Faculty of Science

X-ray Free-Electron Laser Based Methods for Structural and Ultrafast Dynamics Studies of a Photosynthetic Reaction Centre

David Arnlund

Ph.D. thesis
Department of Chemistry and Molecular Biology
University of Gothenburg

UNIVERSITY OF GOTHENBURG

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE NATURAL SCIENCES

X-ray Free-Electron Laser Based Methods for Structural and Ultrafast Dynamics Studies of a Photosynthetic Reaction Centre

DAVID ARNLUND

UNIVERSITY OF GOTHENBURG

Department of Chemistry and Molecular Biology
Göteborg, Sweden
2014

Thesis for the Degree of Doctor of Philosophy in Natural Science

X-ray Free-Electron Laser Based Methods for Structural and Ultrafast Dynamics Studies of a Photosynthetic Reaction Centre

David Arnlund

Cover: Light-induced structural changes of the photosynthetic membrane protein reaction centre from *Bl. viridis* as determined by time-resolved wide-angle X-ray scattering at the Linac Coherent Light Source X-ray free-electron laser. Magnitude of movement enhanced for clarity.

ISBN 978-91-628-9235-7
Available online at http://hdl.handle.net/2077/37222

Department of Chemistry and Molecular Biology
Lundbergslaboratoriet
SE-405 30 Gothenburg
Sweden
Telephone: +46(0)31 – 786 00 00

Printed by Ale Tryckteam
Göteborg, Sweden, 2014

Till Åsa

Abstract

Life on earth is fuelled by the energy of sunlight, which must first be captured and converted into a chemical energy form useful to the cell. This process is known as photosynthesis and the major pathway of this energy conversion is *via* photosynthetic reaction centres. These enzymes convert the energy content of an absorbed photon into a transmembrane potential difference *via* the movements of electrons. Increasing our knowledge of the three-dimensional fold and structural changes that takes place within photosynthetic reaction centres is therefore of considerable importance for understanding biological photosynthesis.

The aim of this work has been to adapt methods for both crystallographic and solution phase structural studies of membrane proteins to the unique properties of X-ray free-electron laser (XFEL) radiation. To accomplish this, a new crystallization technique for the photosynthetic reaction centre from the purple bacterium *Blastochloris viridis* (RC_{vir}) was developed which was suitable for serial femtosecond crystallography (SFX) experiments at an XFEL. Our initial experiments at the Linac Coherent Light Source (LCLS), the world's first XFEL, yielded an SFX structure of RC_{vir} to 8.2 Å resolution. After the LCLS decreased the X-ray wavelength at which the facility could operate, and in combination with improved crystallization conditions, we later resolved the SFX structure of RCvir to 3.5 Å resolution.

Whether or not ultrafast structural changes in RC_{vir} occur in photosynthesis has been debated for two decades. We addressed this question by developing time-resolved wide-angle X-ray scattering (TR-WAXS) studies at the LCLS that could capture rapid structural changes in solubilized samples of RC_{vir}. Proof-of-principle experiments revealed a structural deformation that propagated through the RC_{vir} protein following multi-photon absorption by its cofactors, enabling a protein quake through a photosynthetic protein to be visualized. Further insight was provided by a second TR-WAXS experiment in which this structural signal was observed in the data as the pump laser fluence was decreased to less than one photon absorbed per RC_{vir} molecule. This result implies that, even under physiological conditions of normal sunlight, ultrafast protein structural rearrangements may influence the primary charge separation events of biological photosynthesis.

Contribution report

Paper I	I cultivated cells, purified protein, collected data, and took part in the analysis, writing of the manuscript and producing figures
Paper II	Produced cells and purified protein used for crystallization, collected data, and took part in the analysis, manuscript and figures production.
Paper III	I was responsible for the project, produced cells and required samples. I collected data, lead the data analysis, produced scripts, generated basis spectra and modelled the experimental and theoretical scattering data. I took a major part in writing the manuscript and produced figures.
Paper IV	I was responsible for the project, cultivated cells, produced protein and collected the data. I developed analysis tools for data handling and contributed significantly to the manuscript and figures preparation.

