THESIS FOR THE DEGREE OF DOCTOR OF PHILOSPHY IN NATURAL SCIENCE

Ultrafast Structural Changes in a Bacterial Photosynthetic Reaction Center probed with XFEL Radiation

ROBERT DODS

UNIVERSITY OF GOTHENBURG

University of Gothenburg
Department of Chemistry and Molecular Biology
Gothenburg, Sweden, 2017

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSPHY IN NATURAL SCIENCE

Ultrafast Structural Changes in a Bacterial Photosynthetic Reaction Center probed with XFEL Radiation

Robert Dods

Cover: Global electron density difference map between the ground state of RC_{vir} and the photo-excited state at 300 ps overlaid on the ground state structure.

ISBN 978-91-629-0205-6 (Print)
ISBN 978-91-629-0206-3 (PDF)
Available online at: http://hdl.handle.net/2077/52084

Department of Chemistry and Molecular Biology
Biochemistry and Structural Biology
SE-405 30 Gothenburg
Sweden

Printed by Ineko AB
Gothenburg, Sweden 2017

Abstract

Photosynthesis is the process by which plants and many species of bacteria convert energy from sunlight into chemical energy used to power their metabolism. As these plants and bacteria are eaten, the chemical energy moves up the food chain and thus photosynthesis provides fuel for almost all life on Earth. Photosynthetic reaction centers are the workhorses of photosynthesis. Upon photo-excitation, these multi-domain integral membrane proteins drive an electron transport chain that results in a proton gradient across the cell membrane. The primary electron transport events are of great interest to the scientific community due to their near perfect efficiency and functional role in powering the biosphere. The articles that comprise this thesis deal with one such photosynthetic reaction center, that from the purple non-sulfur bacterium *Blastochloris viridis* (RC_{vir}). Spectroscopic studies of RC_{vir} have revealed that the initial charge-separation reactions occur on a time scale of picoseconds and raise interesting questions about the role of ultrafast structural changes in optimizing the efficiency of the overall process.

As X-ray free-electron lasers (XFELs) have been commissioned, the possibility of studying the initial light-driven reactions of the electron transport process through time-resolved crystallography has been realized. XFELs are powerful new X-ray sources that have a high peak brilliance and a pulse length three orders of magnitude shorter than the most advanced synchrotron source. Through the development of time-resolved crystallographic and solution scattering methods at XFELs, this thesis aims to deliver new information about the role structural changes play in guiding the charge separation reactions of photosynthesis.

A solution scattering experiment was performed to give physiological relevance to previous observations that multi-photon excitation led to quake like movements within RC_{vir} on the order of picoseconds. Oscillatory features were revealed following a single-photon absorption event, but these proved difficult to interpret structurally. This highlighted the need for time-resolved crystallography experiments that could directly visualize these structural changes. After optimizing crystallization methods to produce samples suitable for XFEL sources, a time-resolved crystallography experiment was conducted that captured the protein at two picosecond time-points following photo-excitation. These experiments allowed visualization of conformational changes that evolved over time and it is hypothesized these structural dynamics may play a role in altering the activation energies of the electron transport process.

Contribution list

Paper I I produced all the figures and contributed to editing and formatting the text.

Paper II I produced protein samples suitable for solution scattering at the XFEL. I helped with sample preparation on-site at the XFEL and aided with on-line data analysis.

Paper III I produced sample for all three XFEL beam times discussed in the paper and worked on the crystallization strategies. I processed and refined the structural data, wrote the manuscript and made the figures.

Paper IV I contributed to on-site sample preparation at the beam time at SACLA. I performed on-line data analysis at the LCLS beam time and contributed to processing and refining the structural data.

Paper V I expressed, purified and crystallized protein sample for the experiment. I carried out processing and structural refinement of the data. I analyzed electron difference density maps and carried out partial occupancy modeling on the light-activated states. I contributed to writing of the manuscript and produced the figures.

List of publications

Paper I **Robert Dods** and Richard Neutze. Elucidating ultrafast structural motions in photosynthetic reaction centers with XFEL radiation. *In press.*

Paper II David Arnlund, **Robert Dods**, Despina Milathianaki, Kenneth Beyerlein, Peter Berntsen, Chelsie Conrad, Garret Nelson, Erik Malmerberg, Cecilia Wickstrand, Linda Johansson, Rajiv Harimoorthy, Gisela Branden, Petra Båth, Amit Sharma, Chufeng Li, Yun Zhao, Leonard Chavas, Stella Lisova, Uwe Weierstall, Thomas White, Henry N. Chapman, John C. H. Spence, Garth Williams, Gerrit Groenhof, Sebastien Boutet, Daniel P. DePonte, Anton Barty, Jan Davidsson and Richard Neutze. Ultrafast structural changes in photosynthesis. *Manuscript.*

Paper III **Robert Dods**, Petra Båth, David Arnlund, Kenneth R. Beyerlein, Garrett Nelson, Mengling Liang, Rajiv Harimoorthy, Peter Berntsen, Erik Malmerberg, Linda Johansson, Rebecka Andersson, Robert Bosman, Sergio Carbajo, Elin Claesson, Chelsie E. Conrad, Peter Dahl, Greger Hammarin, Mark S. Hunter, Chufeng Li, Stella Lisova, Despina Milathianaki, Joseph Robinson, Cecilia Safari, Carolin Seuring, Amit Sharma, Garth Williams, Thomas White, Cecilia Wickstrand, Oleksandr Yefanov, Jan Davidsson, Daniel P. DePonte, Anton Barty, Gisela Brändén and Richard Neutze**.** From Macro-Crystals to Microcrystals: a Strategy for Membrane Protein Serial Crystallography. *Submitted manuscript.*

Paper IV Petra Edlund, Heikki Takala, Elin Claesson, Léocadie Henry, **Robert Dods**, Heli Lehtivuori, Matthijs Panman, Kanupriya Pande, Thomas White, Takanori Nakane, Oskar Berntsson, Emil Gustavsson, Petra Båth, Vaibhav Modi, Shatabdi Roy-Chowdhury, James Zook, Peter Berntsen, Suraj Pandey, Ishwor Poudyal, Jason Tenboer, Christopher Kupitz, Anton Barty, Petra Fromme, Jake D. Koralek, Tomoyuki Tanaka, John Spence, Mengning Liang, Mark S. Hunter, Sebastien Boutet, Eriko Nango, Keith Moffat, Gerrit Groenhof, Janne Ihalainen, Emina A. Stojković, Marius Schmidt & Sebastian Westenhoff. The room temperature crystal structure of a bacterial phytochrome determined by serial femtosecond crystallography. *Scientific Reports* **6**, 35279, doi:10.1038/srep35279 (2016).

Paper V **Robert Dods**, Petra Båth, David Arnlund, Robert Bosman, Kenneth R. Beyerlein, Garrett Nelson, Mengling Liang, Despina Milathianaki, Joseph Robinson, Rajiv Harimoorthy, Peter Berntsen, Erik Malmerberg, Linda Johansson, Rebecka Andersson, Sergio Carbajo, Elin Claesson, Chelsie E. Conrad, Peter Dahl, Greger Hammarin, Mark S. Hunter, Chufeng Li, Stella Lisova, Cecilia Safari, Amit Sharma, Garth Williams, Cecilia Wickstrand, Jan Davidsson, Daniel P. DePonte, Anton Barty, Gisela Brändén and Richard Neutze. Ultrafast Time-resolved Serial Femtosecond Crystallography of Photosynthetic Reaction Center. *Manuscript*.

Related Publications

Paper VI 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**, 1552-1557, doi:10.1126/science.aah3497 (2016).

Paper VII Przemyslaw Nogly, Valerie Panneels, Garrett Nelson, Cornelius Gati, Tetsunari Kimura, Christopher Milne, Despina Milathianaki, Minoru Kubo, Wenting Wu, Chelsie Conrad, Jesse Coe, Richard Bean, Yun Zhao, Petra Båth, **Robert Dods**, Rajiv Harimoorthy, Kenneth R. Beyerlein, Jan Rheinberger, Daniel James, Daniel DePonte, Chufeng Li, Leonardo Sala, Garth J. Williams, Mark S. Hunter, Jason E. Koglin, Peter Berntsen, Eriko Nango, So Iwata, Henry N. Chapman, Petra Fromme, Matthias Frank, Rafael Abela, Sébastien Boutet, Anton Barty, Thomas A. White, Uwe Weierstall, John Spence, Richard Neutze, Gebhard Schertler & Jörg Standfuss. Lipidic cubic phase injector is a viable crystal delivery system for time-resolved serial crystallography. *Nature Communications* **7**, 12314, doi:10.1038/ncomms12314

Paper VIII Cecilia Wickstrand, **Robert Dods**, Antoine Royant, & Richard Neutze. Bacteriorhodopsin: Would the real structural intermediates please stand up? *Biochimica et Biophysica Acta (BBA) - General Subjects* **1850**, 536-553, doi:10.1016/j.bbagen.2014.05.021 (2015).

Paper IX Rebecka Andersson, Cecilia Safari, **Robert Dods**, Eriko Nango, Rie Tanaka, Ayumi Yamashita, Takanori Nakane, Kensuke Tono, Yasumasa Joti, Petra Båth, Elin Dunevall, Robert Bosman, Osamu Nureki, So Iwata, Richard Neutze and Gisela Brändén. Serial femtosecond crystallography structure of cytochrome c oxidase at room temperature. *Submitted manuscript.*

Contents

Abbreviations

CSPAD	Cornell-SLAC Pixel Array Detector
CXI	Coherent X-ray Imaging
GPCR	G-protein coupled receptors
GDVN	Gas dynamic virtual nozzle
LCLS	LINAC Coherent Light Source
LCP	Lipidic cubic phase
LDAO	N,N-dimethyldodecylamine N-oxide
P_{960}	Special pair of chlorophylls in RC_{vir}
PSI/PSII	Photosystem I, II
PYP	Photoactive yellow protein
Q_A/Q_B	Quinone molecules in RC_{vir}
QH_2	Ubiquinol
RC_{vir}	Photosynthetic reaction center from *B. viridis*
TR-SFX	Time-resolved serial femtosecond crystallography
TR-WAXS	Time-resolved wide-angle X-ray scattering
SACLA	SPring-8 angstrom compact free-electron laser
SLAC	Stanford Linear Accelerator Center
XFEL	X-ray free-electron laser

1. Introduction

1.1 Photosynthesis

Jan Ingen-Housz was a Dutchman who lived a remarkable life. He spent his formative years hobnobbing in England amongst the company of the likes of Joseph Priestley, Benjamin Franklin and Henry Cavendish. In the late 1760s he was summoned to the court of the Austrian Empress Maria Theresa to inoculate her family against smallpox using emerging inoculation methods, some years before Jenner's discovery of the smallpox vaccine. He settled in Austria as the Empress' court physician and there developed an interest in gaseous exchange in plants. In 1779 he published 'Experiments upon vegetables' wherein he described the production of oxygen by leaves and the uptake of carbon dioxide into the physical mass of the plant[1], thereby founding knowledge of photosynthesis in scientific literature. Over two hundred years later we understand much more about this complex process that converts sunlight into chemical energy and the fantastically efficient cellular machinery that makes it work.

The first oxygenic photosynthetic bacterium evolved over 2 billion years ago. Through the process of splitting water molecules using energy from sunlight to form molecular oxygen, photosynthetic bacteria radically transformed the atmosphere of Earth, making it habitable for the vast diversity of life that exists today[2,3]. Now almost all species of plants and many bacterial species carry out this water splitting reaction. However, non-oxygenic photosynthesis evolved earlier still. In the highly reducing atmosphere of ancient Earth, species evolved that used hydrogen sulfide and hydrogen as reducing agents. Descendants of these bacteria survive today in reductive and hypoxic environments and comprise the group of bacteria known as purple and green photosynthetic bacteria[4,5]. The majority of papers presented in this thesis are studies on the photosynthetic machinery from *Blastochloris viridis*, a purple non-sulfur bacterium. This group of bacteria are named as such due to the variety of pigments they contain making many of them appear purple, although *B. viridis* contains many chlorophyll cofactors and is green. The 'non-sulfur' denomination separates this group from sulfur bacteria in that they have a cyclic flow of electrons and they primarily do not use external reducing agents. Sulfur bacteria use hydrogen sulfide as a reducing agent, resulting in excretion of sulfur compounds.

All photosynthetic organisms are thought to come from a common ancestor, and indeed the proteins that drive photosynthesis are structurally analogous across prokaryotes and eukaryotes[6]. The workhorses of photosynthesis are the two types of photosynthetic reaction centers, of which oxygenic photosynthetic organisms need both. Plants contain

the two reaction centers PSI and PSII embedded in the thylakoid membranes of their chloroplasts. Light is initially absorbed by the special pair of chlorophylls at the core of PSII (P_{680}) or the antenna complexes that surround it. This excites an electron in the special pair and begins an electron transport chain. The oxidized form of P_{680} (P_{680}^{+}) is the strongest known biological oxidizing agent with a redox potential of 1260 mV[7]. This is strong enough to oxidize water at the Mn_4CaO_5 oxygen evolving complex and subsequently reduce the protein's terminal electron acceptor plastiquinone to plastiquinol. Two rounds of the water-splitting reaction also generate four protons and a proton gradient is thereby formed across the thylakoid membrane. This proton gradient drives ATP synthesis, converting light energy into chemical energy. Oxygen is released as a byproduct of splitting two water molecules. The plastiquinol molecule further shuttles electrons *via* various other proteins into a second reaction center, PSI. The electron transport chain in PSI has a ferrodoxin protein as a terminal electron acceptor, and the reduced ferrodoxin drives reduction of NADP+ to NADPH. The ATP and NADPH produced by this process feed into the Calvin cycle, completing the conversion of sunlight into biomass[8]. It is hypothesized that eukaryotes gained the ability to perform photosynthesis through a symbiotic relationship with photosynthetic bacteria, which were then incorporated into the eukaryotic genome, explained by endosymbiotic theory[9-11]. PSI is analogous to the reaction centers found in green sulfur and non-sulfur bacteria, which also have an iron-sulfur protein as a terminal electron acceptor. PSII is structurally analogous to reaction centers found in purple bacteria, which have a quinone molecule as a terminal electron acceptor[6].

1.2 RC_{vir} structure and function

Paper I, **Paper II**, **Paper III** and **Paper V** in this thesis are works on the photosynthetic reaction center from *B. viridis*, RC_{vir}. RC_{vir} was the first membrane protein structure to be solved by X-ray crystallography in work that was later awarded a Nobel Prize in chemistry[12,13]. Since then, more than 20 structures of the protein have been deposited in the Protein Data Bank (PDB) [14-28]. RC_{vir} is a membrane spanning protein consisting of 10 transmembrane helices surrounding a pseudo-symmetrical core of cofactors (Figure 1.1). The protein is made up of four subunits, a light chain (L), a medium chain (M) a heavy chain (H) and a cytochrome-related subunit (C). The L and M subunits span the membrane, while the H subunit caps the protein on the cytoplasmic side of the membrane, and the C subunit caps the periplasmic side.

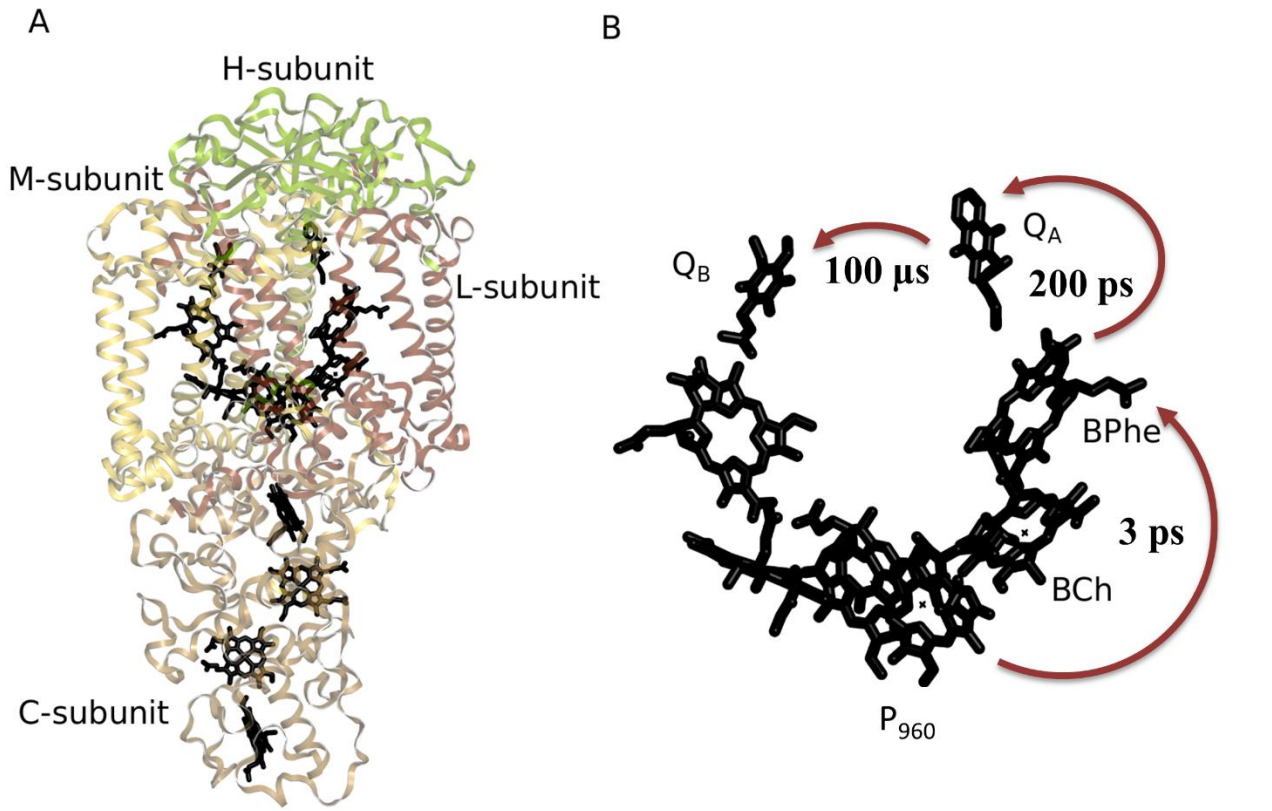


Figure 1.1: *Structure of photosynthetic reaction center from* B. viridis, *showing the arrangement of the subunits and cofactors. The pseudo-symmetrical core of cofactors is enlarged on the right along with the timescales for electron transport.*

The function of RC_{vir} is to drive an electron transport chain that eventually results in the proton gradient required to drive ATP synthesis and thus provide energy for the cell. The fundamentals of this electron transport is that the process begins in the special pair of chlorophyll molecules in the center of the protein, notated P_{960} after the wavelength of light they absorb. Absorption of one photon here excites an electron to a higher energy level which is then translocated up to the pheophytin (BPhe) molecule of the L subunit within 3 ps, creating a charge separated state[29-31]. Due to differences in energy potential, the electron is not transferred back to the special pair, but is shuttled again on to the bound menaquinone molecule of the L branch (Q_A) after about 200 ps[32-35], from which it is transferred further to the mobile ubiquinone molecule (Q_B) on a slightly longer timescale of 100 µs[36]. The photo-oxidized chlorophyll special pair (P_{960}^+) is reduced with an electron *via* the heme cofactors in subunit C. The time-scale of the replenishment of this electron on to the special pair is about 120 ns at physiological temperatures[37,38]. With the absorption of a second photon by P_{960}, a second electron is excited and another round of electron transfer begins. Once two electrons have reached the Q_B cofactor, it becomes doubly protonated forming ubiquinol (QH_2) and is released from the protein into the membrane. QH_2 is re-reduced by cytochrome bc1 resulting in

the net release of two protons into the periplasm, the electrons are transferred back to the C subunit of RC_{vir} *via* carrier protein c2, completing the cyclic flow of electrons. The proton gradient this process creates drives ATP synthase which couples proton translocation along its concentration gradient with synthesis of ATP from ADP and inorganic phosphate[39].

Absorption of the photon does not necessarily have to occur at P_{960} however, it can also occur at the monomeric chlorophylls (BChl) in the L and M subunits, or in LH1, the single chlorophyll-containing light harvesting complex that is associated with RC_{vir}. From these sites the photon energy is transferred to the special pair, from which the electron transport chain continues as before. The monomeric chlorophylls and the pigments in the light harvesting complex have non-overlapping absorbance profiles, broadening the range of light wavelengths that can be used for photosynthesis[40]. Spectroscopic evidence, as well as evidence from mutational studies, show that the electron transport occurs specifically along cofactors in the L subunit of the protein. The reasons for this have been explained in detailed spectroscopic studies on the electron transfer process[41-44], but the reason the protein evolved this unidirectionality is hitherto unknown.

The most remarkable aspect of RC_{vir} and other photosynthetic reaction centers is their efficiency; the quantum yield of the initial photo-oxidation of the special pair approaches unity[45,46]. This is in part due to the fact that the inherently unstable charge separated state rarely collapses before the electron transport process is completed. One hypothesis to explain this is that ultrafast structural movements alter the activation energy for the forwards and backwards steps of the electron transport process. The idea that this electron transport may be conformationally gated has been around for some time and is discussed in the next section. This thesis describes time-resolved crystallographic experiments that explore this idea of conformational gating to confer additional stabilization of the charge separated state.

1.3 Spectral evidence for ultrafast structural changes

Early femtosecond-resolution spectroscopic studies on RC_{vir} and the closely related reaction center from *R. sphaeroides* investigated the temperature dependence of the initial charge-separation reaction and concluded that the reaction rate actually increased upon cooling[47,48]. This resulted in the belief that ultrafast structural movements of the protein were not responsible for influencing this reaction. A 1993 article in Nature called into question this received wisdom with the demonstration that coherent vibrational modes coupled to the excitation of the special pair (P to P*) of the protein persisted for picoseconds even at room temperature, hinting at coherent structural movements on

these timescales[49]. This article was built on by further spectroscopic studies showing not only the residues surrounding the special pair to be important to the frequency of the vibrational modes but also more distant residues[50,51]. Tryptophan absorbance studies on *R. sphaeroides* also implied that protein dynamics controlled the kinetics of the electron transport process[52,53]. Further spectral evidence for ultrafast structural changes is provided by observations that showed the lifetime of the charge-separated state in the reaction center from *R. Sphaeroides* could be extended from 100 ms to 250 s by exposing the protein to bright light. This process was shown to be fully reversible after recovery in the dark, implying that conformational changes in the protein are able to directly tune the charge recombination reaction[54,55]. Direct observation of structural changes by time-resolved crystallography would allow much greater understanding of the implications of this spectroscopic data.

1.4 Synchrotron based time-resolved structural studies of RC_{vir}

There have been several attempts to investigate structural movements in light-activated reaction centers at synchrotron sources. Cryo-trapping studies of illuminated reaction center crystals have shown significant structural changes[56,57]. These experiments were carried out by illuminating reaction center crystals for 150 ms before flash freezing in order to trap the majority of the protein molecules in the charge-separated state. These experiments demonstrated at high resolution that the terminal electron acceptor, the quinone Q_B molecule, was shown to move 4.5 Å and twist around on its axis, moving from the distal binding site in the dark-adapted crystals to the proximal binding site in the light-adapted crystals (Figure 1.2A). This represents a movement towards the bound quinone molecule Q_A from which it accepts an electron.

Due to the nature of X-ray generation at synchrotron sources, the time-resolution of pump-probe experiments is greatly limited compared to spectroscopy by the length of the electron bunch, typically 100 ps without reducing photon flux[58]. However several synchrotron based time-resolved crystallography experiments have been carried out to investigate structural changes at ms time-delays using Laue diffraction[20,25]. Laue diffraction uses a polychromatic X-ray source that samples a large area of reciprocal space with every X-ray pulse. To use this technique in a pump-probe manner, the X-rays arrive at a specified time after excitation of the crystal by a laser pulse. This has the benefit that all data are collected from a single crystal, which removes any systematic errors resulting from differences between individual crystals. However prolonged exposure to both laser flashes and the X-ray beam can limit the diffraction quality of the crystals used. Time-resolved Laue diffraction is an ambitious undertaking, large crystals are needed to reduce X-ray and laser induced damage, and this has to be balanced against the requirement of sufficient laser excitation to increase the occupancy of the

excited state to an observable level[59,60]. The feasibility of time-resolved Laue experiments has been demonstrated by a large body of work on myoglobin and photo-active yellow protein[61-67] but has been mostly limited to small single-domain proteins. Performing time-resolved Laue experiments on reaction center crystals is further complicated by the fact that it is rather large (135 kDa) and as a membrane protein, it is more challenging to form large, well-diffracting crystals. Nevertheless, two time-resolved studies have been published. The first experiment on RC_{vir} crystals grown by vapor diffusion attempted to replicate the results previously shown by cryo-trapping studies, however no structural movements were identified above the noise level of the data[25]. This may be due to the low resolution of the crystals and the X-ray induced damage they received. A second Laue experiment was performed on a different crystal form of the protein, grown in the lipidic sponge phase. This experiment showed convincing difference electron density around a tyrosine residue (L162) that moved towards the excited special pair 3 ms after photo-excitation (Figure 1.2B). It was hypothesized using evidence from free energy calculations that the conserved tyrosine residue became deprotonated and played a role in stabilization of the charge-separated state[20].

