1}cm^{-1}$. At the flash wavelength of 405 nm the absorption coefficient is calculated as 31,184 $M^{-1}cm^{-1}$. Steady-state measurements at a power density of 0.0315 W/cm^2 (at 489 nm) gave a switching rate of 0.32 s^{-1} in rsKiiro. The estimated power density of the 473 nm laser on the chip is around 0.30 W/cm^2 scaling the rate constant to around 2.3 s^{-1} when accounting for change in absorption between 473 nm and 488 nm (a ratio of 0.82:1). Over the one second illumination period, this should preconvert at least 85% of the protein to the trans-state. The average waiting time for a crystal once preconverted was 45 s during which time an estimated 4.3% of the population will thermally recover back to the cis on average. This means the minimum expected cis population before x-ray probing of the crystal will be around 18.6%.

Temperature Increases Tables

A.4.3 Knife Edge

Knife edge scans were performed for experiments at the synchrotron and in the flash photolysis:

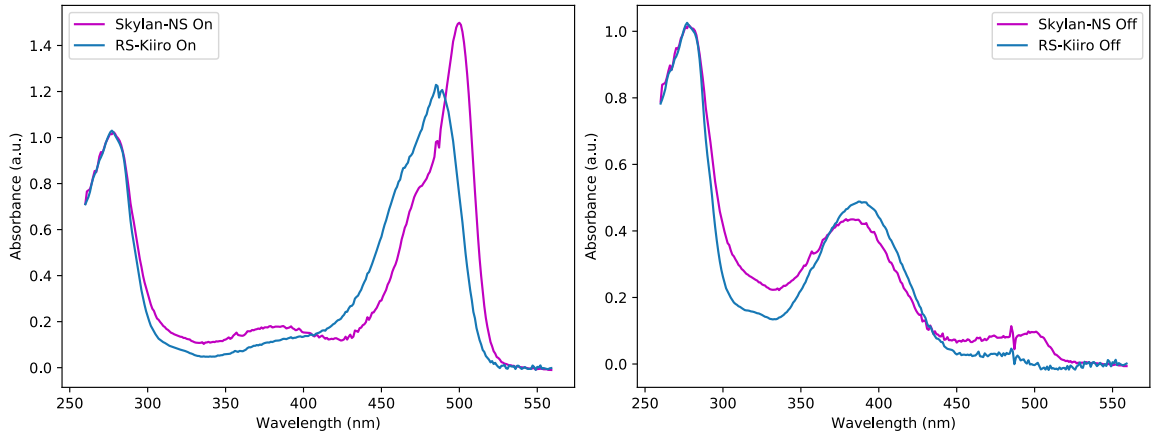
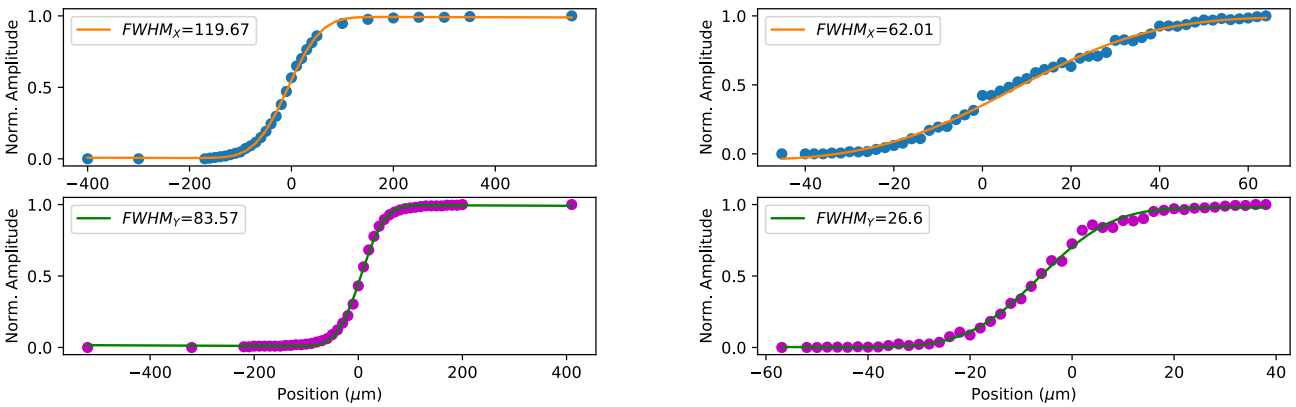


Figure A.8: The fluorescent, 'on' state (left) and 'off' state (right) spectra of Skylan-NS and rsKiiro. Spectra were normalised using the aromatic peak at 280 nm.