List of publications

Paper I Linda C Johansson, **David Arnlund,** Thomas A White, Gergely Katona, Daniel P DePonte, Uwe Weierstall, R Bruce Doak, Robert L Shoeman, Lukas Lomb, Erik Malmerberg, Jan Davidsson, Karol Nass, Mengning Liang, Jakob Andreasson, Andrew Aquila, Saša Bajt, Miriam Barthelmess, Anton Barty, Michael J Bogan, Christoph Bostedt, John D Bozek, Carl Caleman, Ryan Coffee, Nicola Coppola, Tomas Ekeberg, Sascha W Epp, Benjamin Erk, Holger Fleckenstein, Lutz Foucar, Heinz Graafsma, Lars Gumprecht, Janos Hajdu, Christina Y Hampton, Robert Hartmann, Andreas Hartmann, Günter Hauser, Helmut Hirsemann, Peter Holl, Mark S Hunter, Stephan Kassemeyer, Nils Kimmel, Richard A Kirian, Filipe R N C Maia, Stefano Marchesini, Andrew V Martin, Christian Reich, Daniel Rolles, Benedikt Rudek, Artem Rudenko, Ilme Schlichting, Joachim Schulz, M Marvin Seibert, Raymond G Sierra, Heike Soltau, Dmitri Starodub, Francesco Stellato, Stephan Stern, Lothar Strüder, Nicusor Timneanu, Joachim Ullrich, Weixiao Y Wahlgren, Xiaoyu Wang, Georg Weidenspointner, Cornelia Wunderer, Petra Fromme, Henry N Chapman, John C H Spence & Richard Neutze. Lipidic phase membrane protein serial femtosecond crystallography, *Nature methods* **9**:263-265 (2012)

Paper II Linda C. Johansson, **David Arnlund**, Gergely Katona, Thomas A. White, Anton Barty, Daniel P. DePonte, Robert L. Shoeman, Cecilia Wickstrand, Amit Sharma, Garth J. Williams, Andrew Aquila, Michael J. Bogan, Carl Caleman, Jan Davidsson, R. Bruce Doak, Matthias Frank, Raimund Fromme, Lorenzo Galli, Ingo Grotjohann, Mark S. Hunter, Stephan Kassemeyer, Richard A. Kirian, Christopher Kupitz, Mengning Liang, Lukas Lomb, Erik Malmerberg, Andrew V. Martin, Marc Messerschmidt, Karol Nass, Lars Redecke, M. Marvin Seibert, Jennie Sjöhamn, Jan Steinbrener, Francesco Stellato, Dingjie Wang, Weixaio Y. Wahlgren, Uwe Weierstall, Sebastian Westenhoff, Nadia A. Zatsepin, Sebastien Boutet, John C.H. Spence, Ilme Schlichting, Henry N. Chapman, Petra Fromme & Richard Neutze. Structure of a photosynthetic reaction centre determined by serial femtosecond crystallography, *Nature Communications* **4**:2911 (2013)

Paper III **David Arnlund**, Linda C Johansson, Cecilia Wickstrand, Anton Barty, Garth J Williams, Erik Malmerberg, Jan Davidsson, Despina Milathianaki, Daniel P DePonte, Robert L Shoeman, Dingjie Wang, Daniel James, Gergely Katona, Sebastian Westenhoff, Thomas A White, Andrew Aquila, Sadia Bari, Peter Berntsen, Mike Bogan, Tim Brandt van Driel, R Bruce Doak, Kasper Skov Kjær, Matthias Frank, Raimund Fromme, Ingo Grotjohann, Robert Henning, Mark S Hunter, Richard A Kirian, Irina Kosheleva, Christopher Kupitz, Mengning Liang, Andrew V Martin, Martin Meedom Nielsen, Marc Messerschmidt, M Marvin Seibert, Jennie Sjöhamn, Francesco Stellato, Uwe Weierstall, Nadia A Zatsepin, John C H Spence, Petra Fromme, Ilme Schlichting, Sébastien Boutet, Gerrit Groenhof, Henry N Chapman & Richard Neutze. Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser, *Nature methods* **11**:923-926 (2014)

Paper IV **David Arnlund**, Robert Dods, Despina Milathianaki, Kenneth Beyerlein, Peter Berntsen, Chelsie Conrad, Garret Nelson, Erik Malmerberg, Cecilia Wickstrand, Linda C. 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.*

Related publications

Paper V Sebastian Westenhoff, Erik Malmerberg, **David Arnlund**, Linda C. Johansson, Elena Nazarenko, Marco Cammarata, Jan Davidsson, Vincent Chaptal, Jeff Abramson, Gergely Katona, Andreas Menzel & Richard Neutze. Rapid readout detector captures protein time-resolved WAXS, *Nature methods* **7**:775-776 (2010)

Paper VI Erik Malmerberg, Petra H.M. Bovee-Geurts, Gergely Katona, Xavier Deupi, **David Arnlund**, Cecilia Wickstrand, Linda C. Johansson, Sebastian Westenhoff , Elena Nazarenko, Gebhard F.X. Schertler, Andreas Menzel, Willem J. de Grip & Richard Neutze. Conformational activation of visual rhodopsin in native disk membranes, *Science Signalling* (manuscript accepted, 2014)