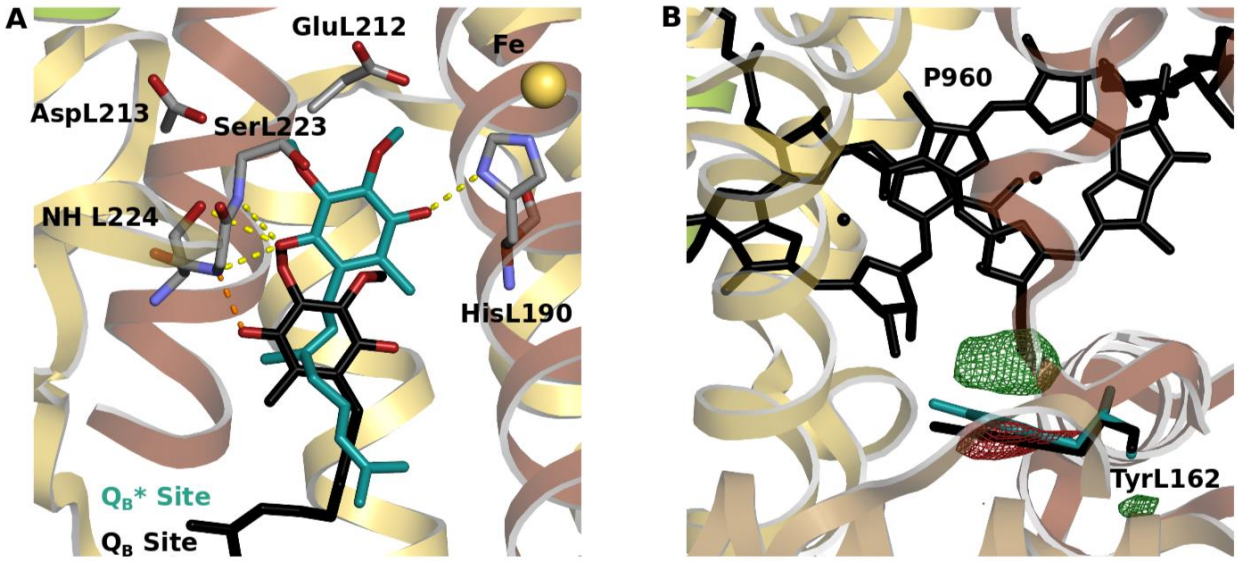


Figure 1.2: A) *The Q_B molecule in* R. sphaeroides *rotates on its axis and changes its hydrogen bonding network after moving from the distal site in the dark-adapted state (black) to the proximal site in the light-activated state (cyan)[56,57]. **B)** Light induced structural change captured by Laue crystallography 3 ms after light exposure[20]. Difference electron density map between light and dark states is contoured at 3 σ, negative electron density is displayed in red and positive electron density is displayed in green. Model of the light-activated conformation is displayed in cyan.*

1.5 Harnessing XFEL radiation for structural biology

1.5.1 What is an XFEL?

X-ray free-electron lasers (XFELs) are powerful X-ray sources that open up new possibilities in the study of ultrafast structural changes such as those being investigated in photosynthetic reaction centers. XFELs generate X-rays by accelerating electrons along a linear accelerator and inducing emission of high energy X-rays through interaction with undulators. The X-rays produced by XFELs have a peak brilliance 10^9 times higher than the brightest synchrotron sources[60]. The pulse lengths of XFELs are less than 100 fs, three orders of magnitude shorter than those possible at synchrotron sources, where the pulse length is limited by electron bunch duration in storage ring facilities[58]. Certain aspects of XFELs make them interesting to structural biologists. Although the repetition rates of currently commissioned XFELs are generally lower than synchrotrons (120 Hz at the Linac Coherent Light Source (LCLS) XFEL, Stanford, California), the high peak brilliance allows every pulse to be used to collect diffraction data and large datasets can be accumulated rapidly. The new European XFEL, expected to be commissioned in 2018, will have a repetition rate of 27,000 Hz[68] and a catch-up game is being played in the development of detectors with read-out rates that can utilize this aspect of the latest XFEL technology. The high peak brilliance allows structural information to be obtained from smaller crystals than ever before and there is considerable work being undertaken to push that even further to single particle X-ray imaging[69-71]. The large photon dose delivered in this short pulse-length allows a new world of ultrafast processes to be explored, from material scientists performing time-resolved near-edge spectroscopy experiments[72], to chemists studying the dynamics of bond breakage[73,74]. For structural biologists it gives the potential to study ultrafast structural movements of proteins, such as those postulated to occur in RC_{vir}.

1.5.2 "Diffraction before Destruction"

The problem with performing a conventional diffraction experiment using powerful XFEL radiation, is that the X-ray dose delivered to the crystal from a single shot of the XFEL beam would cause rapid break-up by radiation damage well before a meaningful rotation series can be performed. Before the crystal explodes however, it will produce a meaningful diffraction pattern, as first described by Neutze *et al*. seventeen years ago, where molecular dynamics simulations were used that demonstrated there was a lag time on the order of femtoseconds after initial X-ray dose before the protein structure has time to respond and reflect this damage[75]. This idea is termed "diffraction before destruction" and is the basis for structural biology experiments at an XFEL[76].

In a synchrotron crystallography experiment, typically a single crystal is rotated through

the X-ray beam, taking a diffraction image after every increment of rotation until all of reciprocal space has been sampled. Due to the destruction of the crystals this is clearly not possible when using a highly focused beam at an XFEL source. Instead a serial method termed serial femtosecond crystallography (SFX) is deployed, passing a stream of crystals of random orientation through the X-ray beam and collecting diffraction images whenever the X-ray pulse hits a crystal. The processing of these single diffraction snap shots into usable structure factor amplitudes is described in chapter 2.5.

In order to carry out successful SFX experiments, there was a requirement for development of rapid-readout detectors as well as a system to deliver a continuous sample of crystals into the XFEL beam. The Cornell-SLAC Pixel Array Detector (CSPAD) developed at the Stanford Linear Accelerator Center (SLAC) was designed specifically to deal with the high repetition rate of the XFEL and the high dynamic range required, its use is discussed further in section 2.3.1[77]. The European XFEL will have a repetition rate of 27 KHz, and work is being undergone to design detectors able to cope with this[78].

Various methods have been developed to deliver crystals into the beam, the first of which was the gas dynamic virtual nozzle (GDVN)[79]. The HPLC-based system pumps a crystal suspension through a glass or ceramic nozzle and focuses the outgoing liquid with helium gas to form a liquid jet. This proved to be the bedrock method for structural studies using XFEL sources, and remains frequently used. However the GDVN system does have a major drawback in that in order to form a steady column of liquid that allows interaction with X-rays before Rayleigh break-up, the liquid moves at speeds of typically 10 ms^{-1}. This is far faster than the repetition rates of currently commissioned XFELs, and results in over 99.9 % of the sample never being probed[80]. Typical SFX experiments using the GDVN sample delivery system may require over 10 mg of purified protein, putting it out of reach for many scientists working on proteins that are difficult to purify. One solution to this problem has come from the development of the lipidic cubic phase (LCP) jet, specifically designed to use proteins embedded in the lipidic cubic phase. LCP is made from a precise mixture of aqueous phase and lipids that somewhat resembles the lipid bilayer of cells and has proved valuable in the crystallization of difficult membrane protein structural targets, including G-protein coupled receptors (GPCRs), a family of receptors of high importance in structure-based drug design[81-86]. This viscous lipid mixture can produce a jet of much slower speeds than an aqueous jet, typically 1000 fold lower, more closely matching the repetition rate of the XFEL and cutting sample consumption significantly[87]. While soluble protein crystals may be sensitive to mixing with LCP, numerous grease carrier matrices have been developed to extend the use of this jet to samples of soluble proteins[88,89].

1.5.3 TR-SFX at an XFEL

XFELs have provided a great leap forward in the time-resolution possible for light-induced time-resolved crystallographic studies. By pumping the crystal samples with a femtosecond laser and probing with the XFEL beam after a specified time delay, time-resolved diffraction studies with a time-resolution down to the hundreds of femtoseconds can be achieved. The large datasets that can be accrued with the repetition rate of XFELs has also placed the idea of a 'molecular movie' of real time structural changes within reach. This has been demonstrated by several landmark time-resolved SFX (TR-SFX) papers.

The first TR-SFX paper for a soluble protein was published in 2014 after experiments carried out at the LCLS XFEL. The experiment examined ultrafast structural changes in photo-active yellow protein (PYP). The authors collected datasets 1 us and 10 ns after photo-excitation. Data collected were used to visualize structural movements in the protein brought about by chromophore isomerization and showed good agreement with time-resolved studies using Laue diffraction[61,90,91]. This article proved the feasibility of producing high quality difference electron density maps using pump-probe TR-SFX[92].

One major TR-SFX experiment on a membrane protein was performed on bacteriorhodopsin at the SPring-8 Angstrom Compact free electron Laser (SACLA), Hyogo, Japan **(Paper VI)**[93] after previous experiments showing the lipidic cubic phase was compatible with time-resolved experiments (**Paper VII)**[94]. Bacteriorhodospin has a wealth of scientific literature exploring its photo-cycle[95-102] and yet the precise structural movements and particularly the amplitude of these structural changes remained controversial[103]. The TR-SFX experiment collected data over 13 time points, tracing structural changes at high resolution on a logarithmic scale from the nanosecond to millisecond time scale. The development of structural movements across time was clear from electron density difference maps and it allowed key questions about the timing of displacement of a key water residue to be finally laid to rest.

A further time-resolved experiment was recently published on PSII from a cyanobacterium[104]. This provided structural answers to questions surrounding the mechanism of the oxygen-evolving manganese cluster. It also showed a rotation of the head group of the reduced quinone (Q_B) cofactor similar to that seen in cryo-trapping studies of the analogous reaction center from *R. sphaeroides*[56,104].

There are several interesting directions being taken in order to carry out TR-SFX experiments on proteins that are not light-induced. A proof-of-principle experiment using a 'mixing-jet' developed at Arizona State University demonstrated the feasibility

of studying enzyme reactions at time-scales down to the millisecond range[105]. Time-resolved experiments based on mixing by diffusion is limited at synchrotrons due to the large size of the crystals needed which hinders substrate diffusion. Conversely, at XFELs, crystals of less than 1 um can be probed, reducing diffusion times by orders of magnitude. Another experiment used strong electric field pulses to stimulate protein dynamics and found concerted protein movements on the sub-microsecond timescale that demonstrated good consistency with the conformational changes induced by substrate binding[106]. A further idea about stimulating protein dynamics with terahertz radiation has been shown to work at a synchrotron but has not yet been carried out at an XFEL source[107].

1.5.4 TR-WAXS at an XFEL

Time-resolved wide-angle X-ray solution scattering (TR-WAXS) is a technique developed at synchrotrons that uses the diffuse scattering of protein molecules in solution to analyze conformational changes. The theory behind solution scattering is discussed in section 2.4 but the key point that theoretical solution scattering curves can be calculated from known atomic coordinates has been crucial to giving structural explanations for WAXS observations. Synchotron TR-WAXS experiments on hemoglobin and myoglobin have been carried out that tracked structural changes following photolysis of a bound carbon monoxide ligand[108-111]. Details of the proton transport mechanism of bacteriorhodopsin and proteorhodopsin were also described using TR-WAXS[112]. Solution scattering studies on bacterial phytochromes have demonstrated the large global structural changes that occur in these proteins upon photo-isomerization of the biliverdin cofactor, and show good agreement with a crystallographic structure of the excited state[113,114]. TR-WAXS studies such as **Paper II** have also been performed using XFEL sources[115-117] and the articles discussed below have built on previous work by greatly extending the time-resolution achieved.

In 2014, a solution scattering study on RC_{vir} was published[115]. This experiment pumped the protein at a high laser power, equivalent to 800 photons per RC_{vir} molecule, at a wavelength corresponding to absorption by the monomeric chlorophyll cofactors. Difference scattering curves were obtained at various time points on a scale from -5 ps to 280 ps after photo-excitation. The results showed significant structural movements in the protein peaking at 7 ps and slowly evolving and damping thereafter. These difference scattering curves were recreated from molecular dynamics simulations on hundreds of pairs of ground state and photo-excited structures. The pairs of structures that gave the best fit to the difference scattering curves were averaged together. The article looked at Cα internal difference matrices and described how the difference scattering curves could be explained by an outward 'breathing' movement of the trans-membrane helices

surrounding the special pair as the energy absorbed by the protein was rapidly distributed to the surrounding solvent. This is termed a 'protein quake' and the significance of the results described in this article is further explored in **Paper II**.

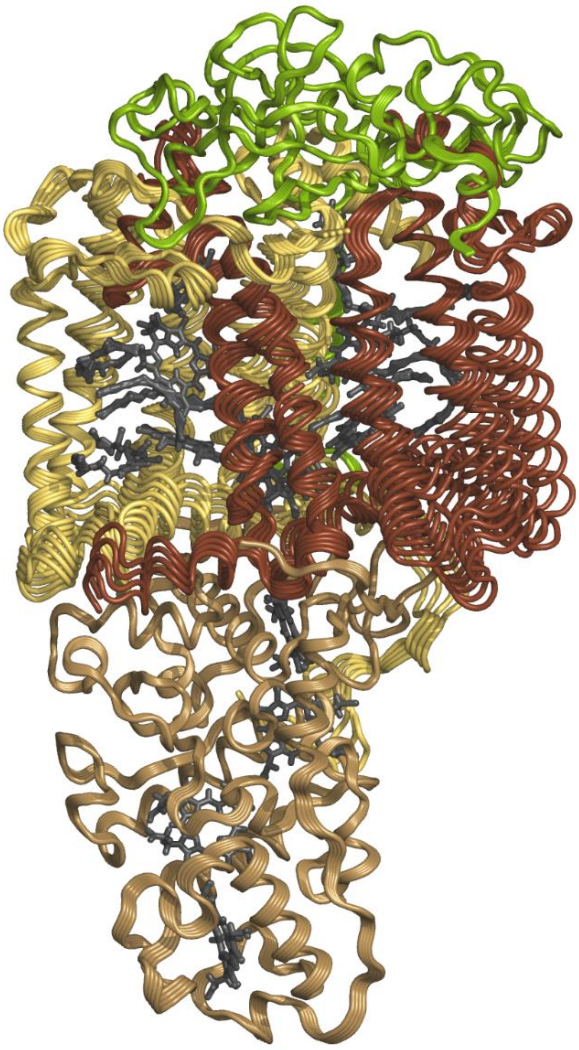


Figure 1.3: *A visualization of the protein quake described by Arnlund et al. Finding the best fits of molecular dynamics simulations to WAXS difference scattering curves resulted in models that showed an outward expansion of helices surrounding the cofactors. These movements arose on the order of picoseconds before then dampening. Structural movements have been exaggerated nine-fold for clarity.*

A TR-WAXS study on myoglobin demonstrated that following photolysis of bound carbon monoxide, the radius of gyration increased ~1 Å after 1 ps with damped structural oscillations on a 3.6 ps timescale as the system approached equilibrium. This provided more evidence that small localized changes in a chromophore brought about by photo-excitation could result in large structural changes over an entire protein[116]. These results were consistent with a later high resolution time-resolved SFX study with sub-picosecond resolution. This study postulated ultrafast helix motions around the chromophore[117].

1.5.5 Scope of this thesis

As described above, there has been much interest in the role structural changes play in the stabilization of the charge-separated state of reaction centers as well as interest in structural dynamics of proteins in general on the picosecond timescale. The aim of this thesis was to utilize emerging structural biology methods using XFEL sources to directly examine the ultrafast structural changes that occur upon photo-excitation of RC_{vir}. The goal was to study whether these changes occur, and if so, what role they may play in the fascinating efficiency of this photosynthetic machine.

The first paper in this thesis (**Paper I**) gives a general overview of the photosynthetic reaction center from *B. Viridis*, providing a base for the thesis from which to build on. It examines the question of conformational gating and explores the possibilities for elucidating ultrafast structural changes using XFEL sources.

Previously published material described a 'protein quake' of structural movements that occur on the picosecond time-scale after a multi-photon absorption event within the monomeric chlorophyll cofactors. By changing the wavelength of the pump-laser to directly excite the special pair and performing a power titration to track structural changes down to a single-photon absorption event, one aim of this thesis was to give a more physiological relevance to these previous findings. This involved expressing and purifying gram quantities of RC_{vir} in solution and collecting data over 60 hours at the LCLS. The signal to noise levels at lower laser powers proved to be very low, and novel data rejection methods were developed to increase the quality of the data. These data are presented in detail in **Paper II**. These data underlined that to truly understand the structural changes observable with solution scattering, time-resolved crystallography would be needed.

In order to perform TR-SFX at the XFEL, crystals measuring less than 20 um in all dimensions were required in order to fit through the narrow tubing used in the GDVN sample delivery system. Over three experiments at the LCLS, micro-crystals were optimized, giving a significant improvement in diffraction quality at every experiment. The improvements in micro-crystal diffraction were gained by implementing a seeding strategy in protein crystallization, and the resolution could be improved from 3.4 Å at the first experiment performed, down to 2.4 Å at the final experiment. This leap in resolution allows a much clearer insight into the structure of the protein. Comparing the structures obtained to those obtained from synchrotron radiation sources, it was noticed that the XFEL structures showed much clearer electron density at highly flexible regions of the protein. This was true even when comparing to synchrotron structures of much

higher resolution. The reasons for this and the implications for structural biologists using XFELs are discussed in **Paper III**. A micro-crystal strategy presented in Paper III was also used in **Paper IV** to collect data on phytochrome crystals at SACLA.

During one TR-SFX experiment at the LCLS, data were collected on RC_{vir} crystals in the ground state and at two time points corresponding to 5 ps and 300 ps after photo-excitation. The task of understanding the structural relevance of electron density difference maps between the photo-excited and ground states is presented in **Paper V**. Through implementation of a Bayesian q-weighting technique[118], maps could be improved to the extent that key structural changes could be modeled occurring around the active branch of the protein on the ultrafast time-scale. The role that these structural changes could play in stabilizing the charge-separated state of photo-activated RC_{vir} is described herein.

2. Methodology

2.1 Expression and Purification

2.1.1 Protein expression

Macromolecular crystallography experiments typically require microgram to milligram quantities of purified protein sample. This requires a robust protein expression and purification protocol in order to produce enough protein and subsequently not to lose too much through the purification procedure. SFX experiments at an XFEL using a GDVN sample delivery system can require hundreds of milligrams, and thus these requirements are even stronger. The majority of structural biology targets are expressed in insufficient quantities in their host cell, and thus a range of expression systems are used by transforming the protein gene in an overexpressed vector into *E. coli* or yeast cells. However, under the right conditions, *B. viridis* can be induced to express RC_{vir} in large quantities, and thus expression of RC_{vir} was carried out in the native host. Under a controlled system of aerobic growth in the dark, followed by anaerobic growth under light, RC_{vir} is highly expressed in levels shown to be inversely proportional to the intensity of the light[119].

2.1.2 Membrane protein purification

The first step in the purification of membrane protein samples from expression systems involves disrupting cells and separating cell lysates away from cell membranes by ultra-centrifugation. This is followed by solubilization of cell membranes in a suitable detergent and finally separating the protein of interest from all other proteins by taking advantage of differences in binding properties, charge or size.

There are a number of mechanical ways to break open cells. Disruption by sonication uses high-frequency sound waves to break open cell membranes while disruption by French press and X-press rely on pressure changes that 'pop' the cells open. Cell disruption of *B. viridis* is achieved by sonication. One drawback of using this technique is the heating of the cells that sonication causes which can damage temperature-sensitive proteins. If sonication was carried out on ice with pauses between sonic pulses, the protein yield was unaffected.

Membrane solubilization has been well researched and it is understood that in order to maintain the folding of membrane proteins a mild detergent should be used that produces a micelle environment as similar as possible to the lipid bilayer of cells. Concentrations of detergents used are normally high for the membrane solubilization step but can be reduced throughout purification to aid creation of crystal contacts.

However the detergent concentration must be kept above a critical micelle concentration at all times to prevent the protein from falling out of solution and aggregating[120]. A large number of commercial detergents are available for this task and the size and polarity of lipids can have an effect on maintaining protein structure and function. LDAO (N,N-dimethyldodecylamine N-oxide) is a zwitterionic detergent and has been shown to be a suitable detergent for purification of RC_{vir}[13].

Proteins can be separated based on their characteristics using column chromatography. By combining several types of chromatography, protein samples of high purity can be produced. Many structural biology targets expressed in non-native hosts are tagged at one terminus with a binding domain such as a poly-histidine tag. Immobilized metal-affinity chromatography can then be used to separate this nickel-binding histidine-tagged protein from all other proteins. The tagged protein will bind favorably to nickel on the column and can be eluted with an increasing concentration of imidazole, which competes for nickel-binding. This is a robust technique and is widely used. Since RC_{vir} is expressed in the native host, it is not tagged, and this type of chromatography cannot be used.

For the purification of RC_{vir}, ion exchange chromatography is used followed by a final step of size exclusion chromatography. The ion exchange step relies on differential binding to a charged medium by differently charged proteins which are then eluted based on the strength of their binding by an increasing salt concentration. The size exclusion step passes a mixture of proteins through a porous column, the size of the pores in the media reflects the effective column volume that differently sized proteins experience. Large proteins experience a reduced effective column volume relative to smaller proteins and will be eluted first. These two steps proved effective in producing purified RC_{vir} samples suitable for WAXS and SFX experiments. The purity of RC_{vir} samples was assessed by comparing absorption at 280 nm with absorption at 830 nm. Absorption at 280 nm reflects tyrosine and tryptophan residues present in most proteins, whereas 830 nm is specific to an RC_{vir} cofactor. An RC_{vir} solution with a ratio of $A_{280/830} < 2.4$ is suitable for crystallization.

2.2 Protein Crystallography

2.2.1 X-ray Crystallography fundamentals

A protein crystal is made up of a repeated, ordered arrangement of protein molecules in a lattice. The smallest repeating unit with translational symmetry is known as a unit cell. The unit cell can be further subdivided into an asymmetric unit which may contain as little as a single protein molecule. The asymmetric unit can describe the unit cell through rotational and translational symmetry operations. Crystals for X-ray crystallography

experiments can be as small as nanometers with the latest X-ray sources like the XFEL, but can also be over a millimeter in all dimensions and easily visible to the naked eye. Even small crystals will contain billions of protein molecules in this ordered arrangement. When X-rays are passed through an electron cloud, they are scattered. When X-rays pass through the electron clouds of an ordered crystal lattice most of this scattering is elastic scattering in which the incoming and scattered X-ray wave have the same wavelength. Although most of this scattering interferes destructively and becomes negated, under specific conditions the scattering is coherent and the intensity of this scattered X-ray beam can be observed as a Bragg reflection. The larger the crystal, the more coherent scatterers are present and the greater the intensity of the reflection. The conditions for coherent X-ray scattering are described by Bragg's law:

$$n\lambda = 2d \sin\theta \qquad (1)$$

Where n is the order of diffraction, λ is the wavelength of the X-rays, θ is the angle between the incident X-rays and lattice plane, and d is the spacing between planes in the crystal lattice. A visual representation of diffraction from a Bragg reflection is shown below in Figure 2.1.

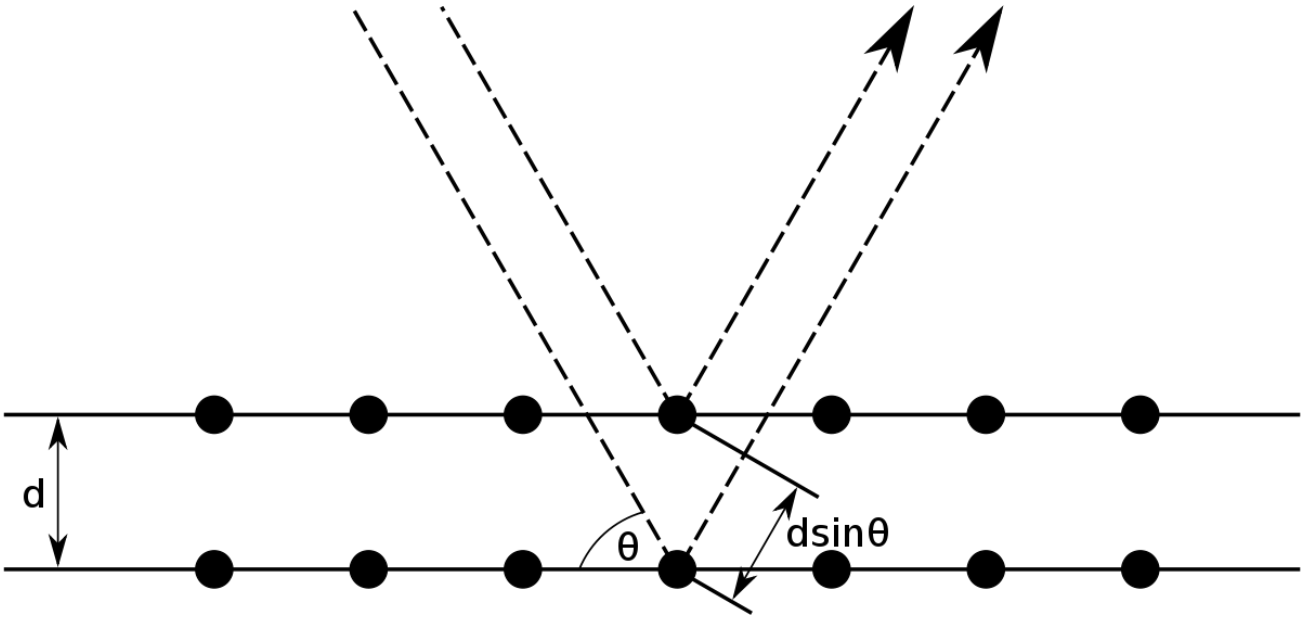


Figure 2.1: *A visualization of the fulfillment of Bragg's conditions, resulting in the coherent scattering of two incident X-ray waves.*

The relative intensities of Bragg reflections are related to the electron density of the unit cell by a Fourier transform, but the entire reciprocal space must be sampled in order to extract meaningful electron density. Conversion of recorded intensities to electron density is further described in section 2.5.1. The Ewald sphere represents the area of

reciprocal space fulfilling the Bragg conditions. When a reciprocal lattice point of the crystal lies on this sphere then a Bragg reflection can be observed. The sphere and the Bragg reflections sampled within it are shown in Figure 2.2 representing a theoretical X-ray crystallography experiment. By rotating the crystal within the X-ray beam, different lattice points in reciprocal space pass through the sphere and a complete sampling of reciprocal space can be achieved.