Method	k (mm^2/mJ)	ΔT (K)	δT
R_{Free}	0.562	112.19	54.83 – 180.96
R_{Work}	0.121	82.73	31.28 – 143.64
F_{Calc}-F_{obs}	0.235	87.26	34.92 – 149.36
Extrapolated	0.218	93.69	40.07 – 157.46
Flash Photolysis	0.274	N/A	N/A

Table A.14: Full results of x-ray crystallographic photolysis modelling and associated 95% confidence intervals

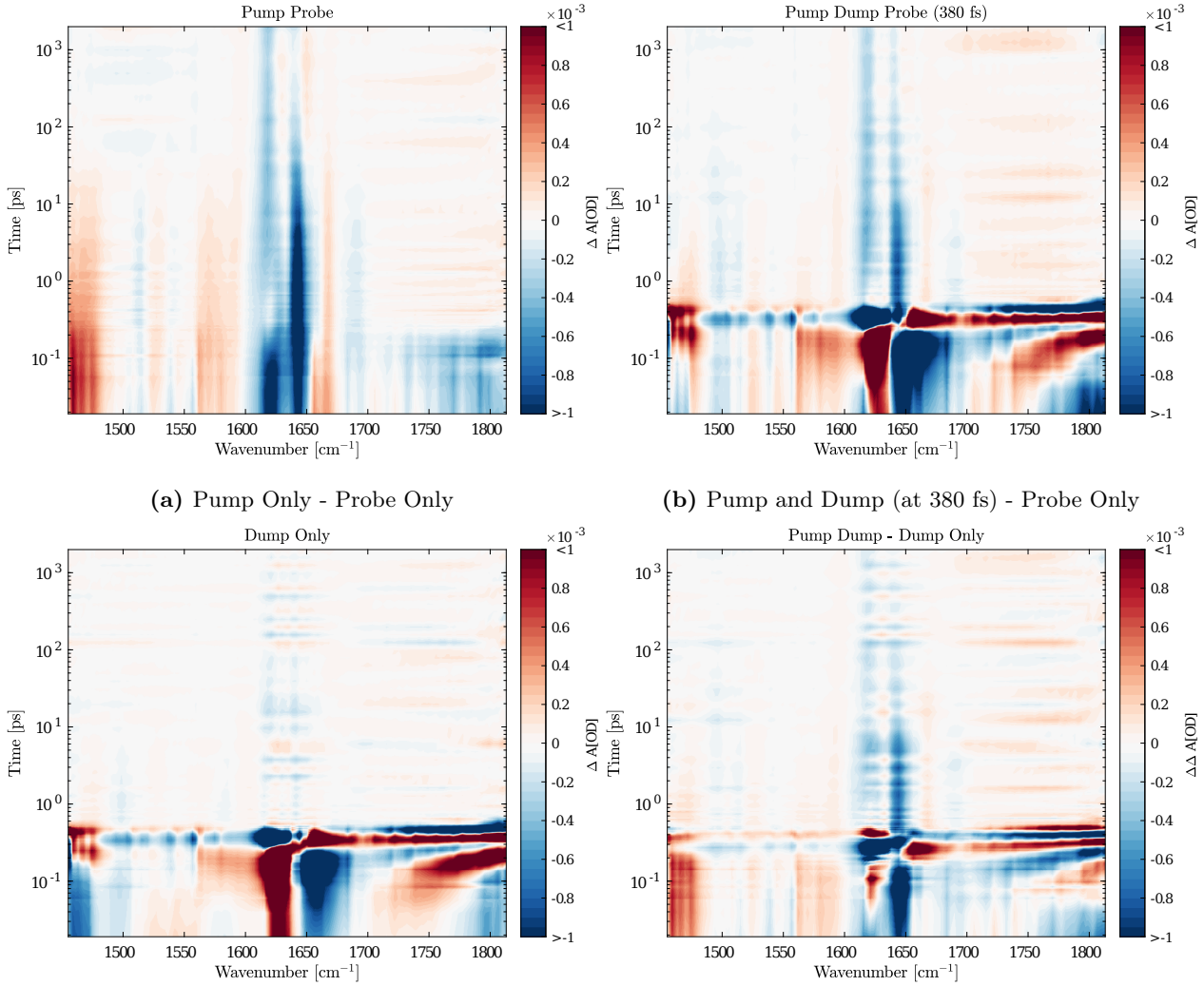


(a) Flash photolysis

(b) TR-SSX experiment

Figure A.9: Knife edge scans to determine the focus size during the flash photolysis experiments and SSX. Plotted is 'knife' position against amplitude of signal (normalised power measurement).