Paper VII Anton Barty, Carl Caleman, Andrew Aquila, Nicusor Timneanu, Lukas Lomb, Thomas A. White, Jakob Andreasson, **David Arnlund**, Sasa Bajt, Thomas R. M. Barends, Miriam Barthelmess, Michael J. Bogan, Christoph Bostedt, John D. Bozek, Ryan Coffee, Nicola Coppola, Jan Davidsson, Daniel P. DePonte, R. Bruce Doak, Tomas Ekeberg, Veit Elser, Sascha W. Epp, Benjamin Erk, Holger Fleckenstein, Lutz Foucar, Petra Fromme, Heinz Graafsma, Lars Gumprecht, Janos Hajdu, Christina Y. Hampton, Robert Hartmann, Andreas Hartmann, Gunter Hauser, Helmut Hirsemann, Peter Holl, Mark S. Hunter, Linda Johansson, Stephan Kassemeyer, Nils Kimmel, Richard A. Kirian, Mengning Liang, Filipe R. N. C. Maia, Erik Malmerberg, Stefano Marchesini, Andrew V. Martin, Karol Nass, Richard Neutze, Christian Reich, Daniel Rolles, Benedikt Rudek, Artem Rudenko, Howard Scott, Ilme Schlichting, Joachim Schulz, M. Marvin Seibert, Robert L. Shoeman, Raymond G. Sierra, Heike Soltau, John C. H. Spence, Francesco Stellato, Stephan Stern, Lothar Struder, Joachim Ullrich, X. Wang, Georg Weidenspointner, Uwe Weierstall, Cornelia B. Wunderer & Henry N. Chapman. Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements, *Nature photonics* **6**:35-40 (2012)

Paper VIII Sébastien Boutet, Lukas Lomb, Garth J. Williams, Thomas R. M. Barends, Andrew Aquila, R. Bruce Doak, Uwe Weierstall, Daniel P. DePonte, Jan Steinbrener, Robert L. Shoeman, Marc Messerschmidt, Anton Barty, Thomas A. White, Stephan Kassemeyer, Richard A. Kirian, M. Marvin Seibert, Paul A. Montanez, Chris Kenney, Ryan Herbst, Philip Hart, Jack Pines, Gunther Haller, Sol M. Gruner, Hugh T. Philipp, Mark W. Tate, Marianne Hromalik, Lucas J. Koerner, Niels van Bakel, John Morse, Wilfred Ghonsalves, **David Arnlund**, Michael J. Bogan, Carl Caleman, Raimund Fromme, Christina Y. Hampton, Mark S. Hunter, Linda C. Johansson, Gergely Katona, Christopher Kupitz, Mengning Liang, Andrew V. Martin, Karol Nass, Lars Redecke, Francesco Stellato, Nicusor Timneanu, Dingjie Wang, Nadia A. Zatsepin, Donald Schafer, James Defever, Richard Neutze, Petra Fromme, John C. H. Spence, Henry N. Chapman & Ilme Schlichting. High-resolution protein structure determination by serial femtosecond crystallography, *Science* **337**:362-364 (2012)

Paper IX Lars Redecke, Karol Nass, Daniel P. DePonte, Thomas A. White, Dirk Rehders, Anton Barty, Francesco Stellato, Mengning Liang, Thomas R.M. Barends, Sébastien Boutet, Garth J. Williams, Marc Messerschmidt, M. Marvin Seibert, Andrew Aquila, **David Arnlund**, Sasa Bajt, Torsten Barth, Michael J. Bogan, Carl Caleman, Tzu-Chiao Chao, R. Bruce Doak, Holger Fleckenstein, Matthias Frank, Raimund Fromme, Lorenzo Galli, Ingo Grotjohann, Mark S. Hunter, Linda C. Johansson, Stephan Kassemeyer, Gergely Katona, Richard A. Kirian, Rudolf Koopmann, Chris Kupitz, Lukas Lomb, Andrew V. Martin, Stefan Mogk, Richard Neutze, Robert L. Shoeman, Jan Steinbrener, Nicusor Timneanu, Dingjie Wang, Uwe Weierstall, Nadia A. Zatsepin, John C. H. Spence, Petra Fromme, Ilme Schlichting, Michael Duszenko, Christian Betzel & Henry N. Chapman. Natively inhibited *Trypanosoma brucei* cathepsin B structure determined by using an X-ray laser, *Science* **339**:227-230 (2013)