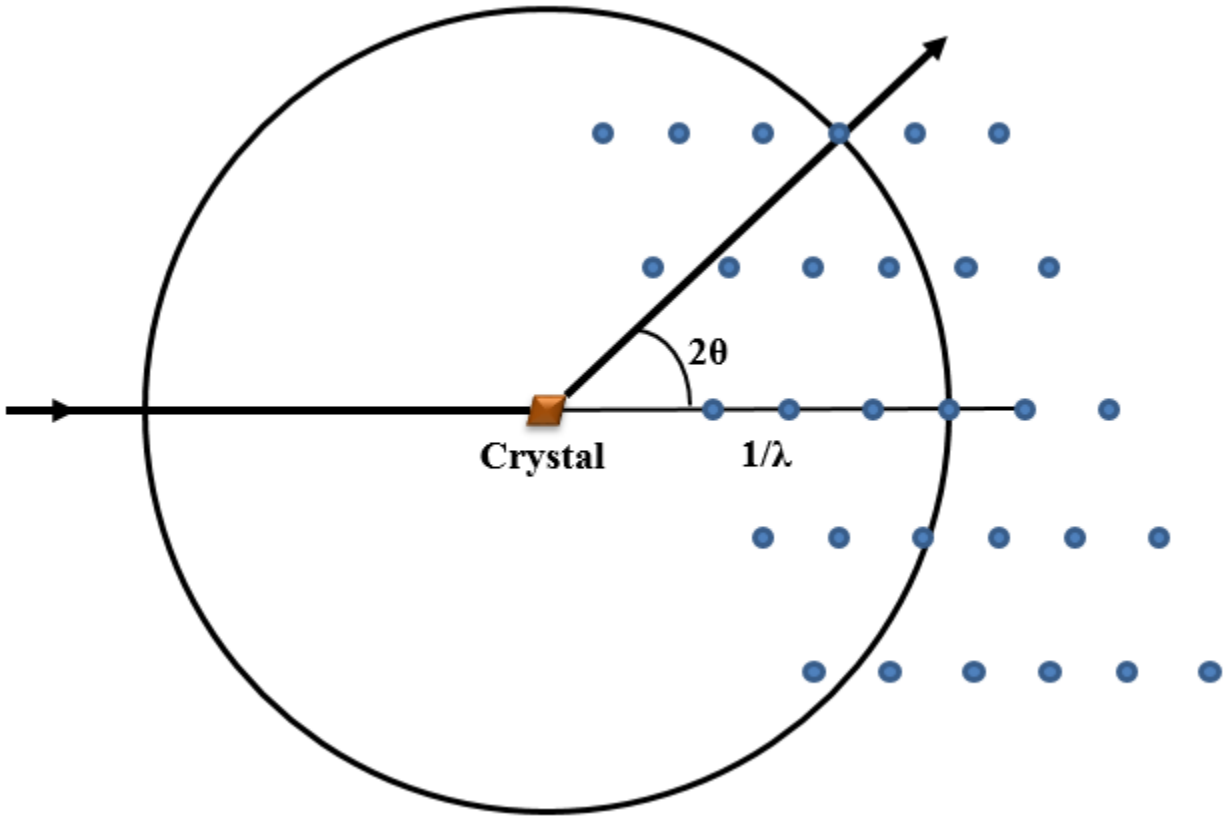


Figure 2.2: *Visualization of the Ewald sphere and the reciprocal lattice. Reciprocal lattice points that lie on the surface of the sphere fulfill the Bragg conditions and result in coherent scattering of the incident X-ray beam.*

2.2.2 Crystallization strategies

Inducing protein molecules to form a protein crystal is not a trivial process. Whereas small molecules such as sodium chloride readily form an ordered salt crystal upon saturation in solution, a saturated solution of protein molecules under the wrong conditions will aggregate together to form disordered precipitates. The formation of these disordered conglomerations of protein molecules is kinetically favored over the formation of a highly ordered crystal lattice, but nevertheless, under the right conditions,

crystallization can be induced. The process is further complicated for membrane proteins by the fact they are extracted from the membrane using detergents and are surrounded by detergent micelles of differing sizes, limiting the space for protein-protein crystal contacts. This, along with the fact that many membrane proteins are poorly expressed and difficult to purify, has made them difficult targets for structural biology. As a result they are relatively underrepresented among deposited structures in the PDB. Many factors have been shown to play a role in inducing crystallization of membrane proteins, including the detergent used, the pH of the crystallization buffer, the salts present in the crystallization buffer and the concentration of protein used. Finding a crystallization strategy for a given protein target generally involves screening many hundreds or thousands of different conditions. Even after initial crystal conditions have been found, they often do not yield large, well-diffracting crystals. Further rounds of optimization screening around initial crystal conditions are often performed in order to yield crystals of sufficient quality for X-ray crystallography experiments.

Although many different general strategies for protein crystallization have been developed, the most common method remains the vapor diffusion method. In a standard vapor diffusion experiment, a purified protein solution is mixed with a precipitant solution and left to equilibrate against a reservoir of more negative osmotic potential that draws water out of the crystallization drop over time. This leads to supersaturation of the crystallization drop and the protein may fall out of the solution as crystal nuclei. As more water is drawn from the solution, the crystal nuclei grow into larger crystals as more and more protein molecules are forced out of solution. Eventually crystal growth stops when the crystallization solution and the reservoir solution reach an equilibrium. The relationship between protein and precipitant concentrations are shown in Figure 2.2. The goal of a standard vapor diffusion experiment is to begin with a sufficiently high protein concentration, such that upon movement of water into the reservoir, the nucleation zone is reached. If crystal nuclei form, the protein concentration will

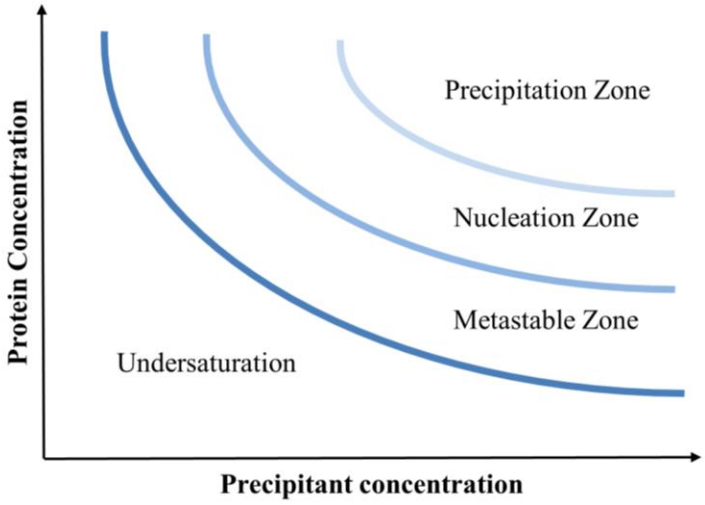


Figure 2.3: *A solubility diagram demonstrating the relationship between precipitant and protein concentrations. Protein molecules will form crystal nuclei in the nucleation zone and these will grow in the metastable zone. Protein aggregation occurs in the precipitation zone.*

begin to fall and the metastable zone will be reached in which growth upon the nuclei occurs but no new nuclei are produced. If too high a precipitant concentration is used, or the drying of the crystallization drop occurs too rapidly, the precipitation zone will be reached and the protein will fall out of solution as disordered aggregates. If too low a protein concentration or precipitant concentration is used, the solution will remain in the undersaturated region and the protein will neither crystallize nor precipitate. Two basic vapor diffusion methods exist, the hanging drop, and the sitting drop, which is displayed in Figure 2.4A. RC_{vir} was first crystallized by the sitting drop method[13], and these conditions were built upon for the work in this thesis in **Papers III and V**. Crystals of RC_{vir} grown by the sitting drop method are displayed in Figure 2.4B

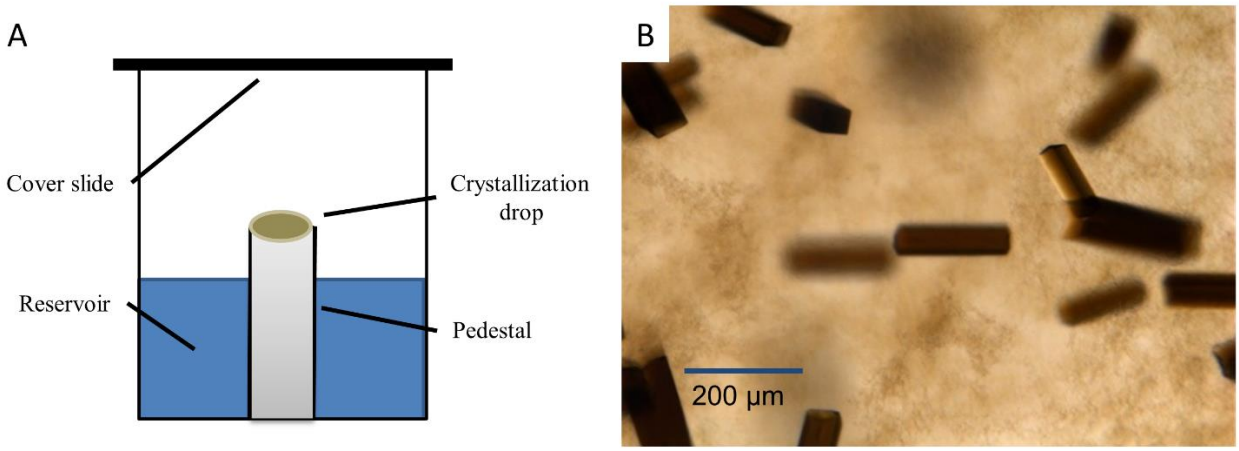


Figure 2.4: A) *Schematic of a sitting drop vapor diffusion experiment. A crystallization drop sits atop a pedestal surrounded by reservoir solution in a sealed environment. Water moves from the crystallization drop to the reservoir following its osmotic potential.* ***B)*** RC_{vir} *crystals grown by the sitting drop method.*

2.3 XFEL experimental set-up

Owing to the extremely high peak brilliance used at an XFEL, the experimental set up used for data collection differs drastically to that conventionally used at synchrotron radiation sources. XFEL pulse lengths on the order of femtoseconds also opens up new capabilities in the time resolution that can be achieved in time-resolved studies. Experiments described in this thesis were mostly carried out at the Coherent X-ray Imaging (CXI) beamline[121] at the LCLS. The experimental set-up described herein is specific to this instrument, although the general principles can be applied to experiments at any XFEL.

2.3.1 Sample delivery and data collection at an XFEL

As described in section 1.5.2, a single pulse from the XFEL will destroy a protein crystal. Thus unlike in synchrotron experiments, it is not feasible to collect an entire dataset from a single macro-crystal. A continuous stream of fresh crystals must be delivered into the path of the X-rays, perpendicular to the beam and at a speed that takes into account the high repetition rate of the XFEL (120 Hz). This is achieved by forming a continuous high velocity microjet using one of various forms of microjet technology. In the experiments described in this thesis, a gas dynamic virtual nozzle (GDVN) was used[79]. This HPLC based system pumps aqueous solution from a reservoir through narrow tubing (75 μm**)** and uses helium gas to focus the jet as it flows out of a nozzle made from glass or ceramics. The reservoir is gradually rotated 180º back and forth in an anti-settling device and can be cooled for temperature-sensitive samples[122]. The X-ray beam is aligned to interact with the microjet before Rayleigh breakup into droplets and a rapid-readout CSPAD detector is used to collect diffraction images at the same repetition rate as the X-ray pulses[77]. The CSPAD detector at the LCLS is made up of 64 panels each comprising of 192 x 185 pixels (pixel size = 110 μm^2)**.** When the X-ray beam encounters a crystal, the X-rays will be diffracted according to Bragg's law and will be recorded as peaks on the detector. Figure 2.5 is an example of a single diffraction image recorded on the CSPAD.

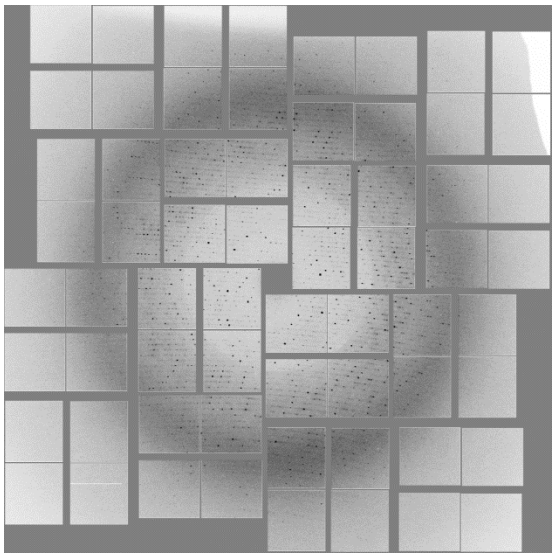


Figure 2.5: *Diffraction image recorded at the CXI beamline at the LCLS XFEL from an* RC_{vir} *micro-crystal.*

The diffraction image obtained from a single crystal represents the protein crystal in a randomly orientated direction and only by collecting diffraction patterns from a large number of crystals can the whole of the reciprocal space be sampled. With the high repetition rate of the XFEL, large datasets are quickly accumulated numbering millions of images. Many of these images will not represent crystal diffraction as it is often not possible to form extremely concentrated crystal dispersions, and when they are formed, it can cause clogging in the narrow tubing of the GDVN. Thus it is important to be able to identify crystal hits and reject all other images to reduce the data set and ease the computational power required for data processing.

Data reduction is performed by CHEETAH[123]. CHEETAH converts the raw data from the detector into the hierarchical HDF5 format and performs a background subtraction using a calibration detector image for which the X-ray beam is turned off. It can also mask out areas of the detector with no interpretable data such as where there is scattering from the jet, shadowing on the detector by the nozzle, or dead pixels on the detector. Finally, the software identifies images as crystalline based on a given number of peaks and rejects all others. Data processing is further described in section 2.5.1.

The set-up for a TR-WAXS experiment is almost identical to that for TR-SFX with the only major difference being that there is no crystal diffraction giving rise to Bragg peaks and CHEETAH retains all diffraction images, sorted into laser-on and laser-off datasets. Due to fluctuations in the X-ray beam and occasional breakup of the microjet, many of these diffraction images do not contain interpretable data. Data rejection is therefore implemented in post-processing and is further described in section 3.2.

2.3.2 Pump-probe time-resolved experiments at an XFEL

The basis of pump-probe time-resolved experiments is that a light-activated protein system is first activated with a laser of specific wavelength (the pump) and thereafter probed with X-rays to observe changes in the perturbed system (the probe). An overview of this experimental set-up is given in Figure 2.6.

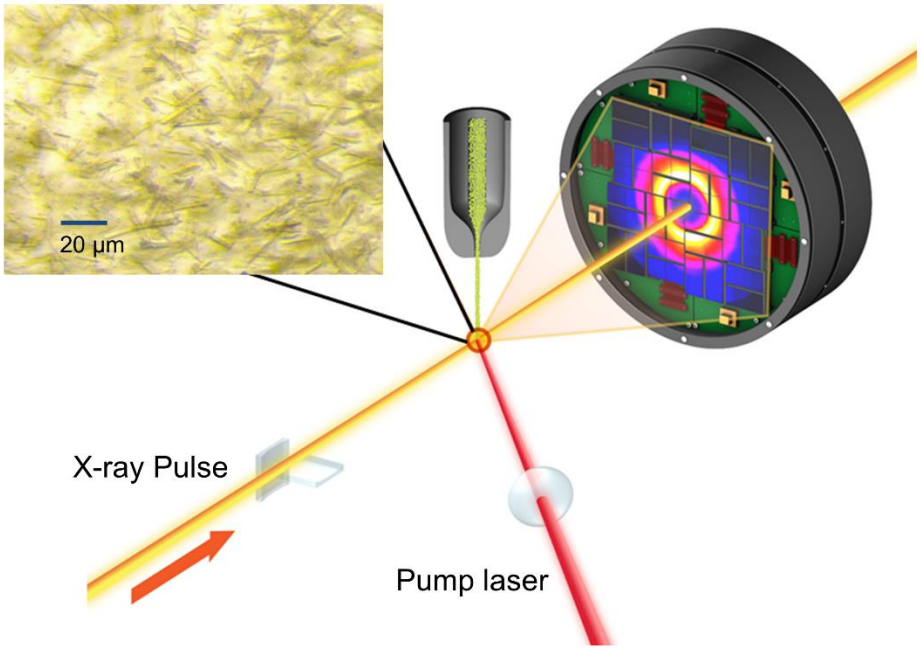


Figure 2.6: *Experimental set-up for a TR-SFX experiment. Micro-crystals are injected into the path of the aligned pump laser and X-ray beam. When X-rays pass through a micro-crystal, a diffraction snapshot will be recorded on the CSPAD detector.*

The time resolution offered by these experiments is limited both by the pulse length of the pump laser and the X-ray probe. A further challenge is to achieve high precision in timing through precise alignment of the X-rays and the pump-laser. The LCLS XFEL has a pulse length of < 50 fs, allowing much faster structural changes in proteins to be detected than at conventional synchrotron sources. To make use of this short pulse length, a femtosecond laser with a similarly short pulse length must also be used. For the experiments presented in **Papers II and V**, a Ti:Sa femtosecond laser with pulse length of 150 fs was utilized that set the hard limit on time-resolution possible.

Precision in timing was achieved by implementing a timing-tool developed at the LCLS[124]. It is possible to operate the Ti:Sa laser at 60 Hz, exactly half that of the LCLS X-rays. This allows collection of interleaved laser-on and laser-off images, helping to reduce systematic errors during data collection due to fluctuations in the flux of the XFEL beam that could present themselves when collecting laser-on and laser-off datasets at different times. With use of the femtosecond laser and the timing tool, precise time measurements can be made from the femtosecond to nanosecond domain. This broad range of available time-points can aid in the creation of molecular movies, following conformational changes from the ultrafast such as cis-trans isomerization to the more delayed structural changes that these events initiate.

2.4 Time resolved solution scattering at an XFEL

2.4.1 Solution scattering theory

Solution scattering is a low resolution X-ray diffraction method using solutions of mono-disperse protein samples in aqueous buffer. Crystallography remains a daunting task for many protein candidates and an advantage of protein solution scattering over X-ray crystallography is that crystalline samples are not needed. X-rays interfere differently with particles randomly orientated in solution than they do with ordered crystalline samples as Bragg's law is not satisfied. As opposed to observing intense Bragg reflections, X-ray diffraction from a protein solution results in a diffuse diffraction pattern that is made up of the sum of the X-ray scattering from all atom pairs in the solution. The scattering vector from an incident X-ray is defined as:

$$q = 4\pi sin(\theta)/\lambda = 2d \qquad (2)$$

Where q is the modulus of the scattering vector, and d is an interaction distance. Thus at higher values of theta, the interaction distance becomes smaller, and it is possible to extract information on the atomic scale. This is the basis of protein wide-angle X-ray scattering (WAXS) that measures scattering on a q-range from q = 0 $Å^{-1}$ to ~2.5 $Å^{-1}$, and

includes all structural information about the protein.

The formula for the total scattering from an incoming X-ray beam interacting with a protein solution is defined as:

$$F(\boldsymbol{q}) = \Sigma_j \left(f_j e^{i\boldsymbol{q}\cdot\boldsymbol{r}_j} \right) \qquad (3)$$

F(**q**) is the amplitude of a scattered wave as a function of the scattering vector **q**, as well as the atomic scattering factor (f_j) and position vector ($\mathbf{r}_j$) of atom j. The atomic scattering factor of an atom can be calculated using quantum chemical calculations. This total absolute scattering is made up of three components, the scattering from the protein molecules, the scattering from the bulk solvent molecules, and the scattering from the excluded volume that the protein occupies. Of these contributions, the scattering from the solvent molecules dominates and must be subtracted from the absolute scattering as background to retrieve any interpretable information about protein scattering.

The solutions measured in protein WAXS are made of many randomly orientated protein molecules which gives rise to symmetrical diffraction patterns that can be thought of as the spherical average of the scattering from all orientations of the protein and the solvent. This spherical averaging can be given by the Debye formula:

$$S(\boldsymbol{q}) = N\langle |F(\boldsymbol{q})|^2 \rangle = N\Sigma_i\Sigma_j (f_i - \rho_s \nu_i)(f_j - \rho_s \nu_j) \frac{sin(\boldsymbol{q}r_{ij})}{(\boldsymbol{q}r_{ij})} \qquad (4)$$

Where S(**q**) is the total scattering from vector **q**, N is the number of molecules in the system, $\mathbf{r}_{ij}$ is the distance between atoms i and j ($|\mathbf{r}_i - \mathbf{r}_j|$) and ($f_i - \rho_s \upsilon_i$) is the contrast amplitude between the bulk solvent and protein atom j. Using this formula, the scattering from a known atomic structure of a protein can be calculated. Unfortunately, coordinates for a protein cannot be calculated from the overall scattering as there can be many structural solutions to a given absolute scattering. This problem is overcome to some extent when trying to interpret collected WAXS data on a protein of known structure. By using molecular dynamics simulations and perturbing the known structure, one can back-calculate scattering data to fit experimental WAXS data.

2.4.2 Time-resolved pump-probe solution scattering at an XFEL

The majority of proteins undergo conformational changes to carry out their function, and these protein dynamics are of great interest to structural biologists who aim to better understand structure/function relationships. Conformational changes can be induced in

a number of different ways. Substrate binding, phosphorylation or oxidation can all induce large-scale structural perturbations in proteins. In this thesis however, experiments were performed on the light-activated protein RC_{vir}. By exciting this protein with a pump laser tuned to the wavelength where it can interact with the protein cofactors, the electron transport process can be induced. Through careful alignment of the XFEL laser and the pump laser, absolute scattering from the protein can be captured at precise time-delays following photo-excitation (see 2.3.2 for further experimental explanation). However, as Figure 2.7 demonstrates, the absolute scattering from photo-excited protein solution and 'dark' protein solution look remarkably similar. This is because scattering from the protein molecules only contributes a small proportion of the absolute scattering. Even in a concentrated protein solution, the number of protein atoms is vastly outnumbered by the number of bulk solvent atoms. Further scattering contributions are provided by the void volume where the protein molecules take up space as well as the detergent micelles surrounding the protein in the case of membrane proteins. Thus protein solution scattering experiments have very low sensitivity and to understand the effect of photo-excitation on the scattering of the protein, one must subtract the absolute scattering of the 'dark' solution from the absolute scattering of the photo-excited solution. Ideally this will remove contributions from the solvent which, in a perfect experimental set-up, would be unchanged by photo-excitation. This process gives rise to difference scattering data. Figure 2.7 shows absolute scattering curves from 'dark' protein solution and laser-pumped protein solution as well as the difference scattering curve they yield.