A.5 Ultrafast IRTA with Optical Control Reveals rsKiiro Isomerisation Mechanism



(a) Pump Only - Probe Only

(b) Pump and Dump (at 380 fs) - Probe Only

(c) Dump (at 380 fs) Only - Probe Only

(d) (Pump and Dump (380 fs) - Probe Only) - (Dump (at 380 fs) Only - Probe Only)

Figure A.10: IRTA difference spectrograms from 0 - 2 ns on a logarithmic scale for the pump probe, pump dump probe, dump only and the double-difference maps of pump dump - dump only.

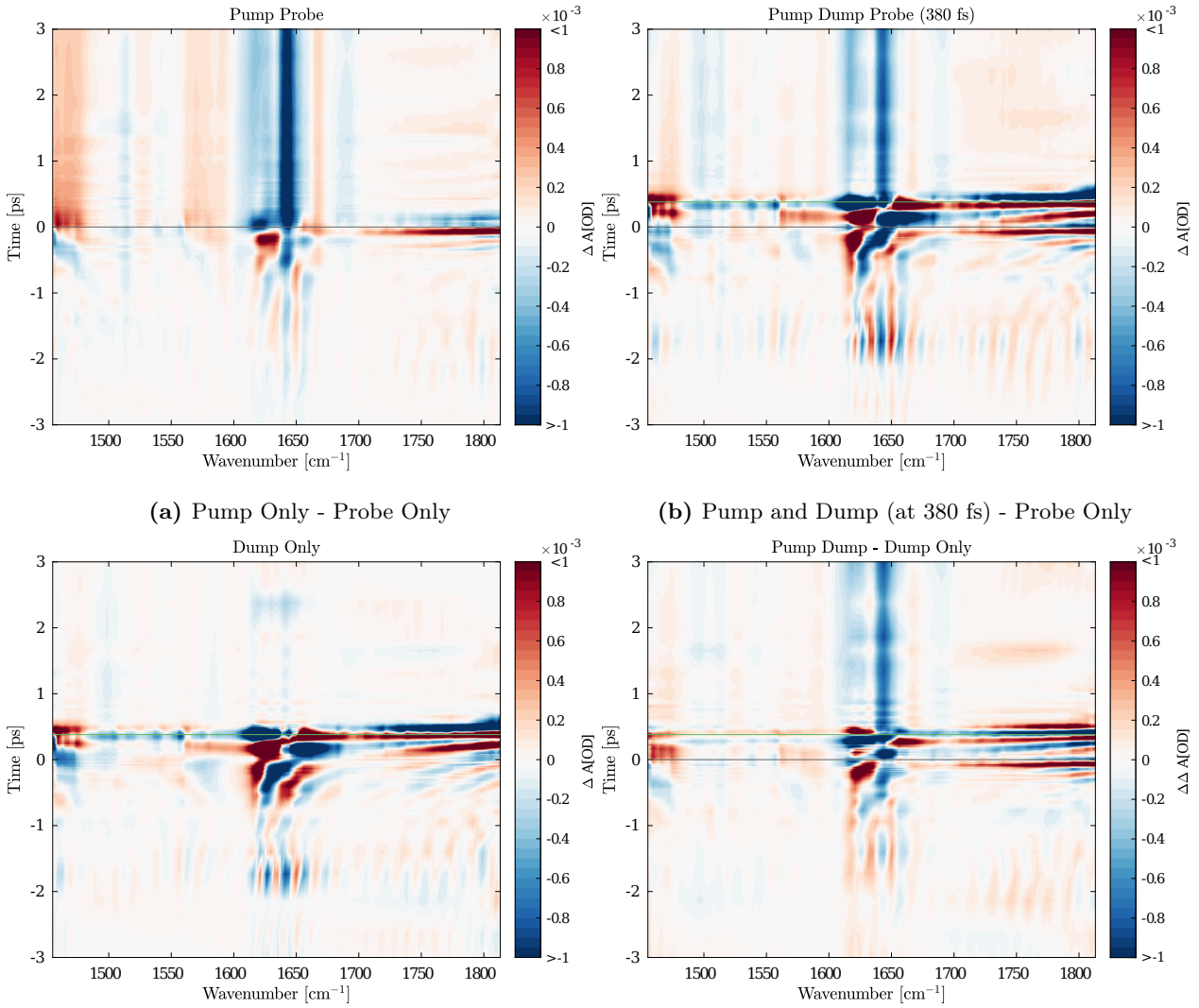
Dominant Component	cis-N ^1H_2O	cis-N-^2H_2O	cis-A	cis-Z-^1H_2O
C=O	1767 (351)	1767 (351)	1697 (366)	1758 (696)
C=C	1687 (300)	1687 (300)	1656 (37)	1677 (295)
phe1	1644 (473)	1643 (512)	1623 (1277)	1643 (18)
phe2	1618 (47)	1613 (22)		1609 (1903)
C=N/C=C	1601 (59)	1596 (46)	1595 (8)	1585 (112)
phe	1541 (127)	1540 (136)	1570 (21200)	1567 (85)