Paper X Andrew Aquila, Mark S. Hunter, R. Bruce Doak, Richard A. Kirian, Petra Fromme, Thomas A. White, Jakob Andreasson, **David Arnlund**, Saša Bajt, Thomas R. M. Barends, Miriam Barthelmess, Michael J. Bogan, Christoph Bostedt, Hervé Bottin, John D. Bozek, Carl Caleman, Nicola Coppola, Jan Davidsson, Daniel P. DePonte, Veit Elser, Sascha W. Epp, Benjamin Erk, Holger Fleckenstein, Lutz Foucar, Matthias Frank, Raimund Fromme, Heinz Graafsma, Ingo Grotjohann, Lars Gumprecht, Janos Hajdu, Christina Y. Hampton, Andreas Hartmann, Robert Hartmann, Stefan Hau-Riege, Günter Hauser, Helmut Hirsemann, Peter Holl, James M. Holton, André Hömke, Linda Johansson, Nils Kimmel, Stephan Kassemeyer, Faton Krasniqi, Kai-Uwe Kühnel, Mengning Liang, Lukas Lomb, Erik Malmerberg, Stefano Marchesini, Andrew V. Martin, Filipe R.N.C. Maia, Marc Messerschmidt, Karol Nass, Christian Reich, Richard Neutze, Daniel Rolles, Benedikt Rudek, Artem Rudenko, Ilme Schlichting, Carlo Schmidt, Kevin E. Schmidt, Joachim Schulz, M. Marvin Seibert, Robert L. Shoeman, Raymond Sierra, Heike Soltau, Dmitri Starodub, Francesco Stellato, Stephan Stern, Lothar Strüder, Nicusor Timneanu, Joachim Ullrich, Xiaoyu Wang, Garth J. Williams, Georg Weidenspointner, Uwe Weierstall, Cornelia Wunderer, Anton Barty, John C. H. Spence, and Henry N. Chapman. Time-resolved protein nanocrystallography using an X-ray free-electron laser, *Optics Express* **20**:2706-2716 (2012)

Table of Contents

Abbreviations

AMO	Atomic, Molecular and Optical Science
BChl	Bacteriochlorophyll *a*
BPhe	Bacteriopheophytin *a*
CSPAD	Cornell-SLAC Pixel Array Detector
CXI	Coherent X-ray Imaging
LCLS	Linac Coherent Light Source
LCP	Lipidic cubic phase
LSP	Lipidic sponge phase
LDAO	Lauryldimethylamine-N-oxide
MD	Molecular dynamics
P_{960}	Special pair in RC_{vir}
PDB	Protein data bank
PEG	Polyethylene glucol
PSI, PSII	Photosystem I, II
Q_A, Q_B	Ubiquinone
QH_2	Ubiquinol
RC_{vir}	Reaction centre from *Bl. viridis*
SAXS	Small-angle X-ray scattering
SFX	Serial femtosecond crystallography
WAXS	Wide-angle X-ray scattering
XFEL	X-ray free-electron laser
Å	Ångström (=10^{-10} m)

1. Introduction

1.1 The biological membrane and membrane proteins

In order to separate environments inside and outside of the cell, a barrier is needed. Nature has solved this by the formation of a biological membrane that effectively keeps unwanted molecules outside the cell, and essential molecules inside creating an environment that could not be obtained if a permeable barrier surrounded the cell. These biological barriers can be further compartmentalized into the sub-cellular organelles such as chloroplasts or mitochondria and the two major building blocks of this barrier are lipids and membrane proteins.

Lipids consist of a polar head group with nonpolar hydrocarbon tails. By connecting tail to tail the lipid double layer in the biological membrane forms a hydrophilic exterior and a hydrophobic interior, preventing ions and other polar solutes from crossing the bilayer. This enables a charge- and concentration gradient to be maintained over the membrane.

Membrane proteins are the communicators, importers and exporters, docking stations and signal propagators of the cell[1]. Each and every protein is comprised of a number of amino acids, of which there only exist twenty different kinds. Despite this limited pool of building blocks the diversity of protein functions is vast. The order in which the amino acids are placed when a protein is synthesized determines the three-dimensional fold of that protein and subsequently its function. The linear string of amino acids orders into secondary structures (α-helices and β-sheets) that can pack together and form more complex tertiary and quaternary structures. A prerequisite of all membrane proteins is that at least a part of its surface is hydrophobic in order to be integrated with the nonpolar part of the lipid bilayer. An integral membrane protein spanning across the lipid bilayer would require more hydrophobic surfaces than a peripheral membrane protein that is loosely attached.

The three-dimensional fold of the membrane protein enables it to be very specific in its role, for example only allowing a specific substrate to be transported or binding to a particular signalling substance. Many membrane proteins carry out essential tasks necessary for cell survival and modifications or malfunctions can be devastating to the cell, hence they are important as both academic as well as pharmaceutical targets[2-4].