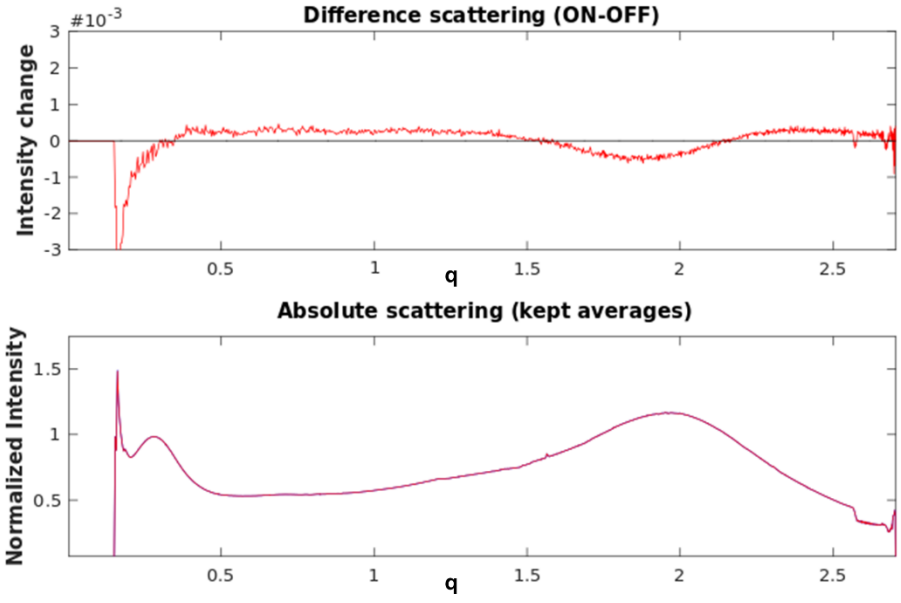


Figure 2.7: *Absolute scattering and the subsequent difference scattering from a 10 minute TR-WAXS run of visual rhodopsin at the LCLS XFEL. The absolute scattering from the 'dark' protein solution and the laser excited protein solution are overlaid and cannot be distinguished by eye.*

2.4.3 Interpretation of difference scattering data

In time-resolved protein solution scattering, the structural information available increases as q increases in the difference scattering curves (Figure 2.8). At very low angles, q < ~ 0.2, information about the overall shape of the protein can be obtained, such as the radius of gyration. Tertiary and secondary structural information such as the movement of α-helices is available at slightly wider angles, q ~ 0.2 to 0.6. At a higher q range, individual amino acid movements can be observed in principle, although owing to the low sensitivity of the experiment and the high intensity of scattering from the solvent at these angles, this is difficult to practically observe.
From a known protein structure, a theoretical absolute scattering can be calculated using the Debye formula. Thus, given two protein structures, a theoretical difference scattering can also be calculated. By using molecular dynamics simulations, it is possible to test a wide variety of protein movements in an attempt to fit theoretical difference scattering to experimental observations.

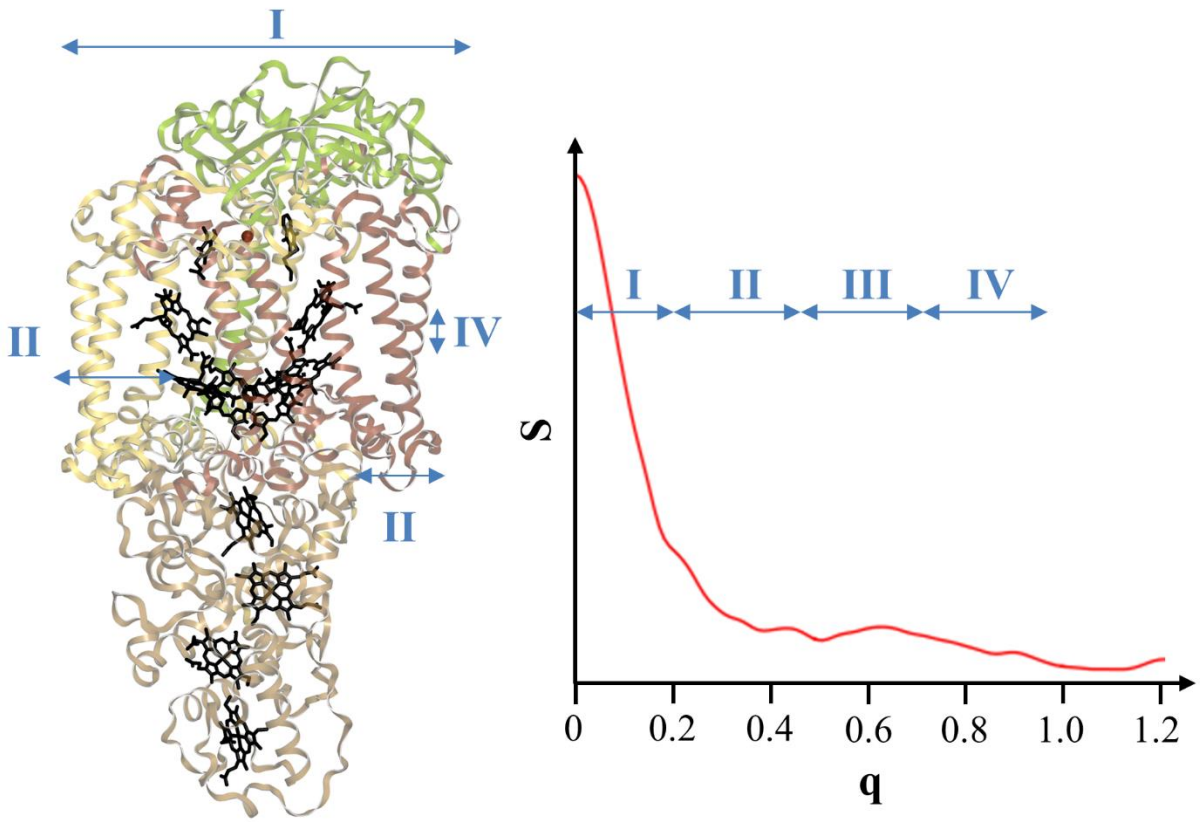


Figure 2.8: *Schematic demonstrating the scale of structural movements observable by TR-WAXS along with the corresponding q-regions on the scattering curve where this structural information is observed. Predicted scattering from RCvir created using CRYSOL*[125].

2.5 Time-resolved serial femtosecond crystallography at an XFEL

2.5.1 Converting XFEL intensities to structure factors for refinement

A TR-SFX experiment at an XFEL collects datasets of 'dark' crystals that have not seen photo-excitation and datasets of crystals at various time points after photo-excitation. These datasets are compiled of single diffraction snapshots of micro-crystals in random orientations. Bragg peaks are not infinitesimal points but have a 3D shape owing to the mosaicity inherent even in micro-crystals as well as the energy spread and divergence of the X-ray beam[126]. The intensities recorded in a single diffraction image are often partial reflections, as a partiality of the reflection passes through the surface of the Ewald sphere. Once integrated using conventional integration programs such as Mosflm[127] or Dirax[128] which determine cell constants and crystal orientations, the intensity measurements can be merged through a Monte Carlo integration method developed by Kirian *et al*[126]. This method merges the partial reflections by integrating over the 3D volume of the Bragg reflection. This gives an intensity output that shows good agreement with data collected at synchrotrons[76] but does not account for the partiality of the reflections. There are efforts in the SFX community to better account for these partialities[129-132]. While with specific data sets these have been shown to improve statistics, overall these calculations still deliver mixed results.

In the experiments described in this thesis, the partiality of reflections was not taken into account and the simple Monte Carlo integration was used. The software suite CrystFEL was used for processing diffraction images, giving as an output a list of symmetry-merged reflections along with their corresponding intensities and errors[133]. CrystFEL can find peaks in diffraction images based on their shape and intensity and calls the indexing programs Mosflm and Dirax to index these images. Given a lack of known space group and cell parameters, the indexing programs will index the diffraction 'snapshots' into any space group it can. The true crystal parameters can often be deduced from the histograms of cell parameters that this *ab initio* indexing outputs and the CrystFEL ambigator module can be used to solve indexing ambiguities. Once crystal parameters have been found, the data can be re-indexed using these. After indexing has been performed the reflection information is merged and scaled. The modulus of the structure factor amplitudes can be solved from given intensities by taking the square root[76].

To assess quality and resolution limits of the data CrystFEL gives information about the signal to noise ratio of the data (SNR), the correlation between two halves of the data ($CC_{1/2}$) and the agreement between two halves of the data (R_{split}). R_{split} is a statistic unique to XFEL data processing and is given by the formula:

$$R_{split} = 2^{-\frac{1}{2}} \frac{\Sigma|I_{even} - I_{odd}|}{\frac{1}{2}\Sigma(I_{even} + I_{odd})} \quad (5)$$

Where I_{even} is the intensity of a reflection from the even numbered diffraction patterns, and I_{odd} is the corresponding reflection from odd numbered patterns. The multiplication by factor $2^{-1/2}$ is to correct for splitting the data into two sets, which would make the data look worse than in reality.

2.5.2 Structure refinement

As demonstrated in Equation 6, the electron density of a protein can be calculated by performing a Fourier transform on the structure factors of the protein.

$$\rho(x, y, z) = \frac{1}{v}\Sigma_h\Sigma_k\Sigma_l|F_{hkl}|e^{-2\pi i(hx+ky+lz)+i\alpha(hkl)} \quad (6)$$

X-ray crystallography experiments measure the intensities of reflected X-rays and these intensities are proportional to the square of the structure factor amplitude. However the phases of the structure factors cannot be directly measured. This has become known as the phase problem of crystallography. There are several solutions to the phase problem. For small molecules, *ab initio* phasing can be performed in a number of different ways which generally rely on probability theory and the relationship between phase angles in a structure[134]. For large macromolecules however, this is currently unachievable. For proteins where a previous structure or a close homologue is already deposited in the PDB, the phases can be taken from these using the molecular replacement method. For RC_{vir}, over 20 structures have been deposited in the PDB so this is the favored approach. Other solutions to the phase problem can be found in protein crystallography by using isomorphous replacement, MAD or SAD phasing. MAD and SAD phasing take advantage of the anomalous scattering of heavy atoms in a protein to discover phase angles of protein reflections. SAD phasing using the anomalous scattering of native sulfur atoms has been used to solve the structure of lysozyme at an XFEL[135]. Isomorphous replacement is performed by soaking crystals in solutions of heavy atoms. The crystals can incorporate these heavy atoms and then the anomalous scattering of these atoms can be used to solve the phase problem.

Once phases have been calculated, they can be combined with structure factor amplitudes and the electron density of the protein can be calculated. The program Phaser can use phases calculated from a known structure along with translation and rotation functions to fit a structure of the model protein into the electron density[136]. The phases contribute more than the amplitudes to the overall electron density and to reduce model

bias, electron density maps are calculated by subtracting the calculated structure factors from two times the observed structure factors, what is known as a $2mF_{obs}$-DF_{calc} map. The model structure can then be manually built in programs such as COOT[137] and iterative refinement can be performed with programs such as Phenix[138] or Refmac[139], until the observed electron density and that calculated from the model converge.

Various statistics measure the success of molecular replacement and structure refinement, two of the most widely cited being R_{work} and R_{free}. R_{work} is a measure of how well the refined model fits the experimental data and is given by the formula:

$$R_{work} = \frac{\Sigma||F_{obs}|-|F_{calc}||}{\Sigma|F_{obs}|} \qquad (7)$$

A theoretical R_{work} value of zero implies a perfect fit between experimental data and structure factors calculated from the model, although this is never actually achieved in a protein crystallography experiment. R_{free} is another measure calculated by exactly the same formula as above but on a small subset of the data made up of random reflections that are never included in the refinement. Usually 5 % of reflections are excluded in the refinement and used to calculate this measure. If the value of R_{free} is much greater than the value of R_{work} it implies that overfitting to the data has occurred.

2.5.3 Analysis of electron density difference maps

Once a structure of the protein in its native state has been refined, differences between the native data set and data sets of excited states can be examined. This is achieved through the creation of an electron density difference map. Conformational changes within the confines of a crystal are by usually relatively small as large movements of whole domains will disrupt the crystal lattice and destroy its diffraction power. It is therefore assumed that the phases change little between the native and excited states. To create electron difference density maps, the phases of the native state are used and the structure factors of the dark state are subtracted from the structure factors of the excited state, creating an F_{obs}-F_{obs} map. To improve signal to noise in the electron density difference maps presented in **Paper V**, a Bayesian method is used to apply a weighting term to the phases[118] using the software CNS[140]. The F_{obs}-F_{obs} map produced reveals regions of electron density that have disappeared from the native state and appeared in the light state and can be interpreted as conformational change between the states. Laser absorption by crystals is generally not completely efficient and uniform; only a proportion of the protein molecules in the crystal will be excited and undergo conformational changes. Therefore, in order to model the structural changes visualized

in the electron density difference map, partial occupancy refinement must be carried out. To perform a partial occupancy refinement of an excited state, alternate conformations are first built of residues in the refined ground state around strong electron difference density peaks. The two conformations are then assigned a percentage occupancy. A refinement is carried out on this structure against the observations from the excited state, with the ground state conformations locked in place, allowing refinement of only the light-activated conformational changes.

3. Ultrafast solution scattering study of photosynthetic reaction center – Paper II

Previous work by our group probed the structural changes in RC_{vir} in solution undergoing multi-photon absorption of XFEL radiation[115]. This work used TR-WAXS analysis along with molecular dynamics simulations to demonstrate that the oscillations observed in WAXS difference scattering curves could be explained by quake-like outward movement of the helices surrounding the monomeric chlorophyll cofactors. These global protein motions arose with a half-maximum of 1.4 ps and decayed with a half-life of 44 ps, peaking at 7 ps. In May 2014, another experiment was carried out at the CXI beamline[121] at the LCLS. The aim of the experiment was to explore if the structural changes observed in a multi-photon absorption event were relevant under more physiological conditions. To this end, a power titration was carried out and oscillations in TR-WAXS difference scattering curves were tracked down to a laser power correseponding to 0.8 photons absorbed per RC_{vir} molecule. In this respect, the work of Levantino *et al*. on myoglobin is relevant[116]. Using TR-WAXS at an XFEL they observed extended motions on picosecond timescales using a laser power equivalent to 1.6 photons absorbed per myoglobin chromophore.

3.1 Pump-probe solution scattering of reaction center at an XFEL

30 $mgml^{-1}$ RC_{vir}, detergent solubilized in a low-salt buffer solution, was injected using a container-less microjet into the path of the XFEL laser using the GDVN system[79]. The protein solution was pumped with an 150 fs pulse-length Ti:Sa 800 nm femtosecond laser with a repetition rate of 60 Hz and a spot size of 221 FWHM. It was then probed with XFEL radiation. 6 keV X-ray pulses were used with 42 fs pulse duration and a repetition rate of 120 Hz, focused to a 10 μm^2 spot. This produced a dataset of interleaved laser-on/laser-off images. The XFEL laser and the pump laser were aligned on the microjet and precise timing for the interaction was achieved using the timing tool[124] at the CXI beamline of the LCLS along with analysis of the heating signal in difference scattering curves[115]. The absolute scattering was recorded using the CSPAD[77] detector. The images were sorted into light and dark datasets using the software CHEETAH[123] by interpreting the readout from a diode within the hutch triggered by the femtosecond laser. This software then radially integrated the intensities recorded by the detector for each shot and saved the data for each run as a radial stack. Matlab was used for further data manipulation. The flux of the X-ray beam at an XFEL can fluctuate over time and by collecting data in this interleaved on/off fashion, systematic errors between laser-on and laser-off datasets can be greatly reduced.

3.2 Processing of difference scattering curves

As explained in section 2.4.2, TR-WAXS is a method that is sensitive to noise, owing to the very small differences in absolute scattering. Moreover, whereas in TR-SFX experiments it is relatively simple to discard detector read outs that do not show crystal diffraction, in TR-WAXS experiments every detector readout is regarded as either a laser-on or laser-off 'hit'. Thus the data collection method is blind to problems in the microjet and this increases overall noise in the dataset. The ability to collect large datasets with the high repetition rate of the XFEL alleviates this problem to some extent, but it was also important to develop strong data rejection and processing techniques specific for time-resolved XFEL experiments to objectively select the best quality data. Fluctuations in X-ray flux, the detector background subtraction as well as the behavior of the microjet can all affect the total scattering recorded on the detector. This can be demonstrated by plotting the sum of the scattering intensity between q = 0.2 to q = 2.4 $Å^{-1}$ for each individual measurement against the ratio in each measurement between the peak in water scattering at q = 2.0 $Å^{-1}$ and a low-scattering plateau found at q = 0.8 $Å^{-1}$. This plot is displayed as a scatter diagram in Figure 3.1**.** Figure 3.1A shows the density of plots on the scatter diagram, and by rejecting all data that fell outside this densest region (outside the ellipse shown in Figure 3.1B), the signal to noise ratio was improved.

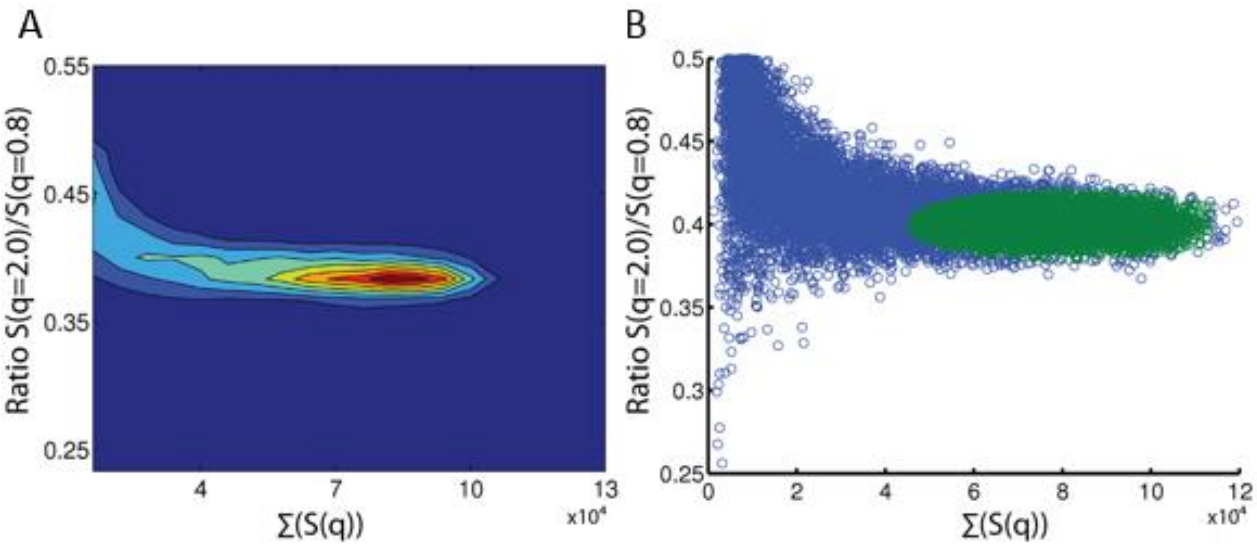


Figure 3.1: *Variations in scattering intensity due to pulse to pulse fluctuations in XFEL flux.* ***A)*** *A contour plot showing the sum of the scattering across the whole q range (0.2 – 2.4) against the ratio of the scattering between the solvent peak (q = 2.0) and a low scattering region (q = 0.8).* ***B)*** *A scatter diagram plotting the sum of all scattering against the ratio of scattering between the solvent peak and a low scattering region. The data within the ellipsoid representing the peak of the contour plot is retained, while all other data is rejected.*

To calculate difference scattering curves from individual runs, data were normalized around the isosbestic point for water heating in X-ray scattering (q = 2.05 to 2.15 $Å^{-1}$) in order to better compare protein scattering between images. Then from each laser-on image, the average of three closest laser-off images on the scatter plot (Figure 3.1B) was subtracted. These individual difference scattering curves were then averaged together to produce the overall difference scattering curve for a given dataset. A slightly simpler data rejection script was implemented 'on-the-fly' at the beam line and allowed feedback from the experiment within around 20 minutes, which greatly aided decision making on-site.

3.3 Power titration

To follow structural changes from a multi-photon absorption setup to a single-photon absorption event, a power titration was carried out, gradually decreasing the pump laser power from 100 µJ down to 1.7 µJ in a step-wise fashion. For this experiment a time point of 5 ps was chosen as this was a timescale dominated by changes in protein scattering as opposed to difference scattering due to solvent heating[115]. In the higher laser powers used, as multiple photons are absorbed per protein molecule, there is a significant transfer of heat after a time delay into the solvent, which is represented in the X-ray scattering by a decrease in the scattering intensity from the solvent, observed as a valley in the scattering at q ~ 2.0. This temperature increase was as much as 4 K at a 100 µJ laser power. As the laser power was lowered, the heating of the sample was reduced. The temperature increase calculated from scattering curves is plotted against the laser power in Figure 3.2. This figure demonstrates that while at higher energies the reduction in temperature jump is quadratic with laser power, below a laser power of

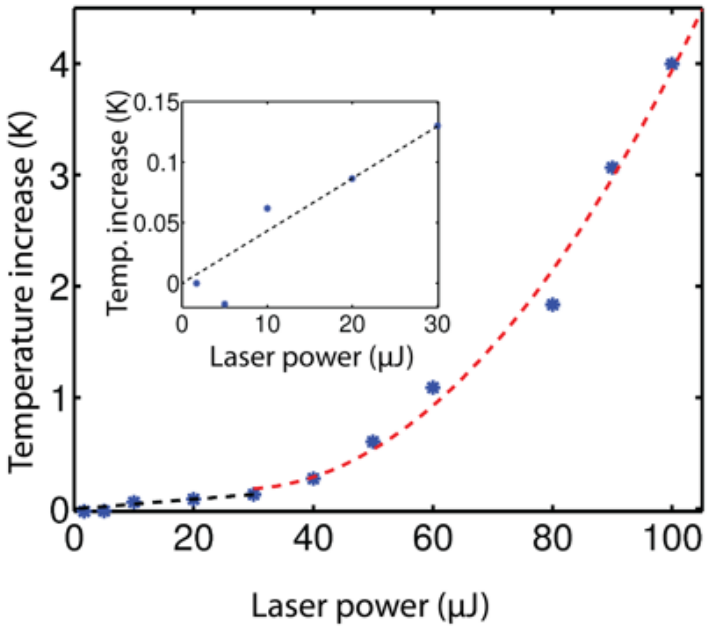


Figure 3.2: *Plot showing the laser power used in a pump-probe TR-WAXS experiment against the increase in sample temperature as calculated from the solvent heating signal. At high laser powers (> 30 µJ) the temperature increase changes quadratically, whereas at low laser powers (< 30 µJ, enlarged in inset) the temperature increase is linear.*

30 μJ the reduction in temperature becomes linear with the laser power. The temperature increase calculated for a laser power of 1.7 μJ (corresponding to 0.8 photons absorbed per protein molecule) is $7e10^{-3}$ K, and any structural changes induced by photon absorption at this laser power are therefore unlikely to be caused by the heating of the sample.

3.4 Effects of single photon absorption on RC_{vir} structure

Figure 3.3 shows the difference scattering curves resulting from the power titration after applying a high-frequency filter to smooth the curves. The dominant features in the curve for the highest laser power (100 μJ, 460 photons per RC_{vir} molecule) is the peak at q ~ 0.4 and the double peak at q ~ 0.6. These features reproduce those seen in previously collected data that could be modeled with molecular dynamics simulations in the protein quake article published in 2014[115]. As the laser power is reduced to 20 μJ (10 photons per molecule), the structure of the dominant features appears to change, the peak at q ~ 0.4 splits into two distinct peaks, while the peaks around q ~ 0.6 change in size relative to one another and appear to shift slightly in q. At a laser power of 1.7 μJ, corresponding to absorption of 0.8 photons per RC_{vir} molecule, there are still oscillatory structural features present in the difference curve although they are much weaker and do not so closely reflect the features seen at higher laser power. These features in the difference curve at low laser power cannot be reproduced simply by extrapolating the movements described in the protein quake paper and a combination of further molecular dynamics simulations as well as comparison with RC_{vir} TR-SFX data is needed to fully understand the ultrafast structural changes that occur in reaction center upon single-photon absorption.

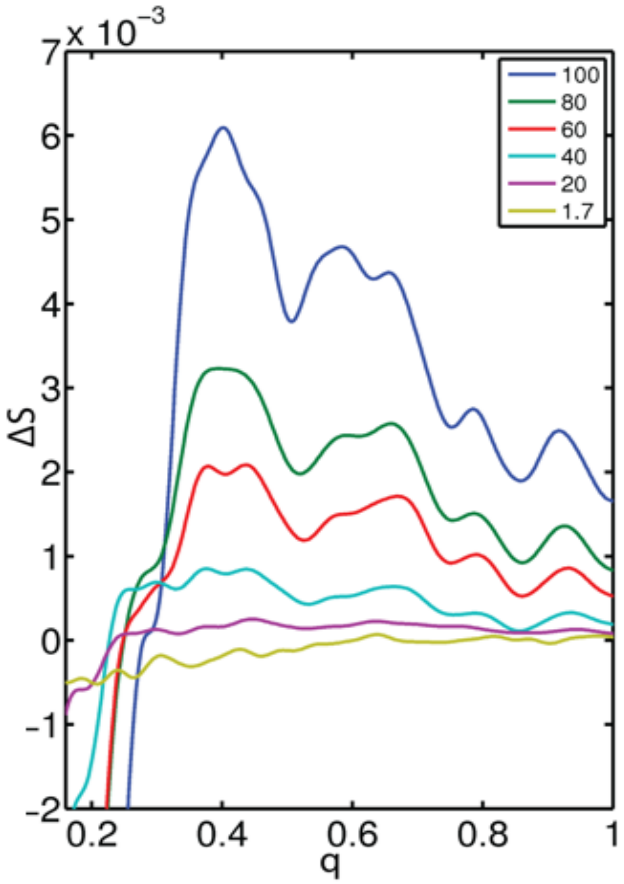


Figure 3.3: *Difference scattering curves observed during a power titration in a pump-probe TR-WAXS experiment on RC_{vir}. The X-rays were aligned to arrive 5 ps after a 800 nm laser pulse at various powers (1.7 - 100 μJ) and the scattering curves were calculated by subtracting the laser-on data from the laser-off data and applying a smoothing function. At high laser powers, there are two dominant features at q ~ 0.4 and q ~ 0.6, representing conformational changes of the protein. As laser power is decreased the intensity of these features is decreased and sub-structures are revealed. At 1.7 μJ, structural oscillations remain but are much dampened.*

3.5 Paper II summary

Solutions of concentrated (30 $mgml^{-1}$) RC_{vir} were produced and a time-resolved wide angle X-ray experiment was performed at the CXI beamline at the LCLS in May 2014. Samples were pumped with a 800 nm femtosecond laser with a repetition rate of 60 Hz and probed with 6 keV X-rays of a 42 fs pulse duration focused to a 10 μm^2 spot at a repetition rate of 120 Hz. This resulted in a dataset of alternating laser-off/laser-on images. Data was recorded in a q-range of 0 - 2.5, encompassing low to medium resolution structural information, and capturing scattering caused by the aqueous solvent from which the temperature increase caused by the laser could be calculated. Constructing difference scattering data from interleaved images on/off images was useful in eliminating noise that results from fluctuations in the XFEL energy from shot to shot and drifts in the overall experiment.