Dominant Component	cis-Z-^2H_2O	trans-A	trans-N-^1H_2O	trans-N-^2H_2O
C=O	1758 (696)	1690 (358)	1738 (181)	1738 (182)
C=C	1676 (331)	1654 (127)	1679 (252)	1679 (252)
phe1	1642 (52)	1628 (699)	1646 (374)	1645 (389)
phe2	1604 (1925)	1593 (113)	1615 (46)	1609 (45)
C=N/C=C	1575 (108)	1551 (1603)	1597 (37)	1592 (28)
phe	1568 (239)	1529 (134)	1543 (106)	1541 (129)

Dominant Component	trans-C-^1H_2O	trans-C-^2H_2O	(trans-N-^1H_2O)R	(trans-N-^2H_2O)R
C=O	1792 (207)	1792 (208)	1752 (178)	1752 (178)
C=C	1663 (75)	1662 (70)	1684 (259)	1684 (257)
phe1	1620 (644)	1613 (1317)	1647 (284)	1646 (304)
phe2	1601 (894)	1587 (138)	1619 (36)	1614 (19)
C=N/C=C	1590 (92)	1587 (138)	1608 (0.4)	1603 (0.7)
phe	1544 (100)	1542 (146)	1542 (96)	1540 (116)

Dominant Component	(trans-C-^1H_2O)R	(trans-C-^2H_2O)R
C=O	1806 (186)	1806 (187)
C=C	1663 (67)	1661 (187)
phe1	1618 (1010)	1615 (1388)
phe2	1604 (474)	1589 (42)
C=N/C=C	1596 (11)	1582 (59)
phe	1543 (83)	1541 (118)

Table A.15: Harmonic frequency calculations at the B3LYP/6-311bg(d,p) level, for the HBDI chromophore from Warren *et al.* [5].