1.2 Photosynthesis

Photosynthesis is of paramount importance for life on earth, converting sunlight to chemical energy used by essentially all life forms. Photosynthesis can be divided into oxygenic and anoxygenic, where the former is the conversion of water and carbon dioxide into carbohydrates and molecular oxygen mainly carried out by higher plants and algae. Anoxygenic photosynthesis can be carried out by bacteria and unlike the oxygen version requires only one out of the two types of photosynthetic reaction centres (type I utilizing [Fe_4S_4] clusters and type II quinone molecules as terminal acceptors)[5]. There is little difference in terms of the photosynthetic machinery driving the process between bacteria and eukaryotes. They are proposed to have to a common ancestor[6-8] and the main hypothesis is that photosynthetic eukaryotes acquired their ability to conserve sunlight through endosymbiosis with cyanobacteria[9, 10]. In both plants and bacteria a light-harvesting antennae captures the incoming light and a reaction centre in the inner core is responsible for the charge separation. This process generates energy that is subsequently used for the synthesis of adenosine triphosphate (ATP)[11], the main chemical energy carrier within the cells. In eukaryotes photosynthesis is carried out in a specific organelle, the chloroplast, whereas in bacteria, which lack organelles, it is carried out in the intracytoplasmic membrane[12].

Photosystem I and II (PSI and PSII) are two of the most studied membrane proteins. In PSI a pair of chlorophyll *a* molecules act as primary donor, a chlorophyll monomer as primary acceptor and a phylloquinone as a secondary acceptor[13]. The [Fe_4-S_4] cluster in the type I reaction centre PSI is functional in the electron transfer. Structural studies of PSI[14] has revealed that the reaction centre apparatus of heliobacteria and green sulphur bacteria resembles PSI[15], whereas the reaction centre of PSII is related to the well-characterized reaction centre of purple bacteria[16-18]. In PSII the oxidation of water to molecular oxygen takes place in the oxygen-evolving complex (OEC), consisting of a cluster of four manganese ions and a calcium ion[19]. Water-splitting in the OEC also requires one chloride ion as well as a specific protein environment[19, 20]. In order for PSII to oxidize water, a potential of up to 1V must be generated[21] and PSII follows a four-flash pattern[22] where the special pair $P_{680}+$ is coupled to a redox-active tyrosine (Tyr_Z) which links $P_{680}+$ to the OEC, and in cooperation with a histidine on subunit D1, facilitates the release of protons to the aqueous phase[21, 23]. The reaction centres of purple bacteria such as *Blastochloris viridis* and *Rhodobacter sphaeroides* are to a high degree structural homologues of PSI and the D1 and D2 subunits of PSII[24, 25] and can therefore be used as model systems for the more complex oxygenic photosynthetic systems, despite their lack of water-splitting capabilities.

1.3 Structural studies of membrane proteins

Membrane proteins are amphiphilic in their nature since they must have both a hydrophobic surface that can be incorporated into the biological membrane as well as a more hydrophilic part where interaction with the polar intra- or extracellular environment takes place. This double nature of membrane proteins makes them difficult targets for structural studies and the first structure to be solved was the that of the reaction centre of *Blastochloris viridis* in 1984[26], a work that four years later was rewarded with the Nobel Prize[27]. The majority of all structures are solved by using X-ray crystallography and since membrane proteins are inheritably difficult to crystallize, most of the structures are from globular proteins[28].

1.3.1 Evolution of X-ray sources

In the early years of macromolecular X-ray crystallography experiments the radiation was generated by sealed-tube instruments based on the Coolidge design[29]. These X-ray sources were sufficiently brilliant to solve the structures of small chemical molecules, but they lacked the power for protein crystallography[30]. If the power load was increased above a critical threshold, the anode material would melt[31] and this lead to the invention of the rotating anode[32] that helped to dissipate the heat. This made more powerful X-rays possible, but it was not until the advent of synchrotrons, first proposed independently by Veksler and McMillian[33, 34], that X-ray based science really took off; though the first biological experiment was not conducted until 1970[35]. The demand of these first generation synchrotrons steadily grew and a key pioneering advancement was the storage ring that allowed particles to circulate for long periods of time, enabling more stable beam conditions. A second generation synchrotron was then constructed in Daresbury, UK, which was the first dedicated synchrotron light source in the world and opened for user experiments in 1981[36]. Strong magnets, so called insertion devices, were developed that could be placed around the storage ring with the effect that beam brilliance at the sample position increased and intensity peaks with adjustable wavelength could be employed. Synchrotrons with insertion devices became the third generation machines and the first one to open its doors was the European Synchrotron Radiation Facility (ESRF)[37]. There are now more than 50 dedicated second and third generation synchrotrons in the world and to date more than 70% of all solved protein structures have been by the use of synchrotron radiation[30].