Data from previous work carried out by the Neutze group on RC_{vir} demonstrated the time-dependent evolution of a 'protein quake' of structural movements following photon absorption on the picosecond timescale[115]. This work aimed to demonstrate that these changes were relevant in a more physiological system with single photon absorption, while also building on the methodology used to process time-resolved solution

scattering data from an XFEL.

A power titration was carried out by pumping the sample at various laser powers in a stepwise fashion from 100 μJ down to 1.7 μJ whereby the lowest power represented 0.8 photons being absorbed per molecule of RC_{vir}. Noise levels were reduced in the data by introducing a novel data rejection technique. Further noise-reduction was carried out using median filtering to smoothen the difference scattering curves.
Difference scattering data at higher laser powers closely reproduced data reported by Arnlund *et al*., with major peaks at q ~ 0.4 and q ~ 0.6. Features that could be reproduced through molecular dynamics simulations that described a 'protein quake'[115]. Although the sub-structure of these features changed at lower laser powers, they could still be identified even in single-photon absorption events.
To truly understand whether the features identified in difference scattering curves at the lowest laser power accurately represent the structural movements identified as the protein quake, work is being continued to replicate these movements using time-resolved serial femtosecond crystallography. The submission of this work has therefore been delayed to be combined with TR-SFX work at two laser powers using an 800 nm laser pump. This is turn is awaiting submission of the TR-SFX work using a 960 nm laser pump presented in **Paper V**.

4. Strategies for micro-crystallization for XFEL studies – Paper III

XFEL technology provides new opportunities for structural biologists with its powerful X-rays, high repetition rate and short pulse duration. However as detailed in section 1.5.2, a single pulse from an XFEL will deliver a dose well above the accepted dose at a synchrotron and will lead to the rapid destruction of protein crystal samples. It is therefore a necessity to deliver into the X-ray beam a stream of well-ordered crystals at high concentration, which must measure less than roughly 20 µm in all directions in order to avoid clogging the sample delivery system. Crystallization efforts of years gone by have focused on producing large crystals, as a large well-ordered crystal will usually diffract to higher resolution than smaller crystals. Thus the scientific literature for producing large quantities of micro-crystals is thin on the ground, although some articles in this field are now being published[16,141-143]. The aim of **Paper III** was to expand this literature and provide general strategies for micro-crystallization while also comparing an RC_{vir} structure obtained at the XFEL to synchrotron structures deposited in the PDB.

4.1 Development of micro-crystals

Development of micro-crystals occurred in three stages across three experiments at the LCLS, with an overall improvement of crystal resolution from 3.4 Å to 2.4 Å. The first experiment was performed on macro-crystals of RC_{vir} that had been mechanically crushed to yield a slurry of crystal shards measuring less than 20 µm in all dimensions. The second experiment used crystals grown from concentrated seed stocks while the third experiment used crystals that had been grown with two rounds of seeding.

Previous SFX experiments on RC_{vir} had used crystals grown in the lipidic sponge phase[15]. These crystals diffracted to 3.4 Å and yet had several drawbacks. The first was a low concentration that was difficult to optimize, this led to a low crystal hit rate of less than 1 %. The second drawback was a long c-axis that resulted in spots that were not well spatially resolved on the detector, leading to a relatively low indexing rate (25 %). To address these problems crystals were grown by a vapor diffusion method adapted from the original method of RC_{vir} crystallization[12]. Crystals grown by vapor diffusion have the advantage that, given they are sufficiently robust, they can be easily concentrated by low speed centrifugation followed by removal of supernatant. The crystals produced by this method were large (roughly 50 x 50 x 250 μm^3) and thus were unsuitable for direct delivery into the XFEL beam (Figure 4.1A). Thus they were mechanically crushed by vortexing using seed beads (Molecular Dimensions). The crystals were periodically examined under a light microscope throughout the crushing

process until they were found to be largely devoid of large crystal pieces greater than 20 μm in all dimensions (Figure 4.1B). This microcrystal slurry was filtered through a 20 μm cutoff filter to ensure removal of any remaining macro-crystals and then delivered into the beam. They were found to diffract to 3.4 Å and could be indexed into a unit cell and space group consistent with other vapor diffusion-grown RC_{vir} crystals. Although some crystal damage was observable in diffraction patterns as smeared Braggs peaks due to mosaicity, the data set was processed with reasonable statistics and a 3.4 Å structure could be refined. This method proved valuable in assessing this crystal forms viability in SFX experiments and a method to produce micro-crystals of the same form was pursued with renewed vigor. The method of crushing crystals in this manner has been replicated in **Paper IV** on phytochrome crystals at SACLA. In this experiment, macro-crystals of phytochrome were crushed by vortexing with seed beads and the diffraction data could be processed to 2.2 Å resolution[144].

Micro-crystal growth was achieved using a micro-seeding method. This method involves initial growth of macro-crystals followed by crushing of these crystals into small pieces of roughly 5 μm or less[145]. The seeding solution can then be spiked into new crystallization set ups which introduces a large number of nucleation sites and encourages crystal growth[146]. Seeding has proved a versatile method and was used for successful growth of ribosome crystals[147,148]. It has also been shown that seeding is a viable technique for crystallization of proteins different to the seeds used[149-152]. Most seeding experiments use highly diluted seed stocks to limit the number of nucleation sites[153], but as the aim in this experiment was to generate showers of micro-crystals, undiluted seed stocks were used. To create seed stocks, RC_{vir} macro-crystals grown by vapor diffusion were crushed using seed beads. Seed stock was then spiked into a new sitting-drop vapor diffusion set up with slightly altered conditions. The protein concentration was slightly lowered, and the volume of the precipitant was also lowered. The idea behind altering crystallization conditions when using seeding methods is to move the protein into the metastable phase, where crystal growth on the many introduced nucleation sites can begin immediately[154]. Under these conditions, micro-crystals grew to lengths averaging around 10 μm (Figure 4.1C). These crystals could be harvested from the sitting drops and concentrated in a similar manner to the crushed macro-crystals. Owing to the lack of crystal damage and the higher degree of homogeneity in crystal size, these crystals diffracted better than crushed macro-crystals, and the data could be processed to 2.8 Å.

To further improve the diffraction quality of RC_{vir} crystals, a second round of seeding was introduced. This has previously been shown to give an improvement in crystal morphology[150]. To achieve this, micro-crystals were grown as before and they were then mechanically crushed into a new seed stock. Micro-crystals appeared to break up

differently to macro-crystals, and instead of producing a slurry of small crystal shards that could not be easily individually identified, the rod-like crystals appeared to break once or twice, leaving a seed stock that was still readily observable as micro-crystals. These were spiked into new sitting-drop vapor diffusion set ups with the same conditions as the original micro-crystal growth. These set ups produced micro-crystal samples of much greater homogeneity with all crystals measuring about 20 µm in size (Figure 4.1D). The data collected using these crystals could be refined to 2.4 Å and the features of the structure obtained are further discussed in 4.3.

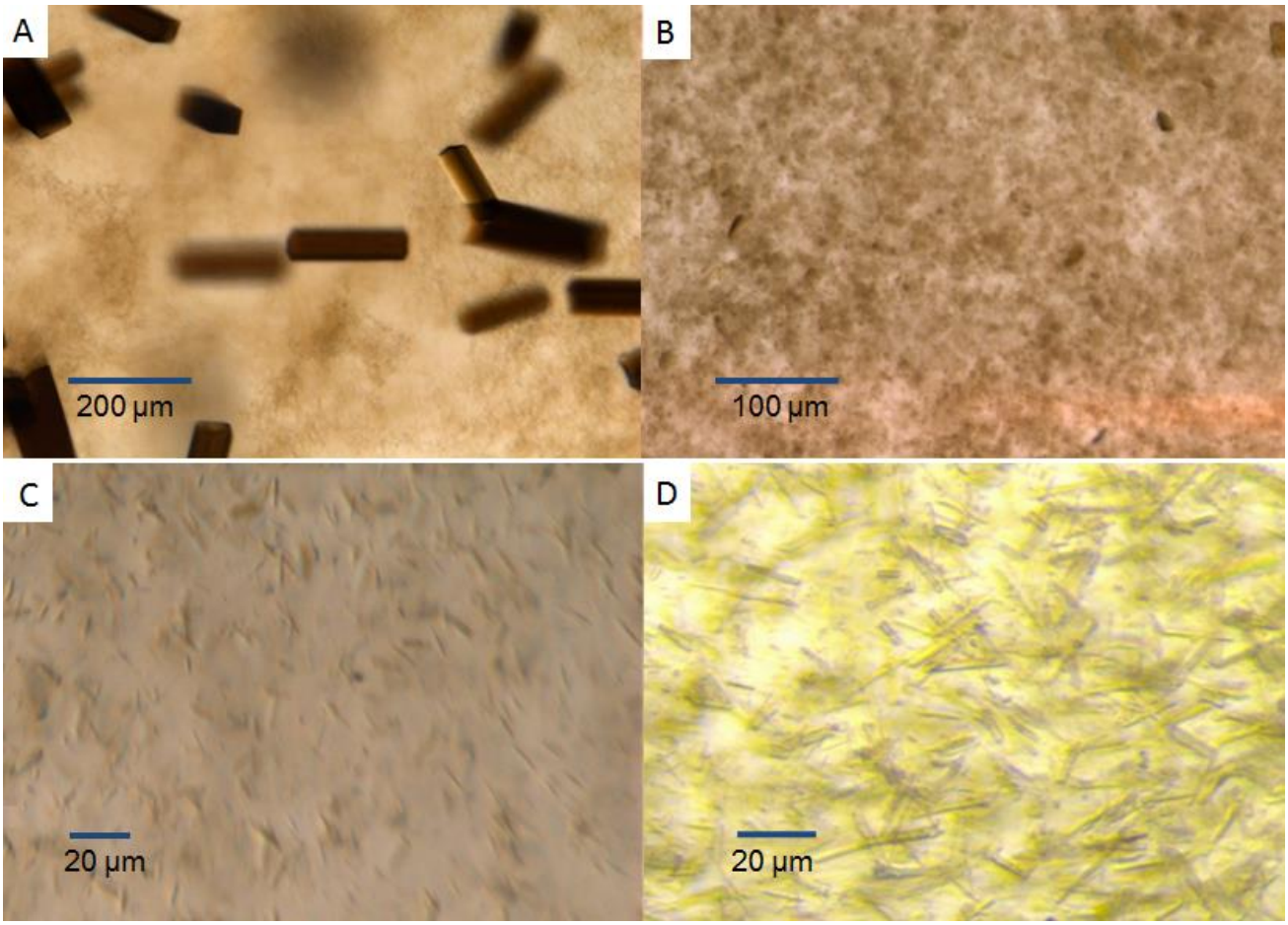


Figure 4.1: *The development of* RC_{vir} *micro-crystals for XFEL SFX experiments in sitting drop vapor diffusion experiments.* ***A)*** *Macro-crystals grown by following a protocol adapted from Michel* et al.[13] ***B)*** *Macro-crystals that have been crushed by vortexing, shown to be devoid of large crystalline pieces. These crystals were used in an SFX experiment and the data was refined to 3.4 Å resolution.* ***C)*** *Micro-crystals grown after one round of seeding measuring between 10-20 µm, giving diffraction data to 2.8 Å.* ***D)*** *Micro-crystals grown after two rounds of seeding giving a more homogeneous crystal suspension with all crystals close to 20 µm. These crystals diffracted to 2.4 Å.*

4.2 Data collection and refinement

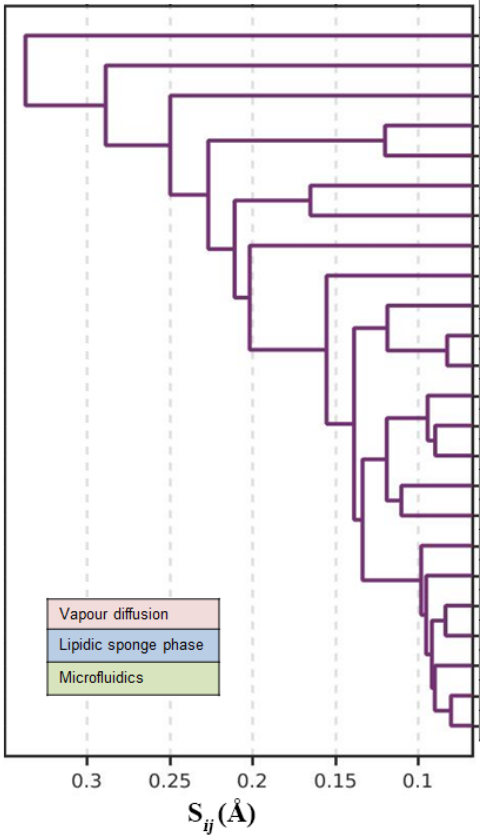
Data were collected from crushed crystals, micro-crystals after one round of seeding, and micro-crystals after two rounds of seeding, across three different experiments at the CXI beamline[121] at the LCLS. All experiments were carried out under a similar experimental set up. Crystals were first concentrated by low speed centrifugation (1000 *g* for 1 min) followed by removal of the supernatant. Crystals were generally concentrated threefold giving an average hit-rate of around 6 %. Further concentration was possible but resulted in increased clogging of the jetting apparatus. Samples were filtered through a 20 µm stainless steel filter (VICI AG International) to remove any fibers or large crystals and transferred to a reservoir placed in an anti-settling device[122]. Crystals were delivered through the GDVN sample delivery system[79] into the X-ray beam and the X-rays were aligned to interact with the jet before Rayleigh breakup into droplets. Diffraction intensities were recorded on the CSPAD[77]. X-ray wavelengths and detector distances differed between experiments and precise values are given in Paper II. Raw data were converted into HDF5 format and crystal diffraction was separated from empty frames using CHEETAH[123]. Integration, merging and scaling of data was performed by CrystFEL[133] with the partialator[155] module used to scale data but not refine partialities. The CrystFEL module geoptimiser[156] was used to refine the geometry of the CSPAD[156]. CrystFEL indexed the patterns using Mosflm[127] and Dirax[128] into the space group $P4_32_12$, the same as other wild-type RC_{vir} structures grown by vapor diffusion. The unit cell of RC_{vir} crystals in all three XFEL experiments is slightly elongated in all axes compared to synchrotron structures, probably due to the fact the experiment is carried out at room temperature.

The reflection intensity data compiled by CrystFEL was passed to CCP4 for structural refinement. Structure factor amplitudes were calculated using the program Truncate[157]. Phases were obtained by molecular replacement in the program Phaser[136] using an RC_{vir} structure from the PDB, the 1.9 Å structure with PDB code 2I5N. Restrained refinement was carried out using multiple rounds of 10 cycle refinement in Refmac[139] with manual modeling using Coot[137] in between. The validity of the final structure was assessed using Molprobity[158] and Procheck[159]. Seven Ramachandran outliers were found, with 2.4 % in the allowed region and the remaining amino acids in the favored region. The final R_{work} value was 16.1 % and the R_{free} value was 18.6 %

4.3 Comparison of XFEL structure with deposited RC_{vir} structures

Since the structure of RC_{vir} was first solved, more than 20 different structures have been deposited to the PDB. The 2.4 Å structure collected at the LCLS has a similar data quality to that found across other structures deposited in the PDB in terms of resolution

and R factors. It also shows a good agreement structurally with the bulk of deposited structures. This is demonstrated in Figure 4.2 where the structures have been clustered based on the similarity of their Cα internal distance matrix using tools developed by Wickstrand *et al*. **(Paper VIII)**. The new RC_{vir} structure clusters closely with all other structures of the same space group, and indeed it is the space group that seems to most strongly determine where an RC_{vir} structure will cluster.



PDB ID	Year	Res (Å)	Collaboration	Xray Source	Temp (K)	Space Group
1R2C	2003	2.86	Norris	Synchotron	293	P 43 21 2
4CAS	2013	3.50	Neutze	XFEL	293	P 21 21 21
4AC5	2011	8.20	Neutze	XFEL	263	P 21 21 21
2X5V	2010	3.00	Neutze	Synchotron	293	P 21 21 2
2X5U	2010	3.00	Neutze	Synchotron	293	P 21 21 2
2WJN	2009	1.86	Neutze	Synchotron	100	P 21 21 2
2WJM	2009	1.95	Neutze	Synchotron	100	P 21 21 2
3D38	2008	3.21	Ismagilov	Synchotron	100	P 31 2 1
1PRC	1988	2.30	Michel	Rot. Anode	271	P 43 21 2
RD14	2014	3.40	Neutze	XFEL	293	P 43 21 2
RD15	2015	2.80	Neutze	XFEL	293	P 43 21 2
RD16	2016	2.40	Neutze	XFEL	293	P 43 21 2
3T6D	2011	1.95	Cogdell	Synchotron	100	P 43 21 2
3T6E	2011	1.92	Cogdell	Synchotron	100	P 43 21 2
2I5N	2006	1.96	Ismagilov	Synchotron	100	P 43 21 2
3G7F	2009	2.50	Norris	Synchotron	100	P 43 21 2
1VRN	2005	2.20	Norris	Synchotron	100	P 43 21 2
2PRC	1997	2.45	Michel	Synchotron	263	P 43 21 2
2JBL	2006	2.40	Ditchfield	Synchotron	263	P 43 21 2
3PRC	1997	2.40	Michel	Synchotron	263	P 43 21 2
1DXR	2000	2.00	Michel	Synchotron	263	P 43 21 2
7PRC	1997	2.65	Michel	Synchotron	263	P 43 21 2
6PRC	1997	2.30	Michel	Synchotron	263	P 43 21 2
5PRC	1997	2.35	Michel.	Synchotron	263	P 43 21 2

Figure 4.2*: Tree diagram analyzing structural differences across all deposited* B. viridis *reaction center structures. Clustering is performed by comparing the average of the internal difference Cα distance matrix. Our deposited structure clusters closely with synchrotron structures of the same space group and not with other room temperature structures. RD14 refers to the structure obtained from crushed micro-crystals (not deposited). RD15 refers to the structure obtained from seeded micro-crystals (not deposited). RD16 refers to crystals obtained after two rounds of seeding (PDB ID 5NJ4).*

The one statistic that differs drastically between the high resolution XFEL structure and synchrotron structures is the multiplicity of the observed intensities. The overall multiplicity of the XFEL dataset was over 1500 (106 in the highest resolution shell), this compares to an average multiplicity of around 5 across all synchrotron structures. This large multiplicity arises due the nature of serial crystallography, where many

thousands of diffraction images are required for successful integration and merging of the data[133], and due to the high repetition rate of the LCLS XFEL at which 432,000 images can be collected in one hour. This has a profound effect on the quality of the electron density data, particularly in areas of high flexibility of the protein such as that associated with lipid molecules as well as in a flexible loop region of the protein on the H subunit.

In Figure 4.3, observable lipids are compared between the XFEL structure and other deposited RC_{vir} structures of the same space group. The XFEL structure showed strong electron density for seven LDAO molecules associated with the protein, five of which are also modelled in most synchrotron structures. Two new LDAO molecules were also observed with strong density associated with the L and H subunits. Interestingly, one lipid molecule observed in most synchrotron structures running parallel to the transmembrane helices of the M subunit was not observed in our XFEL data with sufficient electron density to support modeling. The reasons for this are unknown as it is modeled in the only synchrotron-based room temperature structure that the XFEL structure was compared against (1R2C)[25].

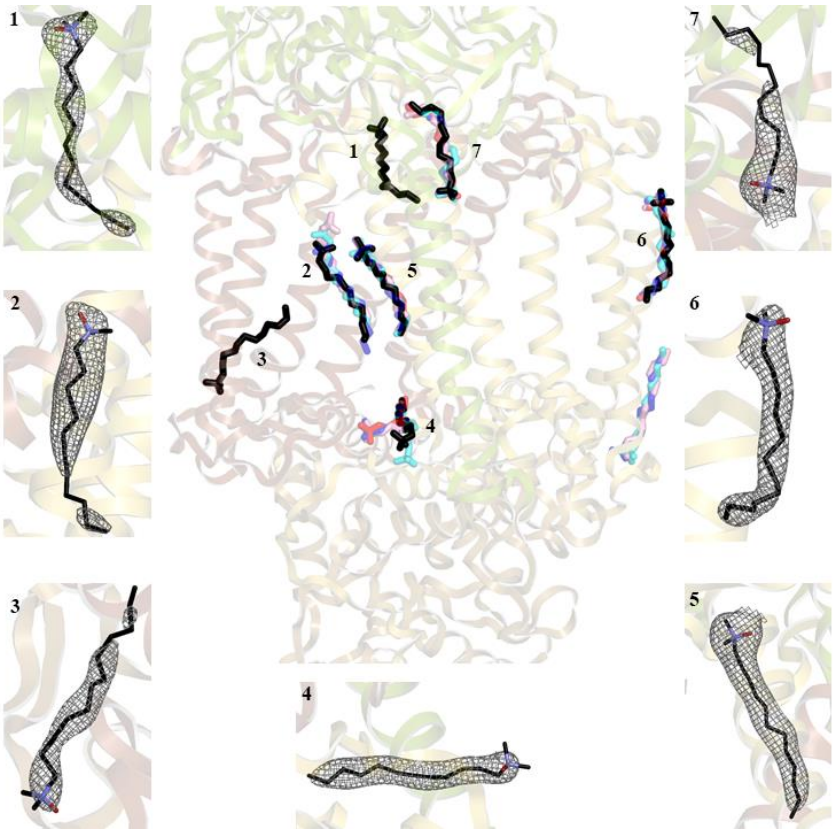


Figure 4.3: *Modelling lipids in the XFEL RC_{vir} structure (black) compared to lipids modelled in a selection of deposited synchrotron structures; 1DXR (blue), 6PRC (pink), 1R2C (cyan) and 2I5N (red). Electron density maps are displayed at 1 σ.*

The flexible loop region of RC_{vir} on the H subunit between Glu45 and Pro54 has not previously been modeled with any occupancy. This is because poorly ordered regions of proteins adopt different conformations in every protein molecule within the crystal, and in synchrotron experiments of low multiplicity, this results in localized regions of poor electron density. Upon initial structural refinement of the high multiplicity XFEL structure, it was noticed that this region gave a strong peak in positive electron difference density in the F_{obs}-F_{calc} map. This is demonstrated in Figure 4.4B where it is compared to the electron density of maps produced from structure factors deposited with the RC_{vir} structure 2I5N. The higher resolution structure (1.9 Å) was used as a model during molecular replacement. Figure 4.4A shows that there is very little evidence in the electron density for this loop region in the synchotron data when comparing it to the XFEL structure. These maps are contoured at 3 σ. After modeling of the main chain of the loop region, a continuous electron density was produced in the $2mF_{obs}$-DF_{calc} map when contouring at 1 σ, as displayed in Figure 4.4C.

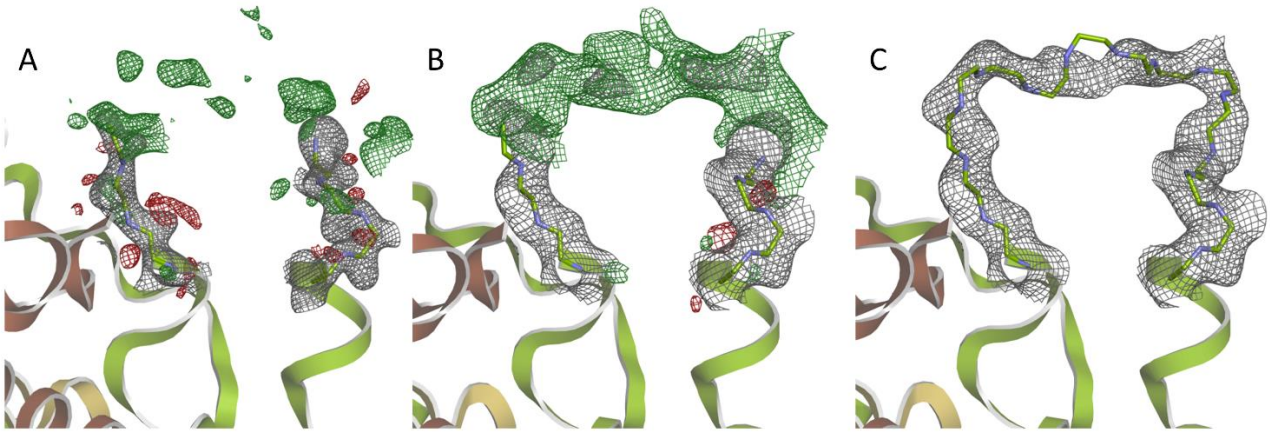


Figure 4.4: *Electron density maps surrounding the flexible loop region ($E45_H$ - $P54_H$) of RCvir comparing synchrotron data to XFEL data. $2mF_{obs}$-DF_{calc} maps are displayed in grey at 1 σ; F_{obs}-F_{calc} maps are displayed in red and green at 3 σ, red indicates negative electron density and green indicates positive electron density.* ***A)*** *Electron density maps for the 1.9 Å RC_{vir} structure 2I5N used as the model in molecular replacement.* ***B)*** *Electron density maps for the XFEL structure before the loop was modelled.* ***C)*** *Electron density map for the XFEL structure after the main chain of the loop was modelled.*

4.4 Paper Summary

Across three successful experiments at the CXI beamline at the LCLS, the resolution of diffracting micro-crystals was improved from 3.4 Å to 2.4 Å. This was achieved with the development of a micro-crystal growth strategy involving two rounds of seeding. The initial method of crushing macro-crystals, while giving lower quality data, was

successful in proving the feasibility of a crystal form for TR-SFX experiments and has shown similar success in related experiments on phytochrome (**Paper IV**). It is a versatile method that can be used to screen various crystal forms of proteins to determine which is best to pursue in further SFX experiments. The growth of micro-crystals from undiluted seed stocks gave a large increase in resolution, allowing the electron density around side-chains to be better elucidated. By incorporating two rounds of seeding into micro-crystal growth, the resolution was pushed yet further to 2.4 Å. It is the aim of this paper to introduce these general micro-crystal growth strategies to the wider XFEL community.