(a) Pump Only - Probe Only

(b) Pump and Dump (at 380 fs) - Probe Only

(c) Dump (at 380 fs) Only - Probe Only

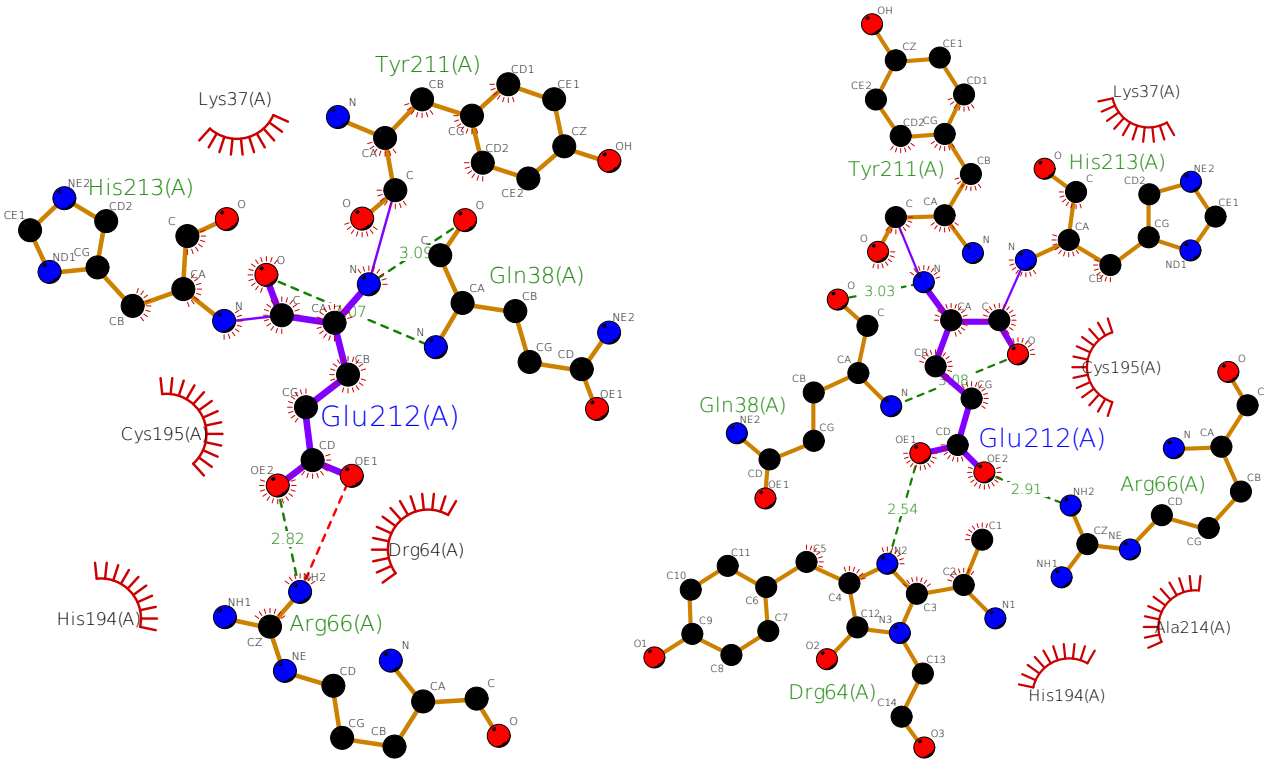
(d) (Pump and Dump (380 fs) - Probe Only) - (Dump (at 380 fs) Only - Probe Only)

Figure A.11: IRTA difference spectrograms from -3 - 3 ps on a linear scale for the pump probe, pump dump probe, dump only and the double-difference maps of pump dump - dump only. The black line shows the arrival of the pump pulse and the green line shows arrival of 515 nm dump pulse.

A.6 Observation of a Neutral to Cation Switching of a GFP-like Protein

Energy Density (mJ/mm^2)	0	0.10	0.53	1.78	6.74	14.43
Resolution (Å)	54.17-1.70					
Space Group	$P2_12_12_1$					
Unit Cell (Å)	$a = 39.6, b = 74.5, c = 78.9, \alpha = 90, \beta = 90, \gamma = 90$					
No. of Reflections	26168	26186	26057	26298	26371	26144
No. of Crystals	12649	10555	14166	18204	13585	10817
Completeness (%)	99.85	99.72	99.65	99.94	99.96	99.81
Redundancy	66.9	50	53.9	93.5	95.2	62.8
I/ σ(I)	4.096	3.302	3.778	4.817	4.363	3.785
R_{split}	17.47	21.37	18.75	14.92	16.31	18.79
CC*	0.992	0.989	0.992	0.994	0.994	0.991
Wilson B factor (Å ^{2})	17.83	17.56	17.51	17.52	17.61	16.77
Refinement:						
No. reflect. all/free	24164 / 1215	23910 / 1141	23983 / 1151	24180 / 1158	24154 / 1156	24122 / 1152
R/R-free	0.171 / 0.211	0.176 / 0.219	0.170 / 0.216	0.156 / 0.195	0.170 / 0.209	0.178 / 0.214
Bonds	0.0097	0.0087	0.0097	0.0113	0.01	0.0099
Angles	1.58	1.554	1.611	1.669	1.632	1.642

Table A.16: TR-SSX Bragg X-ray diffraction data for varying energy densities.