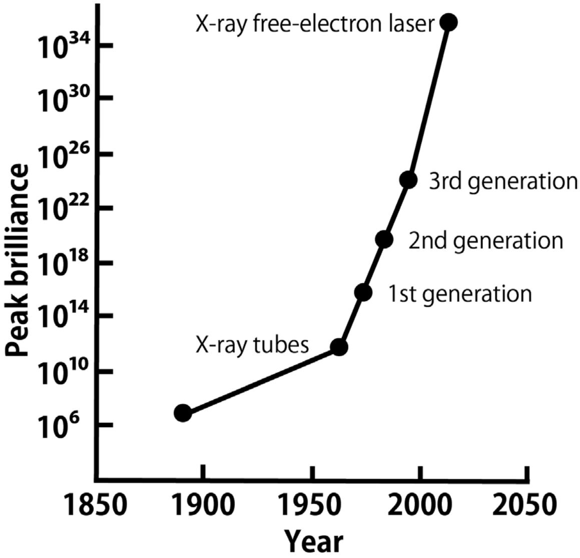

Figure 1. *Illustration of the dramatically increased X-ray peak brilliance as particle accelerators were started to be used as X-ray radiation sources. X-ray free-electron lasers produce X-rays with peak brilliance nine magnitudes higher than a 3rd generation synchrotron.*

A limitation with X-ray radiation from synchrotron sources is that the X-ray induced damage on the material limits the resolution of structural studies on biomolecules[38]. While cryogenically cooled protein crystals can withstand X-ray doses of up to 30 MGy[39, 40] suggestions arose that this dose barrier could be further extended by using extremely short X-ray pulses[41, 42]. Molecular dynamics simulations showed that femtosecond X-ray pulses generated by an X-ray free electron laser (XFEL) could be utilized to collect usable diffraction data of the sample before it exploded due to the intense X-ray brilliance, coining the "diffraction before destruction" principle[43]. The first hard X-ray free electron laser to be constructed was the Linac Coherent Light Source (LCLS)[44] which achieved its first lasing in 2009. This revolutionary X-ray source can produce X-rays with peak energies that are a billion times stronger than those achievable with third generation synchrotrons. Because the pulse duration at the LCLS is extremely short (only a few tens of femtoseconds[45]) the "diffraction before destruction" principle could be experimentally proven by collecting diffraction data from nanometre-sized protein crystals before radiation damage onset[46]. In addition to having short and intense X-ray pulses, the X-ray beam is extremely focused[47] (down to 0.1 µm) providing a method where sub-micron sized crystals, otherwise unsuitable for conventional data collection at a synchrotron, can be used by injecting a crystal containing solution into the X-ray beam, allowing one shot per crystal to be obtained. For time-resolved experiments the maximum time-resolution obtainable at an XFEL is in the femtosecond regime, compared to the ~100 picosecond limit at synchrotrons[48]. This allows for exciting experiments not previously conceivable. The Japanese XFEL, SACLA, has since opened[49] and several XFELs have begun construction and will be operational within the next few years[50, 51].

1.3.2 Protein crystallography

In 1912, Friedrich and Knipping, assistants of Max von Laue, used a crystal as a diffraction grating, giving direct proof for the existence of lattices in crystals and for the wave nature of X-rays, experimentally proving von Laue's theory[52]. Due to this, we can get a visual look of proteins in a three-dimensional way, something that is not possible with sources providing radiation with longer wavelength than X-ray. Since proteins are nanometre sized and bond distances between individual atoms are in the order of ~2 Å, X-rays with low Angstrom wavelengths are needed to resolve the detail of protein molecules. Protein crystallography is by far the most successful technique for solving protein structures, with over 90,000 structures deposited into the Protein Data Bank[53]. The most common method within protein crystallography is to produce the target protein and crystallize it in its ground state and subsequently solve that structure. Since proteins are large biomolecules their crystals diffract X-ray beams much less than do crystals of small molecules and as proteins consist mainly of C, N and O, which are light elements and scatter less then heavier atoms, protein crystallographers prefer to have a high intensity source.

1.3.3 Conformational dynamics

A conformational change is defined as a protein changing from one energetically favourable state to another and can be as large as entire subunits moving several Angstroms, or as subtle as one amino acid rotating its side chain. The causes for a conformational change can be many: ligand binding, pH change, temperature increase or absorption of a photon to name a few. The steps leading from one conformation to another are usually very well coordinated and time efficient, hence investigating the molecular machinery involved in conformational dynamics is of major importance for our understanding of how proteins work and how they have evolved over time.

For crystallographic studies of conformational dynamics methods such as low-temperature trapping[54-61] and time-resolved Laue diffraction have successfully been applied to characterize several different systems in the pico to millisecond regime such as photoactive yellow protein[62-68], carbon monoxide in complex with myoglobin[69-74] or haemoglobin[75], and photosynthetic reaction centre[76, 77]. In low-temperature trapping an intermediate population is increased prior to the collection of the crystallographic data. Complimentary techniques can be used to assess the intermediate population such as Raman[78-80] and optical[81] spectroscopy and the intermediate can then be kinetically trapped by flash freezing the crystal. Time-resolved Laue diffraction has the advantage of not requiring freeze-trapping of long-lived intermediates as one collects the data at ambient temperatures. The reaction is instead invoked *in situ* and afterwards probed with a polychromatic X-ray pulse rather than a monochromatic X-ray pulse generally used for protein crystallography. This increases the X-ray flux on the crystal as well as enabling full reflections to be recorded without rotation of the crystal.