The power of the XFEL to produce high quality structures was also detailed in **Paper III**. The near 1000 fold increase in the multiplicity of XFEL data sets over conventional synchrotron experiments led to electron density maps that supported modeling of a flexible loop region of the protein for the first time. This is notable particularly as the deposited XFEL structure is lower in resolution than the best obtained from synchrotron sources. Lipid molecules could also be observed that were not present in earlier synchrotron structures. The relevance of this structure is further validated by the fact it clusters closely with the bulk of other deposited RC_{vir} structures when scoring by an internal difference matrix.

The ultimate goal of this thesis is to investigate conformational gating on an ultrafast timescale through time-resolved SFX experiments on RC_{vir}. In order to observe these structural differences that may be very small, it is important to have high quality crystal data where the position of side chains can be accurately interpreted. The micro-crystal development in Paper III laid the groundwork for TR-SFX experiments on RC_{vir}, allowing high quality difference electron density maps to be obtained where small movements of side chains and cofactors could be observed (**Paper V**).

5. Time-resolved serial crystallography study of photosynthetic reaction center – Paper V

Photosynthetic reaction centers have an extremely efficient photo-chemistry, harnessing energy from sunlight and converting it into the chemical energy that provides fuel for most life on Earth. The initial charge separation reaction, with an electron transferring from the special pair of chlorophylls (P_{960}) to the L branch pheophytin ($BPhe_L$), occurs across 9 Å in less than 3 picoseconds at room temperature[30]. This is too fast to be explained by the electron transfer theory developed by Marcus[160,161]. As such, there has been much debate on whether the monomeric $BChl_L$ is involved in the charge separation reaction as an intermediary in a two-step electron tunneling process from P_{960} to $BPhe_L$[31,48,162-164]. The electron transport chain has been well studied spectroscopically and ultrafast structural changes have been implicated in the overall efficiency of the process in closely related reaction centers from other purple bacteria[49,51-53,165]. A deeper understanding of this process would extend fundamental knowledge about the nature of electron transport within biological systems. The advent of XFEL sources with their sub 100 fs pulse length provides the opportunity to study structural changes on the picosecond time-scale. The aim of the experiments in **Paper V** was to investigate, using TR-SFX, if structural changes occur at two time points after photo-excitation of RC_{vir}, Δt = 5 ps and 300 ps. At 5 ps after photon absorption, P_{960} has become excited and transferred an electron to $BPhe_L$ forming a charge separated state. After 300 ps, the electron has been further transferred to the bound menaquinone (Q_A) cofactor in the L subunit (Figure 5.1). Conformational changes were found in amino acids around the special pair and along the L subunit

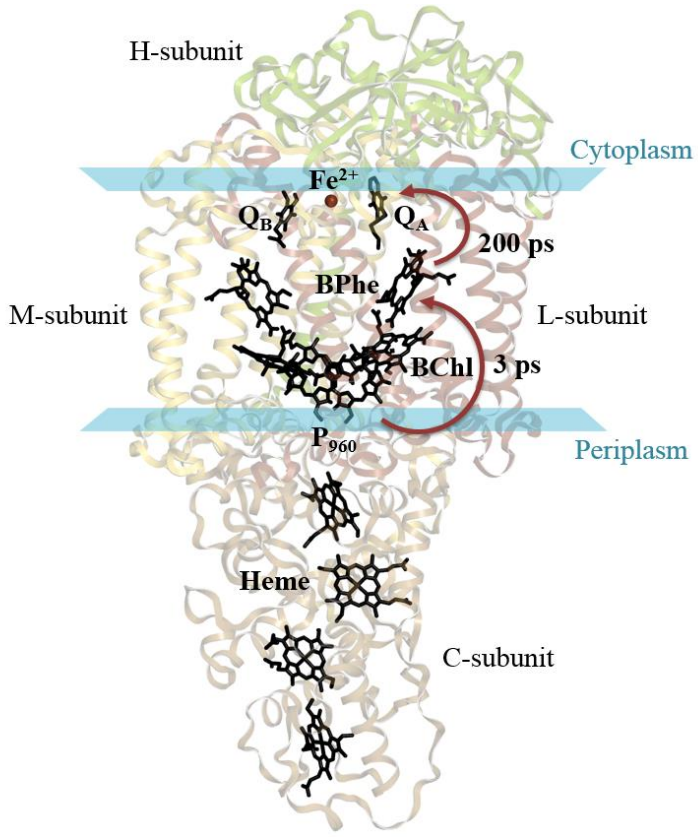


Figure 5.1: *Structure of the photosynthetic reaction center from* B. viridis *showing the position of the protein in the membrane and the positions of cofactors. The timescales of the initial electron transfer events are displayed.*

that evolved between 5 ps and 300 ps. A key finding which confirms the successful photo-activation of RC_{vir} in micro-crystals was the shortening of a hydrogen bond between Q_A and a histidine residue that coordinates the non-heme iron. The scale and implications of these ultrafast structural changes are discussed herein.

5.1 Pump-probe TR-SFX at an XFEL

Micro-crystals were grown from one round of seeding as described in **Paper III**. These crystals were used in a TR-SFX experiment at the CXI beamline[121] at the LCLS in April 2015. The crystals were concentrated by low speed centrifugation (1000 *g* for 1 min) and enough supernatant was removed so as to concentrate the crystals three-fold. This concentrated micro-crystal solution was filtered through a 20 µm stainless steel filter (VICI AG International) to remove any larger crystals that could clog the injection apparatus. The sample was loaded into the reservoir, which was placed on an anti-settling device[122], and the crystals were then injected into the X-ray beam through the GDVN[79] system. A Ti:Sa 960 nm femtosecond laser of 150 fs pulse-length and 11.75 µJ energy focused to a spot size of 190 FWHM was aligned with the XFEL beam and precise timing was achieved using the time-sorting device[124] installed at the CXI beamline. As opposed to the 800 nm laser used in **Paper II**, the 960 nm laser directly excites the special pair of chlorophylls, P_{960}. After a single photon-absorption event, the absorption spectrum of the protein at 960 nm is bleached and it effectively becomes invisible to further laser excitation[30,166]. Conformational changes observed in the data are therefore due to a single photon absorption event. Diffraction data was collected on the ground state of the protein and after time delays corresponding to 5 ps and 300 ps following photo-excitation. The X-rays had an energy of 6.6 keV, a 36 fs pulse duration and a repetition rate of 120 Hz, focused to a 10 μm^2 spot. Diffraction intensities were recorded on the CSPAD[77]. The software CHEETAH[123] sorted diffraction images from empty frames by analyzing numbers of diffraction peaks. Diffraction images were converted into the HDF5 format and sorted into laser-on and laser-off data based on an event code in the instrument. The data was then passed to the software suite CrystFEL for processing.

5.2 Data Processing

CrystFEL was used to optimize the geometry (geoptimiser[156]), index and integrate the data (indexamajig calling Moslfm[127] and Dirax[128]), and scale and merge diffraction intensities (partialator[155]). The space group identified for the crystals in the ground state and the excited states was $P4_32_12$, with cell parameters a = b = 226.4, c = 113.7, $\alpha = \beta = \gamma = 90$ °, consistent with the XFEL structures presented in **Paper III**. The reflection

intensity data compiled by CrystFEL was passed to CCP4 for structural refinement. Initial focus was on the ground state structure from which the excited state structures would later be refined. Structure factor amplitudes were calculated using the program Truncate[157]. Phases were obtained by molecular replacement in the program Phaser[136] using the RC_{vir} structure collected from a later XFEL SFX experiment, the 2.4 Å structure presented in **Paper III** (PDB ID 5NJ4) as a model. Restrained refinement was carried out using multiple rounds of 10 cycle refinement in Refmac[139] with manual modeling using Coot[137]. The structure was refined to 2.8 Å and the validity of the final structure was assessed using Molprobity[158] and Procheck[159]. 99.6 % of amino acids lay in allowed regions of the Ramachandran plot, with five outliers. $2mF_{obs} - DF_{calc}$ maps of the areas surrounding the special pair of chlorophylls and the Q_A cofactor are displayed in Figure 5.2. The refined structure showed a good fit to the electron density, allowing side chains to be modelled with some precision. The quality of the data was sufficient to examine subtle movements of individual amino acids between the ground and photo-excited states through analysis of electron density difference maps.

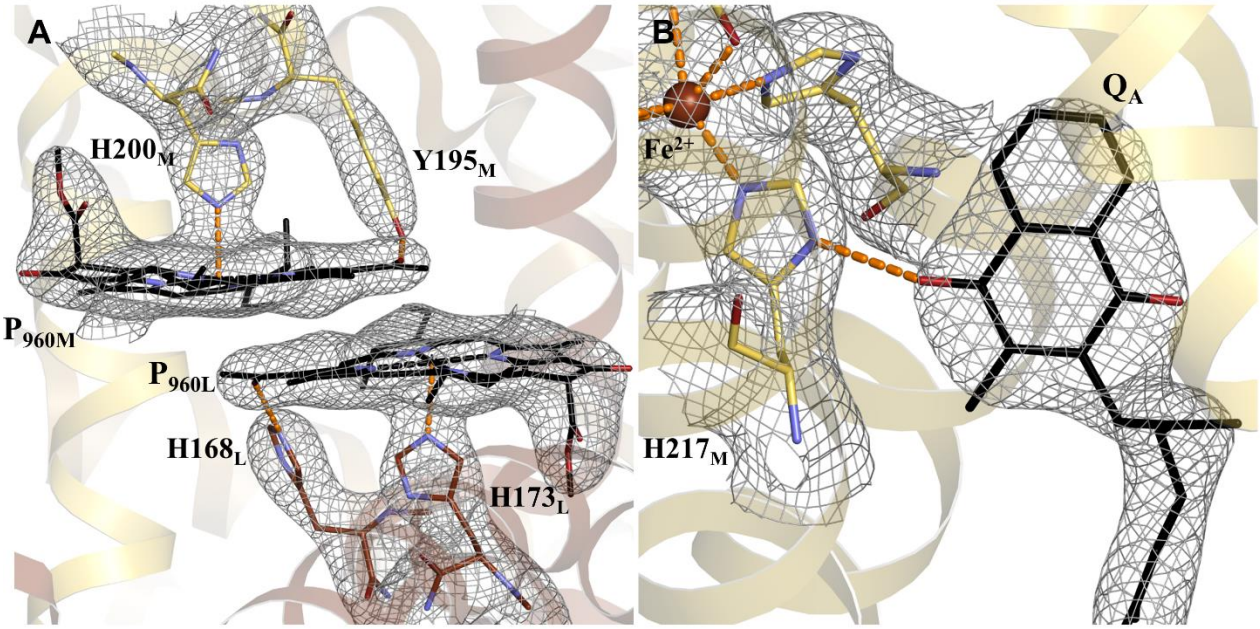


Figure 5.2: *$2mF_{obs} - DF_{calc}$ maps of the special pair and the ubiquinone binding pockets contoured at 1.5 σ overlaid on the ground state structure refined to 2.8 Å resolution. The side chains of the amino acids surrounding the cofactors show a good fit to the density.*

Electron density difference maps between photo-excited and dark states of the protein were produced as a guide to refining the charge-separated states. The observed structure factor amplitudes of the ground state were subtracted from those of the photo-excited state ($|F_{obs}^{light}| - |F_{obs}^{dark}|$) and phases were used from the refined ground state. To improve signal to noise in the difference maps, a Bayesian weighting was applied to the

phases[118]. In areas around significant electron density difference peaks that indicated meaningful structural changes, alternate conformations of cofactors and amino acids were built into the density and given an occupancy of 30 %. The modelled structural changes were refined against the observed structure factors of the excited states using 10 rounds of refinement in Phenix[138].

5.3 Interpretation of difference maps

Figure 5.3 displays the global electron density difference maps between the ground state and the two time delays, Δt = 5 ps and 300 ps. Ultrafast structural changes in photosynthetic reaction centers have not been reported from crystallographic data and it was unknown whether there would be any significant electron density differences on such short timescales. What is therefore remarkable about the overview of the difference Fourier maps is the emergence of difference density around the special pair of chlorophylls and extending up the L branch of the protein towards Q_A, along the path of electron transport. These maps also show evolution of difference density between the 5 ps and 300 ps time delays. At Δt = 5 ps, difference density clusters mostly around the oxidized special pair (P_{960}^+), the L-branch pheophytin and transmembrane helices adjacent to them (E_L and D_M). At Δt = 300 ps, many of the difference density peaks become stronger and significant difference electron density can be seen around the Q_A, which by this time delay has been reduced.

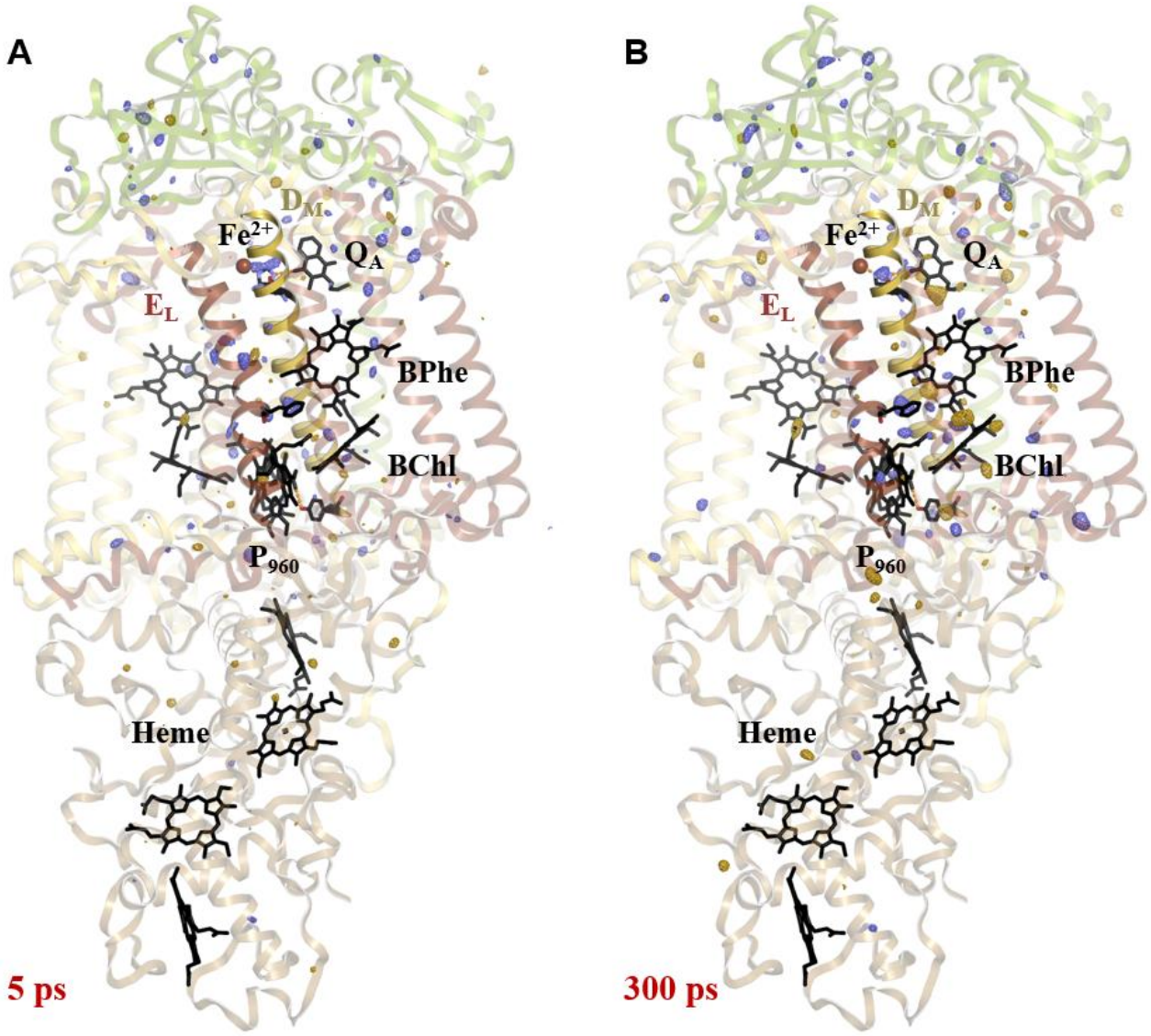


Figure 5.3: *Global difference electron density maps ($F_{obs}^{light} - F_{obs}^{dark}$) between photo-excited (Δt = 5 ps, left; Δt = 300 ps, right) and ground state RC_{vir} overlaid on the refined ground state structure. Maps for both time delays are contoured at 4.5 σ with positive electron density shown in blue and negative electron density shown in gold. Two transmembrane helices with particularly strong difference electron density are highlighted (E_L and D_M).*

Paired negative and positive difference electron density features above and below one of the chlorophyll molecules in the special pair (P_{960M}) implied that the two chlorophylls move slightly closer together after the photo-oxidation event. Strong difference electron density features also appear around a tyrosine residue for Δt = 300 ps ($Y195_M$). This polar residue has a hydrogen bond with P_{960M} ($Y195_M$). The difference density features indicate that this residue readjusts its conformation to maintain its bonding to the special pair. There are no such features on the symmetrically equivalent residue with a hydrogen bond to P_{960L} ($H168_L$). These features are displayed in Figure 5.4A and B.

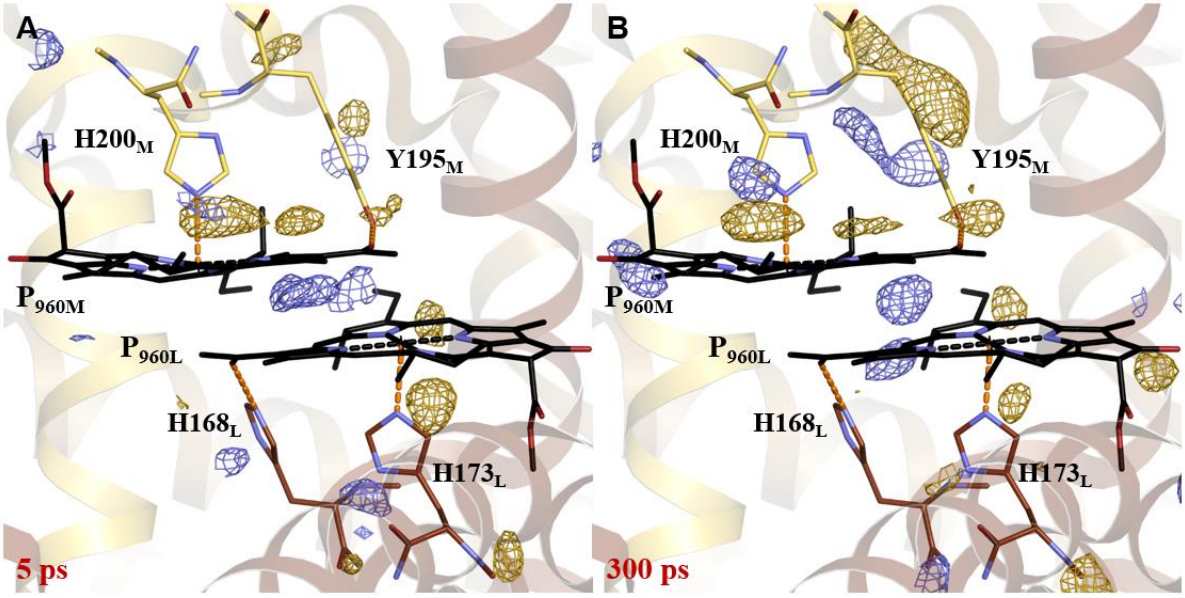


Figure 5.4: *Difference electron density maps around the special pair region contoured at 3.4 σ. Positive electron density is displayed in blue and negative electron density is displayed in gold.* ***A)*** *Difference Fourier map for Δt = 5 ps contoured to 3.4 σ revealing paired difference electron density above and below P_{960M} as well as weak difference electron density around $Y195_M$.* ***B)*** *Difference Fourier map for Δt = 300 ps contoured to 3.4 σ. The difference electron density around P_{960M} is maintained while the difference density around $Y195_M$ becomes stronger, indicating a structural movement to maintain its hydrogen bond as P_{960} moves.*

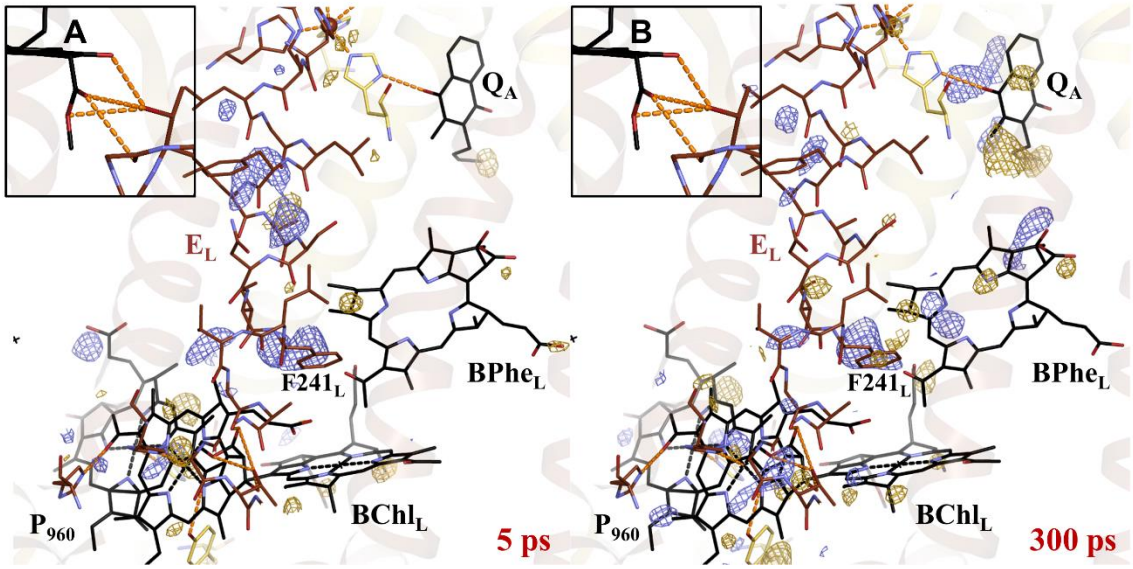


Figure 5.5: *Difference electron density maps at Δt = 5 ps* ***(A)*** *and Δt = 300 ps* ***(B)*** *around the E_L helix extending from the special pair of chlorophylls to the bound quinone molecule. Maps are contoured at 3.7 σ with positive density in blue and negative density in gold. Strong positive features can be observed extending up the helix and particularly around $F241_L$ which has π-stacking interactions with $BPhe_L$. The hydrogen bonding network between P_{960L} and E_L is shown in the inset.*

P_{960L} has a hydrogen bonding network with a helix in the L subunit, E_L (Figure 5.5 inset). This transmembrane helix contains some of the strongest difference density peaks in the entire difference Fourier map at both time delays, displayed in Figure 5.5. At Δt = 5 ps, strong positive features extend up the helix towards $BPhe_L$ and Q_A, but have little corresponding negative features. While requiring appropriate controls and deeper analysis, it is tempting to speculate that these positive features may arise due to selectively enhanced normal modes, as have been observed in a THZ-stimulated study on lysozyme[107]. At both Δt = 5 ps and 300 ps, strong features are observed around $F241_L$, indicating a light-induced movement towards $BPhe_L$.