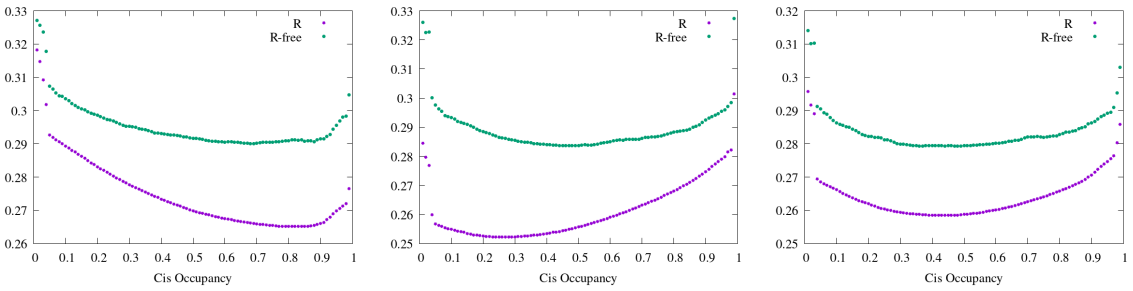


(a) pH 8.4 **(b)** pH 5.5

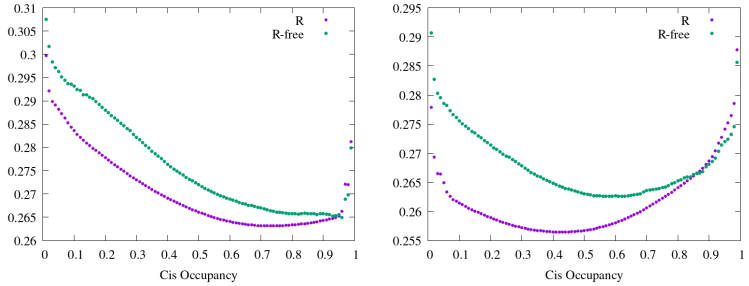
Figure A.12: LigPlot+ [329] plots showing the hydrogen bonding environment (green lines) around the residue Glu212, interacting residues, spectator residues and distances (in Å) for different crystal structures.

Dominant Composition	pD 7.8 On	pD 7.8 Off	pH 7.8 On	pH 7.8 Off
C=O	1665	1688	1665	1690
Arg 66 asym($CN_3H_5^+$)	1655		1674	
C=C	1624	1644	1628	1650
Arg 66 sym($CN_3H_5^+$)	1594		1609	
Phenol-1	1574	1615	1572	1617
Phenol-2		1600		N.D.
C=N/C=C	1549	1561	1547	1561
Phenol-3	1497	1524	1497	1518
$^{13}C_5$	1350	1358	1349	1362
Phenol	1152	1177	1150	1177

Table A.17: On and Off state assignments from Warren *et al.*[5] of the RSFP, Dronpa in ^2H_2O and ^1H_2O at pD/pH 7.8. All values in cm^{-1}.



(a) pH 8 Dark **(b)** pH 8 500 nm **(c)** pH 8 500 nm then 400 nm



(d) pH5 Dark **(e)** pH5 400 nm

Figure A.13: (a-e): Minimisation of crystallographic R-factors for JvT75 TR-SSX data at pH 5.5 and pH 8 under different laser illumination conditions.

	pH 8.4 Dark	pH 8.4 500 nm	pH 8.4 400 nm
Resolution (Å)	39.490-1.75	39.485-1.75	39.460-1.75
a (Å)	39.6	39.52	39.6
b (Å)	74.5	74.34	74.44
c (Å)	78.98	78.97	78.92
Completeness (%)	100	100	100
No. Crystals	10790	9342	11917
Red.	351.7	463.1	655.2
SNR	2.834102	3.301172	3.528936
R_{split}	26.52	21.19	20.2
$CC_{1/2}$	0.9574746	0.9663334	0.9723949
CC*	0.989078	0.9914023	0.9929775
B-factor (Å ^{2})	29.58	26	25.49
R_{Free}	20	19	19
R_{Work}	24	23	22

Table A.18: Crystallographic statistics for JvT75 collected at room temperature using fixed target serial crystallography.

	pH 8.4 500 400 nm	pH 5.5 Dark	pH 5.5 400 nm
Resolution (Å)	39.465-1.75	53.953-1.75	79.210-1.75
a (Å)	39.57	39.6	39.5
b (Å)	74.38	74.43	74.43
c (Å)	78.93	78.32	79.21
Completeness (%)	100	100	100
No. Crystals	7696	9015	13104
Red.	397.5	266.4	380.7
SNR	2.99898	2.870947	3.504225
R_{split}	23.52	25.9	21.23
$CC_{1/2}$	0.9445716	0.95066	0.9702119
CC*	0.9856449	0.987272	0.9924116
B-factor (Å ^{2})	25.63	29.84	31.51
R_{Free}	20	20	19
R_{Work}	23	23	23

Table A.19: Crystallographic statistics for JvT75 collected at room temperature using fixed target serial crystallography.