A structure solved with protein crystallography gives a snap shot of a protein fixed in a crystal lattice. The natural environment for proteins is not however in crystal form, they are dynamic molecules often undergoing structural rearrangements as they perform their various tasks. It has also been shown that pH, high salt and other crystallization conditions could potentially negatively affect the reaction conditions[82]. An argument could be made that the movements seen in time-resolved crystallography are the structural rearrangements permitted within the crystal and that they do not fully describe the conformational changes. Large-scale structural changes, such as subdomain or helical movements, could potentially break the crystal contacts between molecules in the crystal lattice, severely limiting the diffraction power of that crystal. As such solution based X-ray methods has evolved complementing the information one can extract from crystallography.

Time-resolved small- and wide-angle X-ray scattering (SAXS and WAXS) are two closely related techniques that have emerged as tools for studying protein conformation dynamics in solution. SAXS probes for structural information in the small angle (~0.3 – 4°) area of the diffuse scattering pattern providing low-resolution data (~20 – 250 Å) related to general shape and size of the measured molecule. Producing a reliable SAXS profile requires relatively small sample consumption. Protein complexes and large multimeric systems have been characterized with SAXS. Sample delivery systems and strategies for data collection have improved permitting automatic sample characterization and time-resolved studies[83]. There have been many advances on the software side of SAXS and several tools are now publicly available for structural reconstruction of the one-dimensional SAXS profile[84-86]

As one goes to wider angles (~3 – 20°) of the diffuse scattering pattern (Figure 2) information about the secondary and tertiary fold of proteins can be probed[87], but this regime scatters X-rays to a much smaller extent (1 to 3 orders of magnitude lower[88]) than does the smaller angles and thus WAXS requires higher concentration of the protein sample. Improvements of X-ray sources, in particular the availability of third-generation synchrotrons, have provided the X-ray brilliance required for collecting WAXS data comparable to a resolution of ~2.5 Å[87]. The recent advent of the X-ray free electron laser dramatically improves the time required for statistically reliable time-resolved WAXS data as both the pulse flux and repetition rate are considerably higher.

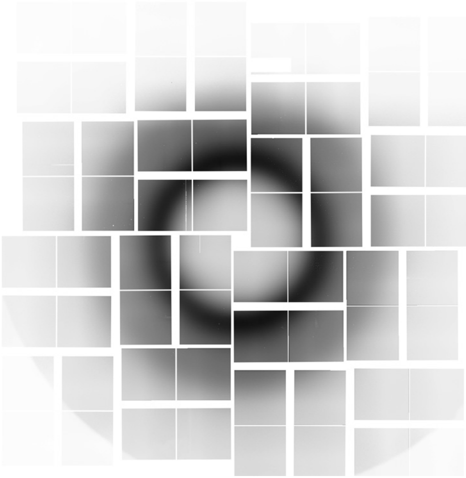

Figure 2. Diffuse scattering pattern collected on a CSPAD detector.

Originally time-resolved WAXS was used to study transient intermediates in heavy-atom containing small molecules that could be photo-activated[89-92]. Distinguishing between the background solvent heating and the signal related to the structural changes of the probed molecule was a fundamental challenge in these early experiments[93, 94] and it was not until 2008 that the first conformational changes in a biological system was published[95], making WAXS a less mature technique than SAXS. After this initial report on carbon monoxide dissociation from haemoglobin, showing that the scattering changes observed and those predicted from crystallographic structures were consistent with each other, several time-resolved studies using WAXS have been performed on protein systems, including membrane proteins[96-104].

1.4 Structure and function of reaction centre from *Bl. viridis*

The first membrane structure to be solved was the reaction centre from the purple bacterium *Blastochloris viridis* (henceforth referred to as RC_{vir}), using X-ray crystallography[26, 105]. The structure revealed that RC_{vir} spans the membrane and consists of four subunits: L, M, C and H. Associated to the protein are also a number of cofactors: the special pair P_{960}, bacteriopheophytin, bacteriochlorophyll *b*, menaquinone A and ubiquionone B, a non-heme iron and four hemes.