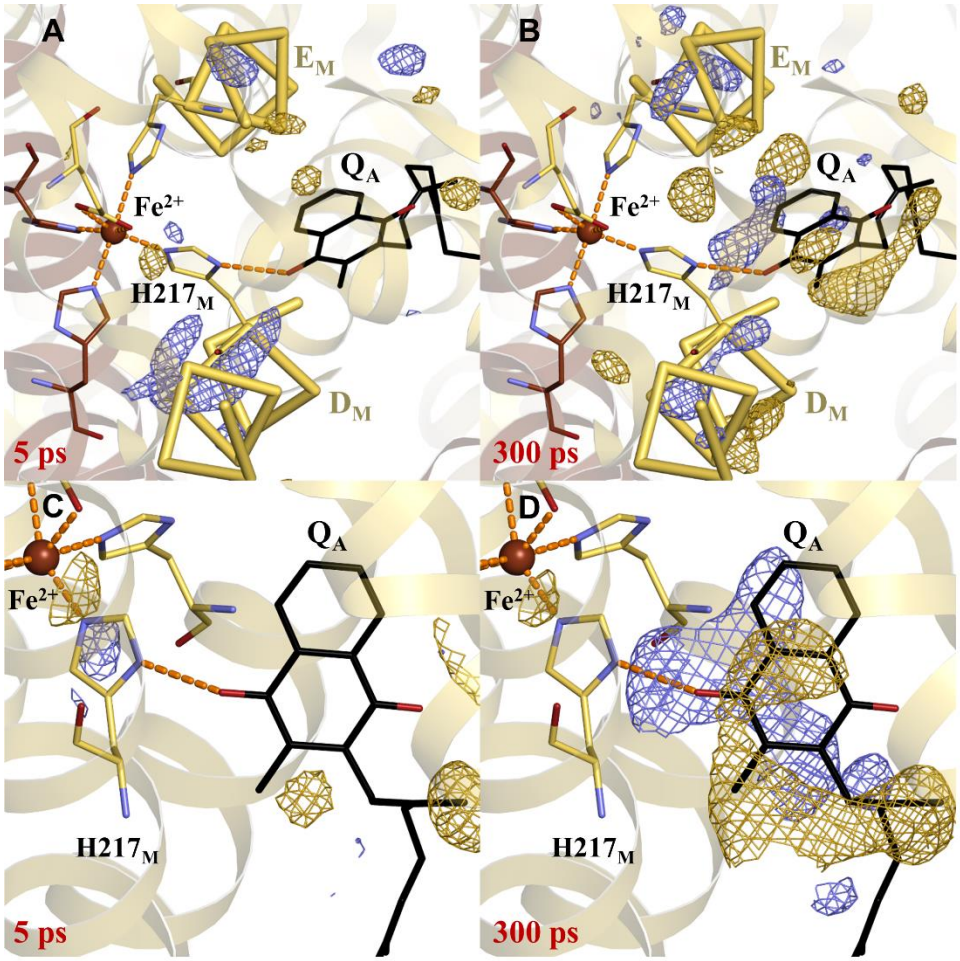


Figure 5.6: *Electron density difference maps around the Q_A binding pocket at two angles after time delays of 5 ps and 300 ps. Positive electron density is displayed in blue with negative density in gold* ***A-B)*** *Little electron difference density around the Q_A can be observed after 5 ps, although helices D_M and E_M do show some relatively strong positive density. At a time delay of 300 ps, a significant difference density around the quinone can be observed as well as around the transmembrane helices surrounding it, suggesting a rearrangement of the Q_A binding pocket. Maps contoured at 3.7 σ* ***C-D)*** *Close up view of the electron difference density around Q_A. After 300 ps, the Q_A is mostly reduced and difference density surrounding it can be interpreted as a movement towards $H217_M$. Maps contoured at 3 σ.*

Electron transport proceeds from P_{960} to $BPhe_L$ and then to Q_A, which is reduced after around 200 ps[34]. The difference Fourier map for Δt = 5 ps shows little change around the Q_A cofactor itself, although there are features around adjacent helices. At Δt = 300 ps, after which the cofactor is mostly in its reduced form, there are significant positive and negative features (Figure 5.6A-D), suggesting a rearrangement of the cofactor in its binding pocket with a movement towards a histidine residue ($H217_M$) coordinating the non-heme iron. Viewed as whole, the electron difference density peaks observed seem to emanate out from the special pair and along the transmembrane helices closely associated with the cofactors involved in electron transport, $BPhe_L$ and Q_A.

5.4 Implications of structural changes on electron transport

The general principle of conformational gating in photosynthetic reaction centers is that ultrafast structural changes alter the activation and reorganization energies for the forward and backward reactions of the electron transport process, so that an electron leaving P_{960} is favored over an electron returning to P_{960}^+. An initial look at the global difference Fourier maps presented in Figure 5.3 reveal that differences in the electron density of ground and photo-excited states are in regions that could well have an effect on the free energies of electron transport steps. Significant features are found around P_{960}, $BPhe_L$ and Q_A as well as around the helices closest to these cofactors. Conformational changes in these regions were modelled with structural refinement of the photo-excited states and the refined structures will provide a basis from which to examine theoretically how these movements may affect electron transport.

Partial occupancy refinement of the conformational changes around photo-oxidized P_{960}^+ show the two chlorophylls moving 0.1 Å closer together as well as a sliding of the pair 0.3 Å away towards the M-subunit. Subtle movements in the histidine residues that coordinate the special pair indicate a weakening of the ligating bonds. A movement of a nearby H-bonded tyrosine residue ($Y195_M$) is also seen (Figure 5.7A). Mutations in $Y195_M$ have been shown to have an effect on the low frequency protein vibrational modes observed in time-resolved spectroscopic studies[50]. It can therefore be hypothesized that movements in this residue couple low frequency dynamical modes of P_{960}^+ to the wider protein matrix surrounding it.

Significant difference density peaks were observed in helix E_L adjacent to the special pair. Structural refinement in this region shows a displacement of the helix by as much as 0.5 Å after a time delays of 5 ps (Figure 5.7B). Although this seems to be an extremely rapid movement, similar scale helix displacements have been observed in a multi-photon absorption TR-WAXS study on RC_{vir}[115] as well as in a TR-SFX study on myglobin[117]. The relatively large collective structural movements in this transmembrane

helix E_L as well as the adjacent helix D_M may also effect the free energies of electron transport. E_L has a hydrogen bonding network with P_{960L} and D_M is adjacent to the binding pocket of both the $BPhe_L$ and Q_A cofactors. Structural changes in this region directly link the photo-oxidation event in P_{960} to the electron acceptor cofactors. The movement of residue $F241_L$ around 0.3 Å towards the binding pocket of $BPhe_L$ at a time delay of 5 ps affects π-stacking interactions, and this may alter the energy barrier for the electron transfer.

5 ps after photo-excitation, the region around the Q_A binding pocket shows little structural change from the ground state. However after 300 ps, the reduced Q_A has moved 0.2 Å towards $H217_M$ which coordinates the non-heme iron. Structural refinement also reveals substantial displacement of the helices surrounding Q_A (Figure 5.7C). The shortening of the H-bond between Q_A^- and $H217_M$ can be explained by increased electro-negativity of the reduced ketone group on Q_A^-. The movements of helices D_M and E_M around the Q_A^- binding pocket may help tune the free energy barrier of the reverse electron transfer reaction that would collapse the charge separated state, thereby optimizing the efficiency of the process.

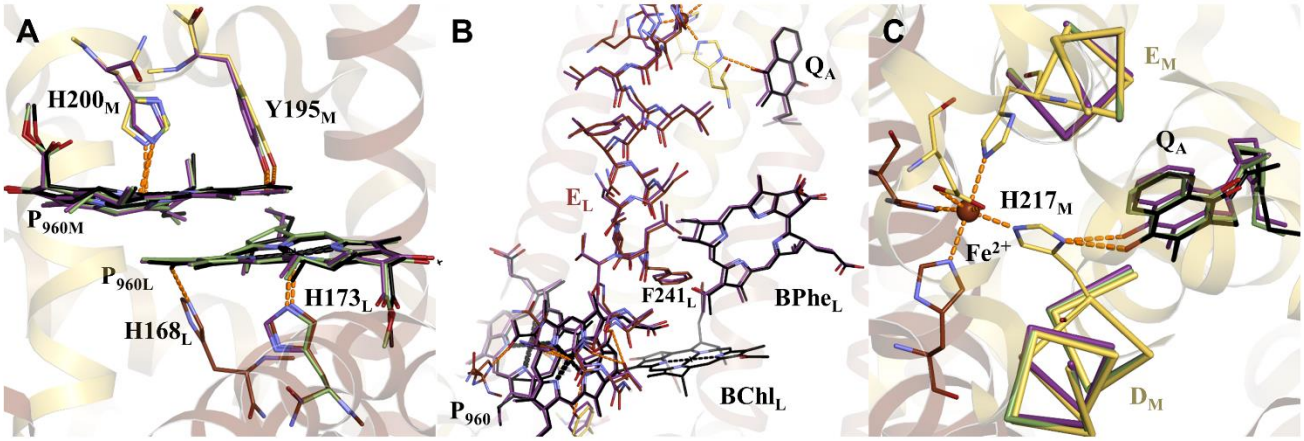


Figure 5.7: *Results of structural refinement of photo-excited structures (5 ps green; 300 ps purple) overlaid on the ground state structure.* ***A)*** *Structural refinement on the region around the special pair shows the chlorophyll P_{960M} moving 0.1 Å towards P_{960L}. $Y195_M$ moves with the cofactor to maintain its hydrogen bond.* ***B)*** *Movements of up to 0.7 Å can be seen in helix E_L, in close proximity to the cofactors involved in electron transport (5 ps structure is omitted here for clarity).* ***C)*** *At a time delay of 5 ps, the Q_A cofactor shows negligible conformational change while helix D_M moves subtly towards the quinone's binding pocket. At a time delay of 300 ps, the Q_A undergoes a substantial movement towards the non-heme iron and helix D_M shows further displacement.*

5.5 Paper Summary

A TR-SFX experiment was carried out at the CXI beamline at the LCLS in order to investigate postulated ultrafast structural changes associated with the primary charge separation reactions of RC_{vir}. Two time delays were chosen to study, Δt = 5 ps and 300 ps after photo-excitation. These time-delays correspond to important times in the electron transport process. After 5 ps, the special pair of chlorophylls has become photo-oxidized and an electron is transferred along the L-subunit of the protein to a pheophytin cofactor. After 300 ps, an electron has been transferred from the reduced pheophytin to a bound quinone molecule on the cytoplasmic side of the membrane. Micro-crystals grown using a seeding method diffracted to 2.8 Å resolution and allowed a detailed visualization of the protein structural changes.

Electron density difference maps between the two photo-excited states and the ground state revealed dominant features around the special pair and along adjacent transmembrane helices closely associated with the $BPhe_L$ and Q_A cofactors that are intermediary electron acceptors in the electron transport chain. Structural refinement of the excited states showed significant conformational changes in this region, the active branch of electron transport. Subtle movements in amino acids that hydrogen bond to the special pair may couple the low frequency dynamical modes of P_{960}^+ to the protein matrix. Ultrafast transmembrane helix movements showing Cα displacement of ~ 0.5 Å could influence free energy barriers along the electron transport chain and aid in the charge transfer process, helping to prevent charge recombination. Interestingly, movement of the helices adjacent to the binding pocket of Q_A occurs before Q_A becomes reduced and subsequently moves. This raises fascinating questions about the nature of the observed ultrafast structural changes, and whether they may help fine tune electron transport from $BPhe_L$ to Q_A.

Submission of this article is awaiting the results of molecular dynamics and quantum mechanical simulations on ground state and charge-separated reaction centers. The aim of this further analysis is to reproduce the structural changes observed in the crystallographic data. A follow-up TR-SFX experiment was carried out on the well-diffracting (2.4 Å) crystals presented in **Paper III**, with data collected at 4 time points, corresponding to Δt = 1, 5, 20 and 300 ps after photo-excitation. This experiment was hampered by lower population of the excited states due to the slightly larger crystals used. However with careful rejection of the largest crystals from data refinement, the results presented in this manuscript could be reproduced. A further aim is to build on the results described here with the data from this later TR-SFX experiment.

6. Concluding remarks and future perspectives

The papers comprising this thesis explore the technological capabilities of the X-ray free-electron laser (XFEL) technology in regards to exploring the early structural events associated with photo-excitation of the photosynthetic protein RC_{vir} from *B. viridis*. **Paper I** gives a detailed review of the current literature surrounding bacterial photosynthetic reaction centers and postulates on whether ultrafast structural movements could occur on the picosecond time scale that stabilize the charge-separated state. This paper also discusses the time-resolution achieved in time-resolved wide-angle X-ray scattering (TR-WAXS) experiments on RC_{vir} as well as the possibilities of visualizing ultrafast conformation changes using time-resolved serial femtosecond crystallography (TR-SFX). **Papers II, III and V** use the methods and ideas described in **Paper I** to build on literature knowledge of RC_{vir} by utilizing XFEL technology.

Paper II describes a TR-WAXS study of RC_{vir} that aimed to give a more physiological relevance to an article from Arnlund *et al*. that described a 'protein quake' of structural movements radiating out from the special pair of chlorophylls following multi-photon excitation[115]. In the experiment described in **Paper II**, the monomeric chlorophylls were excited with a 150 fs pulse-length Ti:Sa femtosecond laser. A power titration was also performed in order to observe the evolution of difference scattering curves from a multi-photon absorption event down to a single-photon absorption event. Although the difference scattering curves do change in sub-structure at low laser powers, oscillations resulting from conformational changes are still observed. This hints that in a physiological setting, conformation changes do occur on the picosecond time-scale and may contribute to the efficiency of the charge-separation reaction.

Small crystals measuring no more than 20 μm in all dimensions are generally required for XFEL SFX experiments owing to the narrow tubing they must pass through before delivery into the X-ray beam. **Paper III, IV and V** are all SFX studies and **Paper III** in particular describes the steps that can be taken to produce well-diffracting micro-crystals from larger macro-crystals. Over three experiments at the LCLS, resolution of RC_{vir} micro-crystals was improved from 3.4 Å to 2.4 Å, giving more detailed electron density maps and allowing side-chains to be well modelled. The improvements in resolution from mechanically crushed macro-crystals to micro-crystals grown from seeds facilitated obtaining the convincing time-resolved SFX (TR-SFX) data described in **Paper V**. This paper also demonstrates the power of the XFEL in delivering high multiplicity data sets that can better capture flexible regions of proteins that show very weak electron density in synchrotron data. **Paper IV** describes SFX experiments on the light-sensing domain of bacterial phytochromes, and compares room-temperature structures obtained at the XFEL to a cryo-cooled synchrotron structure. The data

collected in this experiment at the SACLA XFEL was obtained from crystals that were crushed as per the method described in **Paper III**.

In **Paper V**, TR-SFX experiments were performed on RC_{vir} micro-crystals grown from seeds. The crystals were excited with a 960 nm Ti:Sa femtosecond laser in order to directly excite the special pair. The laser was aligned with the X-ray beam in order to capture diffraction data two time-points after photo-excitation, 5 ps and 300 ps. Ground state data was also collected. Difference electron density maps were produced comparing the structure factor amplitudes between the laser-on excited states and the laser-off ground state. These maps showed strong peaks around the special pair of chlorophylls and on the L-branch of the protein where the electron transport chain progresses, indicating that structural movements were occurring on this time-scale. Furthermore, the maps showed an evolution between the 5 ps and 300 ps state, with difference electron density becoming stronger in some areas, and weaker in others. At the bound quinone (Q_A) site, which becomes reduced 200 ps after photo-excitation, there was strong difference electron density peaks at 300 ps, and very little difference electron density at 5 ps. Modelling of the structural changes revealed that this quinone molecule, once reduced, shortened its hydrogen bond with a non-heme iron coordinating histidine residue. Structural changes around the special pair indicating a shortening of the distance between the chlorophyll molecules and a movement of a hydrogen-bonded tyrosine residue. Transmembrane helices adjacent to the special pair and Q_A also showed large displacements upon structural refinement which appear to link the photo-oxidation event at P_{960} to subsequent steps in the electron transport chain. These ultrafast structural changes may play a role in conformational gating of the electron transport process, increasing the activation barrier for the collapse of the charge-separated state.

Paper VI is a related paper to this thesis and describes a molecular movie of the green-light sensing bacteriorhodopsin. This paper used TR-SFX at an XFEL to obtained datasets from 13 different time-points after photo-excitation and modelled the evolving structural changes as they propagated over time. The next step in the saga of the RC_{vir} protein is to try and replicate the scope of these results. Work is being done on crystallizing the protein in the lipidic cubic phase that would greatly reduce sample consumption and allow more data to be collected from less protein. These crystals will be spiked with ubiquinone to allow it occupy the mobile quinone (Q_B) site, which is not occupied in the crystals used in **Papers III and V**. This will allow structural changes to be followed over a longer time-scale and may make production of an RC_{vir} molecular movie possible. A molecular movie of RC_{vir} would give a much deeper understanding of conformational gating of the electron transport process.

7. Acknowledgements

Five years ago I made the rash decision to apply to ad I found online for a PhD position in some small town I hadn't heard of in Sweden. I was flown over for an interview and on the morning I would meet my future supervisor, I debated for a long time whether to wear the suit I had brought or not. I decided not to, and in hindsight, it was a good choice. The man who would be my supervisor rocked up late wearing an earring and an old T-shirt from some beamline or another, and I bluffed my way through a conversation about rugby and cricket. Somehow I managed to land the job, and thus begun my grand Swedish adventure. Five years later I am well versed in Swedish life. I know my Håkan from my Sanna Nielsen, I've sampled the delights of surströmming, and I understand the solemn importance of the thrice-daily fika. I've learnt a great deal during my time at the Lundberg lab, not only about the wonders of XFEL radiation, but also about friendship, love and life. I would like to acknowledge some of the people who have been there with me along the way.

The casually dressed man who interviewed me that day was **Richard**. I would like to thank you for giving me this opportunity, I've had an incredible time at the lab here. The LCLS beamtimes were some of the most nerve-wracking, exciting and rewarding times of my PhD. You always manage to be optimistic about results even when I don't quite see it, and your guidance in this time-resolved SFX work has been fantastic. I think we have achieved something really great. Because of you I will always live by the maxim, 'shoot first; ask questions later'.

Gisela, you are the calm and measured yin to Richard's sometimes chaotic yang. You have been an excellent co-supervisor. I'm sure you must have been a bit annoyed the last few months when I kept bursting in to your room to consult you on random issues, but it never showed. So thank you for the help with crystallization and the thesis, and for calming me down when I have been going crazy. You did a good job picking your first PhD student too, **Cece** is a great friend and colleague.

I would like to thank **Martin**, my examiner for always answering my queries about how this whole thesis thing works.

To my predecessors and my protégé in the Neutze RC division, it's been an amazing time together. **Linda** and **David**, we bonded over Lili and Susie during a night shift in Chicago and I made two good friends, thanks for the proofreading David. **Petra**, I've worked closely with you over the last few years and your creative scientific mind and love of J-pop has inspired me. You're going to breeze through your defense when it comes time.

To the rest of the Neutze gang, **Cecilia**, **Elin**, **Greger**, **Rajiv**, **Jennie**, **Peter**, **Rebecka**, **Alex**, **Rob**, **Daniel** and **Amit**, it's been great working with you and getting to know you. **Rebecka**, that painting of you was my muse as I wrote up this thesis. **Elin**, I feel I didn't know you that well, and was a bit nervous about being one on one with you on a 10 hour flight to Japan, but it was awesome! I'm really glad I can now call you a friend. **Cecilia**, thanks for all the hard work you put into keeping this place somewhat safe to spend time in and for the work you did on the BR and RC papers. Good luck with the new addition to the family!

My office mate and toast master **Pia**, you have tolerated smelly lunchboxes, a dead tree and my feet on your chair, thanks for being a good roomie. I'm so excited to come to your wedding and see you marry one of the loveliest men I know!

Vlad, you gave me use of your computer on my first day in the office, and I made one of my best friends ever, you're amazing. A few years after meeting you I met **Joe**, who stuck around with me while you vanished off to Stockholm. One day soon you will both join me on the other side and I can't wait to be there to support you.

The Westenhoff group, **Sebastian**, **Elin**, **Emil**, **Oskar**, **Leo**, **Linnéa, Joachim, Mattÿs** (and formally **Ash**), it has been great collaborating with you on the phytochrome project, and also on the beer drinking project. **Oskar**, you were a worthy successor to the beer club crown no matter what I have said in the past. **Ash**, Stavros is going to miss you and **Liane**, good luck with the new life at Diamond! **Leo** and **Elin**, I'm looking forward to our next De Vet Du concert!

Rosie, **Rhawnie**, **Parveen**, **Weixao**, **Kristina**, **Florian** and **Stefan**, it's been a pleasure working with you all! Big thanks to **Rosie** for helping me touch up my CV and cover letter, and having my favourite laugh in the corridor.

Gergely, Majo, Maja and Viktor, thank you for all the good times and good luck with all your projects. **Majo**, estas muy caliente! **Maja**, keep showing your kids that science is awesome. **Gergely**, thank you so much for answering all my crystallographic queries.

Johanna and **Davide**, the new(ish) kids on the block. You have been great additions to the corridor and I wish you all the best with your group as you build up our EM expertise.

Stefan, **Annette**, **Ida**, **Mike**, **Mikael**, **Alex B**, **Karin**, the 'golden oldies'. It was great getting to know you all, whether it was through climbing, or drinking wine, or making

risotto, or chasing each other on an island.

I would also like to thank all the Masters students and project workers who have come and gone and brightened up the place. Particularly **Jordi**, the good-hearted Belgian who opened his house to me and my stinking cyclist friends on our trip through Europe.

Lars and **Bruno**, this lab wouldn't work without you. There has been countless times when I've had to run downstairs for a glitchy computer or a broken machine, and you've always been there to help.

Jasper, I may have been 45 minutes late to our first date, but I am truly glad you stuck around anyway. After only a month together you made the hop over to England to meet all of my friends at once, and then you came to my brother's wedding and met all of my extended family at once. You're brave, you're passionate, and I love you. Not only have we had great adventures together, but you also brought back the fluff-monster from Greece. Our little family, you, me and **Stavros**, just warms my heart to think to about. Here's to many more years chasing those pesky rabbits together!

Finally I would like to thank my friends and my family back home. I've got some inspirational parents who always have an ear to lend, and some weird and wonderful friends.

8. References

1 Ingenhousz, J. *Experiments upon vegetables: discovering their great power of purifying the common air in the sun-shine, and of injuring it in the shade and at night. To which is joined, a new method of examining the accurate degree of salubrity of the atmosphere*. (Printed for P. Elmsly and H. Payne, 1779).

2 Bekker, A. in *Encyclopedia of Astrobiology* (eds Ricardo Amils *et al.*) 1-9 (Springer Berlin Heidelberg, 2014).

3 Dismukes, G. C. *et al.* The origin of atmospheric oxygen on Earth: The innovation of oxygenic photosynthesis. *Proceedings of the National Academy of Sciences* **98**, 2170-2175 (2001).

4 Madigan, M. T., Martinko, J. M. & Parker, J. *Brock biology of microorganisms*. Vol. 11 (prentice hall Upper Saddle River, NJ, 1997).

5 Pfennig, N. Phototrophic green and purple bacteria: a comparative, systematic survey. *Annual Reviews in Microbiology* **31**, 275-290 (1977).

6 Blankenship, R. E. Origin and early evolution of photosynthesis. *Photosynthesis Research* **33**, 91-111 (1992).

7 Rappaport, F., Guergova-Kuras, M., Nixon, P. J., Diner, B. A. & Lavergne, J. Kinetics and Pathways of Charge Recombination in Photosystem II. *Biochemistry* **41**, 8518-8527 (2002).

8 Ort, D. R. & Yocum, C. F. *Oxygenic photosynthesis: the light reactions*. Vol. 4 (Springer Science & Business Media, 1996).

9 Dagan, T. *et al.* Genomes of Stigonematalean Cyanobacteria (Subsection V) and the Evolution of Oxygenic Photosynthesis from Prokaryotes to Plastids. *Genome Biology and Evolution* **5**, 31-44 (2013).

10 Whatley, J. M., John, P. & Whatley, F. R. From Extracellular to Intracellular: The Establishment of Mitochondria and Chloroplasts. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **204**, 165-187 (1979).

11 Zimorski, V., Ku, C., Martin, W. F. & Gould, S. B. Endosymbiotic theory for organelle origins. *Current Opinion in Microbiology* **22**, 38-48 (2014).

12 Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. X-ray structure analysis of a membrane protein complex. *Journal of Molecular Biology* **180**, 385-398 (1984).

13 Michel, H. Three-dimensional crystals of a membrane protein complex. *Journal of Molecular Biology* **158**, 567-572 (1982).

14 Lancaster, C. R. D. & Michel, H. Refined crystal structures of reaction centres from Rhodopseudomonas viridis in complexes with the herbicide atrazine and two chiral atrazine derivatives also lead to a new model of the bound carotenoid1. *Journal of Molecular Biology* **286**, 883-898 (1999).

15 Johansson, L. C. *et al.* Structure of a photosynthetic reaction centre determined by serial femtosecond crystallography. *Nature Communications* **4**, 2911 (2013).

16 Johansson, L. C. *et al.* Lipidic phase membrane protein serial femtosecond crystallography. *Nat Meth* **9**, 263-265 (2012).

17 Roszak, Aleksander W. *et al.* New insights into the structure of the reaction centre from Blastochloris viridis: evolution in the laboratory. *Biochemical Journal* **442**, 27 (2012).

18 Ponomarenko, N. S. *et al.* Structural and spectropotentiometric analysis of Blastochloris viridis heterodimer mutant reaction center. *Biochimica et biophysica acta* **1788**, 1822-1831 (2009).

19 Li, L. *et al.* Simple Host—Guest Chemistry To Modulate the Process of Concentration and

Crystallization of Membrane Proteins by Detergent Capture in a Microfluidic Device. *Journal of the American Chemical Society* **130**, 14324-14328 (2008).

20 Wöhri, A. B. *et al.* Light-Induced Structural Changes in a Photosynthetic Reaction Center Caught by Laue Diffraction. *Science* **328**, 630-633 (2010).

21 Wöhri, A. B. *et al.* Lipidic Sponge Phase Crystal Structure of a Photosynthetic Reaction Center Reveals Lipids on the Protein Surface. *Biochemistry* **48**, 9831-9838 (2009).

22 Lancaster, C. R. D., Hunte, C., Kelley Iii, J., Trumpower, B. L. & Ditchfield, R. A Comparison of Stigmatellin Conformations, Free and Bound to the Photosynthetic Reaction Center and the Cytochrome bc1 Complex. *Journal of Molecular Biology* **368**, 197-208 (2007).

23 Li, L. *et al.* Nanoliter microfluidic hybrid method for simultaneous screening and optimization validated with crystallization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19243-19248 (2006).

24 Baxter, R. H. G., Seagle, B.-L., Ponomarenko, N. & Norris, J. R. Cryogenic structure of the photosynthetic reaction center of Blastochloris viridis in the light and dark. *Acta Crystallographica Section D* **61**, 605-612 (2005).

25 Baxter, R. H. G. *et al.* Time-resolved crystallographic studies of light-induced structural changes in the photosynthetic reaction center. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 5982-5987 (2004).

26 Deisenhofer, J., Epp, O., Sinning, I. & Michel, H. Crystallographic refinement at 2.3 Å Resolution and Refined Model of the Photosynthetic Reaction Centre fromRhodopseudomonas viridis. *Journal of Molecular Biology* **246**, 429-457 (1995).

27 Lancaster, C. R. D., Bibikova, M. V., Sabatino, P., Oesterhelt, D. & Michel, H. Structural Basis of the Drastically Increased Initial Electron Transfer Rate in the Reaction Center from a Rhodopseudomonas viridis Mutant Described at 2.00-Å Resolution. *Journal of Biological Chemistry* **275**, 39364-39368 (2000).

28 Lancaster, C. R. D. & Michel, H. The coupling of light-induced electron transfer and proton uptake as derived from crystal structures of reaction centres from Rhodopseudomonas viridis modified at the binding site of the secondary quinone, QB. *Structure* **5**, 1339-1359 (1997).

29 Martin, J.-L., Breton, J., Hoff, A. J., Migus, A. & Antonetti, A. Femtosecond spectroscopy of electron transfer in the reaction center of the photosynthetic bacterium Rhodopseudomonas sphaeroides R-26: Direct electron transfer from the dimeric bacteriochlorophyll primary donor to the bacteriopheophytin acceptor with a time constant of 2.8 ± 0.2 psec. *Proceedings of the National Academy of Sciences* **83**, 957-961 (1986).

30 Breton, J., Martin, J.-L., Migus, A., Antonetti, A. & Orszag, A. Femtosecond spectroscopy of excitation energy transfer and initial charge separation in the reaction center of the photosynthetic bacterium Rhodopseudomonas viridis. *Proceedings of the National Academy of Sciences* **83**, 5121-5125 (1986).

31 Wasielewski, M. R. & Tiede, D. M. Sub-picosecond measurements of primary electron transfer in Rhodopseudomonas viridis reaction centers using near-infrared excitation. *FEBS Letters* **204**, 368-372 (1986).

32 Parson, W. W. & Ke, B. Primary photochemical reactions. *Photosynthesis* **1**, 331-385 (2012).

33 Davis, M. S., Forman, A., Hanson, L. K., Thornber, J. P. & Fajer, J. Anion and cation radicals of bacteriochlorophyll and bacteriopheophytin b. Their role in the primary charge separation of Rhodopseudomonas viridis. *The Journal of Physical Chemistry* **83**, 3325-3332 (1979).

34 Holten, D., Windsor, M. W., Parson, W. W. & Thornber, J. P. Primary photochemical processes in isolated reaction centers of Rhodopseudomonas viridis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **501**, 112-126 (1978).

35 Deprez, J., Trissl, H. W. & Breton, J. Excitation trapping and primary charge stabilization in Rhodopseudomonas viridis cells, measured electrically with picosecond resolution. *Proceedings of the National Academy of Sciences* **83**, 1699-1703 (1986).
36 Okamura, M. Y., Paddock, M. L., Graige, M. S. & Feher, G. Proton and electron transfer in bacterial reaction centers. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1458**, 148-163 (2000).
37 Ortega, J. M. & Mathis, P. Effect of temperature on the kinetics of electron transfer from the tetraheme cytochrome to the primary donor in Rhodopseudomonas viridis. *FEBS Letters* **301**, 45-48 (1992).
38 Ortega, J. M. & Mathis, P. Electron transfer from the tetraheme cytochrome to the special pair in isolated reaction centers of Rhodopseudomonas viridis. *Biochemistry* **32**, 1141-1151 (1993).
39 Capaldi, R. A. & Aggeler, R. Mechanism of the F1F0-type ATP synthase, a biological rotary motor. *Trends in Biochemical Sciences* **27**, 154-160 (2002).
40 Drews, G. Structure and functional organization of light-harvesting complexes and photochemical reaction centers in membranes of phototrophic bacteria. *Microbiological reviews* **49**, 59 (1985).
41 Steffen, M. A., Lao, K. & Boxer, S. G. Dielectric Asymmetry in the Photosynthetic Reaction Center. *Science* **264**, 810-816 (1994).
42 Moore, L. J., Zhou, H. & Boxer, S. G. Excited-State Electronic Asymmetry of the Special Pair in Photosynthetic Reaction Center Mutants: Absorption and Stark Spectroscopy. *Biochemistry* **38**, 11949-11960 (1999).
43 Michel-Beyerle, M. E. *et al.* Unidirectionality of charge separation in reaction centers of photosynthetic bacteria. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **932**, 52-70 (1988).
44 Deisenhofer, J. & Norris, J. R. *Photosynthetic Reaction Center*. (Academic Press, 2013).
45 Wraight, C. A. & Clayton, R. K. The absolute quantum efficiency of bacteriochlorophyll photooxidation in reaction centres of Rhodopseudomonas spheroides. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **333**, 246-260 (1974).
46 McConnell, I., Li, G. & Brudvig, G. W. Energy conversion in natural and artificial photosynthesis. *Chemistry & biology* **17**, 434-447 (2010).
47 Lauterwasser, C., Finkele, U., Scheer, H. & Zinth, W. Temperature dependence of the primary electron transfer in photosynthetic reaction centers from Rhodobacter sphaeroides. *Chemical Physics Letters* **183**, 471-477 (1991).
48 Fleming, G. R., Martin, J. L. & Breton, J. Rates of primary electron transfer in photosynthetic reaction centres and their mechanistic implications. *Nature* **333**, 190-192 (1988).
49 Vos, M. H., Rappaport, F., Lambry, J.-C., Breton, J. & Martin, J.-L. Visualization of coherent nuclear motion in a membrane protein by femtosecond spectroscopy. *Nature* **363**, 320-325 (1993).
50 Rischel, C. *et al.* Low frequency vibrational modes in proteins: Changes induced by point-mutations in the protein-cofactor matrix of bacterial reaction centers. *Proceedings of the National Academy of Sciences* **95**, 12306-12311 (1998).
51 Vos, M. H. *et al.* Coherent dynamics during the primary electron-transfer reaction in membrane-bound reaction centers of Rhodobacter sphaeroides. *Biochemistry* **33**, 6750-6757 (1994).
52 Wang, H. *et al.* Unusual Temperature Dependence of Photosynthetic Electron Transfer due to Protein Dynamics. *The Journal of Physical Chemistry B* **113**, 818-824 (2009).
53 Wang, H. *et al.* Protein Dynamics Control the Kinetics of Initial Electron Transfer in

Photosynthesis. *Science* **316**, 747-750 (2007).
54 Andréasson, U. & Andréasson, L.-E. Characterization of a semi-stable, charge-separated state in reaction centers from Rhodobacter sphaeroides. *Photosynthesis Research* **75**, 223-233 (2003).
55 van Mourik, F., Reus, M. & Holzwarth, A. R. Long-lived charge-separated states in bacterial reaction centers isolated from Rhodobacter sphaeroides. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1504**, 311-318 (2001).
56 Stowell, M. H. B. *et al.* Light-Induced Structural Changes in Photosynthetic Reaction Center: Implications for Mechanism of Electron-Proton Transfer. *Science* **276**, 812-816 (1997).
57 Fritzsch, G., Koepke, J., Diem, R., Kuglstatter, A. & Baciou, L. Charge separation induces conformational changes in the photosynthetic reaction centre of purple bacteria. *Acta Crystallographica Section D* **58**, 1660-1663 (2002).
58 Wulff, M. *et al.* Time-resolved structures of macromolecules at the ESRF: Single-pulse Laue diffraction, stroboscopic data collection and femtosecond flash photolysis. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **398**, 69-84 (1997).
59 Fromme, P. XFELs open a new era in structural chemical biology. *Nat Chem Biol* **11**, 895-899 (2015).
60 Neutze, R. Opportunities and challenges for time-resolved studies of protein structural dynamics at X-ray free-electron lasers. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369** (2014).
61 Ihee, H. *et al.* Visualizing reaction pathways in photoactive yellow protein from nanoseconds to seconds. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 7145-7150 (2005).
62 Perman, B. *et al.* Energy Transduction on the Nanosecond Time Scale: Early Structural Events in a Xanthopsin Photocycle. *Science* **279**, 1946-1950 (1998).
63 Schmidt, M., Graber, T., Henning, R. & Srajer, V. Five-dimensional crystallography. *Acta Crystallographica Section A* **66**, 198-206 (2010).
64 Ren, Z. *et al.* A Molecular Movie at 1.8 Å Resolution Displays the Photocycle of Photoactive Yellow Protein, a Eubacterial Blue-Light Receptor, from Nanoseconds to Seconds. *Biochemistry* **40**, 13788-13801 (2001).
65 Aranda, R., IV, Levin, E. J., Schotte, F., Anfinrud, P. A. & Phillips, G. N., Jr. Time-dependent atomic coordinates for the dissociation of carbon monoxide from myoglobin. *Acta Crystallographica Section D* **62**, 776-783 (2006).
66 Schotte, F. *et al.* Watching a Protein as it Functions with 150-ps Time-Resolved X-ray Crystallography. *Science* **300**, 1944 (2003).
67 Šrajer, V. *et al.* Photolysis of the Carbon Monoxide Complex of Myoglobin: Nanosecond Time-Resolved Crystallography. *Science* **274**, 1726-1729 (1996).
68 Cartlidge, E. European XFEL to shine as brightest, fastest x-ray source. *Science* **354**, 22-23 (2016).
69 Barty, A. Single molecule imaging using X-ray free electron lasers. *Current Opinion in Structural Biology* **40**, 186-194 (2016).
70 Aquila, A. *et al.* The linac coherent light source single particle imaging road map. *Structural Dynamics* **2**, 041701 (2015).
71 Bogan, M. J. *et al.* Single Particle X-ray Diffractive Imaging. *Nano Letters* **8**, 310-316 (2008).
72 Gaudin, J. *et al.* Towards simultaneous measurements of electronic and structural properties in ultra-fast x-ray free electron laser absorption spectroscopy experiments. *Scientific Reports* **4**,

4724 (2014).
73 Motomura, K. *et al.* Electronic decay and fragmentation dynamics of iodomethane, multiply core-ionized by photoabsorption of intense XFEL pulses. *Journal of Physics: Conference Series* **488**, 032043 (2014).
74 Nagaya, K. *et al.* Femtosecond charge and molecular dynamics of I-containing organic molecules induced by intense X-ray free-electron laser pulses. *Faraday Discussions* **194**, 537-562 (2016).
75 Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406**, 752-757 (2000).
76 Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73-77 (2011).
77 Hart, P. *et al.* The CSPAD megapixel x-ray camera at LCLS. *Proc. SPIE* **8504**, 85040C (2012).
78 Henrich, B. *et al.* The adaptive gain integrating pixel detector AGIPD a detector for the European XFEL. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **633, Supplement 1**, S11-S14 (2011).
79 DePonte, D. *et al.* Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics* **41**, 195505 (2008).
80 Weierstall, U. Liquid sample delivery techniques for serial femtosecond crystallography. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369** (2014).
81 Cherezov, V. Lipidic cubic phase technologies for membrane protein structural studies. *Current Opinion in Structural Biology* **21**, 559-566 (2011).
82 Cherezov, V. *et al.* High-Resolution Crystal Structure of an Engineered Human β₂-Adrenergic G Protein–Coupled Receptor. *Science* **318**, 1258-1265 (2007).
83 Santos, J. S. *et al.* Crystal Structure of a Voltage-gated K+ Channel Pore Module in a Closed State in Lipid Membranes. *Journal of Biological Chemistry* **287**, 43063-43070 (2012).
84 Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P. & Landau, E. M. X-ray Structure of Bacteriorhodopsin at 2.5 Angstroms from Microcrystals Grown in Lipidic Cubic Phases. *Science* **277**, 1676-1681 (1997).
85 Liao, J. *et al.* Structural Insight into the Ion-Exchange Mechanism of the Sodium/Calcium Exchanger. *Science* **335**, 686-690 (2012).
86 Li, D. *et al.* Crystal structure of the integral membrane diacylglycerol kinase. *Nature* **497**, 521-524 (2013).
87 Weierstall, U. *et al.* Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature Communications* **5** (2014).
88 Conrad, C. E. *et al.* A novel inert crystal delivery medium for serial femtosecond crystallography. *IUCrJ* **2**, 421-430 (2015).
89 Sugahara, M. *et al.* Grease matrix as a versatile carrier of proteins for serial crystallography. *Nat Meth* **12**, 61-63 (2015).
90 Jung, Y. O. *et al.* Volume-conserving trans–cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nat Chem* **5**, 212-220 (2013).
91 Schotte, F. *et al.* Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proceedings of the National Academy of Sciences* **109**, 19256-19261 (2012).
92 Tenboer, J. *et al.* Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* **346**, 1242-1246 (2014).
93 Nango, E. *et al.* A three-dimensional movie of structural changes in bacteriorhodopsin. *Science*

354, 1552-1557 (2016).
94 Nogly, P. *et al.* Lipidic cubic phase injector is a viable crystal delivery system for time-resolved serial crystallography. *Nature Communications* **7**, 12314 (2016).
95 Lanyi, J. K. Mechanism of Ion Transport across Membranes: BACTERIORHODOPSIN AS A PROTOTYPE FOR PROTON PUMPS. *Journal of Biological Chemistry* **272**, 31209-31212 (1997).
96 Ulrich Haupts, Jörg Tittor, a. & Oesterhelt, D. CLOSING IN ON BACTERIORHODOPSIN: Progress in Understanding the Molecule. *Annual Review of Biophysics and Biomolecular Structure* **28**, 367-399 (1999).
97 Edman, K. *et al.* High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature* **401**, 822-826 (1999).
98 Luecke, H. *et al.* Coupling photoisomerization of retinal to directional transport in bacteriorhodopsin1. *Journal of Molecular Biology* **300**, 1237-1255 (2000).
99 Rouhani, S. *et al.* Crystal structure of the D85S mutant of bacteriorhodopsin: model of an O-like photocycle intermediate1. *Journal of Molecular Biology* **313**, 615-628 (2001).
100 Royant, A. *et al.* Helix deformation is coupled to vectorial proton transport in the photocycle of bacteriorhodopsin. *Nature* **406**, 645-648 (2000).
101 Sass, H. J. *et al.* Structural alterations for proton translocation in the M state of wild-type bacteriorhodopsin. *Nature* **406**, 649-653 (2000).
102 Subramaniam, S. & Henderson, R. Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* **406**, 653-657 (2000).
103 Wickstrand, C., Dods, R., Royant, A. & Neutze, R. Bacteriorhodopsin: Would the real structural intermediates please stand up? *Biochimica et Biophysica Acta (BBA) - General Subjects* **1850**, 536-553 (2015).
104 Suga, M. *et al.* Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature* **543**, 131-135 (2017).
105 Kupitz, C. *et al.* Structural enzymology using X-ray free electron lasers. *Structural Dynamics* **4**, 044003 (2017).
106 Hekstra, D. R. *et al.* Electric-field-stimulated protein mechanics. *Nature* **540**, 400-405 (2016).
107 Lundholm, I. V. *et al.* Terahertz radiation induces non-thermal structural changes associated with Fröhlich condensation in a protein crystal. *Structural Dynamics* **2**, 054702 (2015).
108 Cho, H. S. *et al.* Protein structural dynamics in solution unveiled via 100-ps time-resolved x-ray scattering. *Proceedings of the National Academy of Sciences* **107**, 7281-7286 (2010).
109 Ahn, S., Kim, K. H., Kim, Y., Kim, J. & Ihee, H. Protein Tertiary Structural Changes Visualized by Time-Resolved X-ray Solution Scattering. *The Journal of Physical Chemistry B* **113**, 13131-13133 (2009).
110 Cammarata, M. *et al.* Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. *Nat Meth* **5**, 881-886 (2008).
111 Cammarata, M., Levantino, M., Wulff, M. & Cupane, A. Unveiling the Timescale of the R–T Transition in Human Hemoglobin. *Journal of Molecular Biology* **400**, 951-962 (2010).
112 Andersson, M. *et al.* Structural Dynamics of Light-Driven Proton Pumps. *Structure* **17**, 1265-1275 (2009).
113 Björling, A. *et al.* Ubiquitous Structural Signaling in Bacterial Phytochromes. *The Journal of Physical Chemistry Letters* **6**, 3379-3383 (2015).
114 Takala, H. *et al.* Signal amplification and transduction in phytochrome photosensors. *Nature* **509**, 245-248 (2014).
115 Arnlund, D. *et al.* Visualizing a protein quake with time-resolved X-ray scattering at a free-

electron laser. *Nat Meth* **11**, 923-926 (2014).
116 Levantino, M. *et al.* Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nature Communications* **6**, 6772 (2015).
117 Barends, T. R. M. *et al.* Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350**, 445-450 (2015).
118 Ursby, T. & Bourgeois, D. Improved Estimation of Structure-Factor Difference Amplitudesfrom Poorly Accurate Data. *Acta Crystallographica Section A* **53**, 564-575 (1997).
119 Lang, F. & Oesterhelt, D. Microaerophilic growth and induction of the photosynthetic reaction center in Rhodopseudomonas viridis. *Journal of bacteriology* **171**, 2827-2834 (1989).
120 Privé, G. G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397 (2007).
121 Liang, M. *et al.* The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *Journal of Synchrotron Radiation* **22**, 514-519 (2015).
122 Lomb, L. *et al.* An anti-settling sample delivery instrument for serial femtosecond crystallography. *Journal of Applied Crystallography* **45**, 674-678 (2012).
123 Barty, A. *et al.* Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *Journal of Applied Crystallography* **47**, 1118-1131 (2014).
124 Harmand M *et al.* Achieving few-femtosecond time-sorting at hard X-ray free-electron lasers. *Nat Photon* **7**, 215-218 (2013).
125 Svergun, D., Barberato, C. & Koch, M. H. J. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography* **28**, 768-773 (1995).
126 Kirian, R. A. *et al.* Femtosecond protein nanocrystallography—data analysis methods. *Opt. Express* **18**, 5713-5723 (2010).
127 Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D* **67**, 271-281 (2011).
128 Duisenberg, A. Indexing in single-crystal diffractometry with an obstinate list of reflections. *Journal of Applied Crystallography* **25**, 92-96 (1992).
129 Ginn, H. M. *et al.* A revised partiality model and post-refinement algorithm for X-ray free-electron laser data. *Acta Crystallographica Section D* **71**, 1400-1410 (2015).
130 Kroon-Batenburg, L. M. J., Schreurs, A. M. M., Ravelli, R. B. G. & Gros, P. Accounting for partiality in serial crystallography using ray-tracing principles. *Acta Crystallographica Section D* **71**, 1799-1811 (2015).
131 White, T. A. Post-refinement method for snapshot serial crystallography. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369** (2014).
132 Sharma, A. *et al.* Asymmetry in serial femtosecond crystallography data. *Acta Crystallographica Section A* **73**, 93-101 (2017).
133 White, T. A. *et al.* CrystFEL: a software suite for snapshot serial crystallography. *Journal of Applied Crystallography* **45**, 335-341 (2012).
134 Drenth, J. *Principles of protein X-ray crystallography*. (Springer Science & Business Media, 2007).
135 Nakane, T. *et al.* Native sulfur/chlorine SAD phasing for serial femtosecond crystallography. *Acta Crystallographica Section D* **71**, 2519-2525 (2015).
136 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).
137 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*

Crystallographica Section D **60**, 2126-2132 (2004).
138 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66**, 213-221 (2010).
139 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **53**, 240-255 (1997).
140 Brunger, A. T. *et al.* Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallographica Section D* **54**, 905-921 (1998).
141 Ibrahim, M. *et al.* Improvements in serial femtosecond crystallography of photosystem II by optimizing crystal uniformity using microseeding procedures. *Structural Dynamics* **2**, 041705 (2015).
142 Kupitz, C. *et al.* Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature* **513**, 261-265 (2014).
143 Koopmann, R. *et al.* In vivo protein crystallization opens new routes in structural biology. *Nature Methods* **9**, 259-262 (2012).
144 Edlund, P. *et al.* The room temperature crystal structure of a bacterial phytochrome determined by serial femtosecond crystallography. *Scientific Reports* **6**, 35279 (2016).
145 Luft, J. R. & DeTitta, G. T. A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallographica Section D* **55**, 988-993 (1999).
146 Bergfors, T. Seeds to crystals. *Journal of Structural Biology* **142**, 66-76 (2003).
147 Yonath, A. Approaching Atomic Resolution in Crystallography of Ribosomes. *Annual Review of Biophysics and Biomolecular Structure* **21**, 77-93 (1992).
148 Makowski, I. *et al.* Single crystals of large ribosomal particles from Halobacterium marismortui diffract to 6 Å. *Journal of Molecular Biology* **193**, 819-822 (1987).
149 D'Arcy, A., Villard, F. & Marsh, M. An automated microseed matrix-screening method for protein crystallization. *Acta Crystallographica Section D* **63**, 550-554 (2007).
150 D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W. & Marsh, M. Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallographica Section F* **70**, 1117-1126 (2014).
151 Shaw Stewart, P. D., Kolek, S. A., Briggs, R. A., Chayen, N. E. & Baldock, P. F. M. Random Microseeding: A Theoretical and Practical Exploration of Seed Stability and Seeding Techniques for Successful Protein Crystallization. *Crystal Growth & Design* **11**, 3432-3441 (2011).
152 Stura, E. A. & Wilson, I. A. Applications of the streak seeding technique in protein crystallization. *Journal of Crystal Growth* **110**, 270-282 (1991).
153 McPherson, A. in *EJB Reviews 1990* (eds P. Christen & E. Hofmann) 49-71 (Springer Berlin Heidelberg, 1991).
154 Khurshid, S., Haire, L. F. & Chayen, N. E. Automated seeding for the optimization of crystal quality. *Journal of Applied Crystallography* **43**, 752-756 (2010).
155 White, T. A. *et al.* Recent developments in CrystFELThis article will form part of a virtual special issue of the journal on free-electron laser software. *Journal of Applied Crystallography* **49**, 680-689 (2016).
156 Yefanov, O. *et al.* Accurate determination of segmented X-ray detector geometry. *Opt. Express* **23**, 28459-28470 (2015).
157 French, S. & Wilson, K. On the treatment of negative intensity observations. *Acta Crystallographica Section A* **34**, 517-525 (1978).
158 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular

crystallography. *Acta Crystallographica Section D* **66**, 12-21 (2010).
159 Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**, 283-291 (1993).
160 Marcus, R. A. & Sutin, N. Electron transfers in chemistry and biology. *Biochimica et Biophysica Acta (BBA) - Reviews on Bioenergetics* **811**, 265-322 (1985).
161 Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I. *The Journal of Chemical Physics* **24**, 966-978 (1956).
162 Chan, C. K., DiMagno, T. J., Chen, L. X., Norris, J. R. & Fleming, G. R. Mechanism of the initial charge separation in bacterial photosynthetic reaction centers. *Proceedings of the National Academy of Sciences* **88**, 11202-11206 (1991).
163 Tiede, D. M., Kellogg, E. & Breton, J. Conformational changes following reduction of the bacteriopheophytin electron acceptor in reaction centers of Rhodopseudomonas viridis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **892**, 294-302 (1987).
164 Woodbury, N. W. & Allen, J. P. in *Anoxygenic Photosynthetic Bacteria* (eds Robert E. Blankenship, Michael T. Madigan, & Carl E. Bauer) 527-557 (Springer Netherlands, 1995).
165 Orzechowska, A. *et al.* Coupling of collective motions of the protein matrix to vibrations of the non-heme iron in bacterial photosynthetic reaction centers. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1797**, 1696-1704 (2010).
166 Knapp, E. W. *et al.* Analysis of optical spectra from single crystals of Rhodopseudomonas viridis reaction centers. *Proceedings of the National Academy of Sciences* **82**, 8463-8467 (1985).