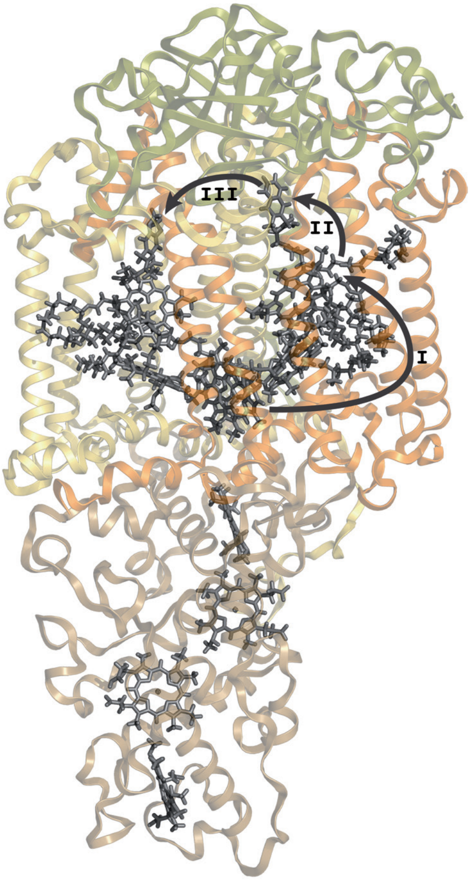

***Figure 3**. Light-driven charge separation reactions in RC_{vir} initiates with an electron transferring from P_{960} to the BPhe (I) followed by a transfer to Q_A (II) and subsequently to Q_B (III). This process is repeated twice to yield a QH_2 that gets released into the membrane from the Q_B site. Only the L-branch is active in the electron transfer.*

In the membrane, RC_{vir} is surrounded by a single type of light-harvesting complex (LH1) that captures incoming light and transfers the energy to the reaction centre, resulting in a charge separation between P_{960} and menaquinone A. This is followed by an electron transfer to the mobile ubiquinone B and during this transfer the $P_{960}+$ is re-reduced by the cythochrome *c* subunit. Re-reduction of the $P_{960}+$ enables a second charge separation event to proceed in identical fashion transferring another electron to the ubiquinone B, which then picks up two protons and gets released as QH_2 into the membrane[106], schematically shown in Figure 3. In RC_{vir} only the L-branch is active in the electron transfer[21] (Figure 3). The cytochrome *c* subunit is then re-reduced by a cyclic electron flow *via* cythochrome *bc*$_1$ and the soluble carrier *c*$_2$ and this coupled electron transfer leads to a net transmembrane potential change of two protons per photon absorbed (Figure 4). This potential difference is later utilized to produce ATP[107].

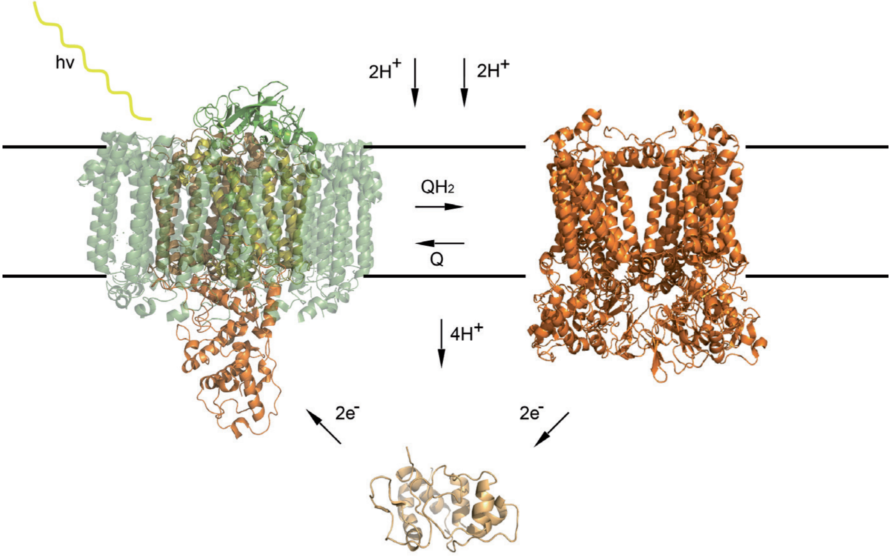

Figure 4. Electron flow and proton pumping presented schematically. Light activation of reaction centre surrounded by light harvesting complex I (left) releases ubiquinol (QH$_2$*) into the membrane that diffuses to the cytochrome bc*$_1$ *complex (right) where it is re-oxidized to quinone (Q) with accompanying proton depositions to the periplasmic side. The soluble carrier cytochrome c*$_2$ *re-reduces the cytochrome c subunit of the reaction centre.*

1.5 Scope of the thesis

The aim of this thesis has been to develop novel methods for structural studies on the photosynthetic reaction centre using the X-ray free electron laser at the LCLS.

The first step was to produce micron sized crystals of RC$_{vir}$ in large enough quantities to perform a proof-of-principle experiment at the Linac Coherent Light Source (LCLS) X-ray free-electron laser (XFEL). RC$_{vir}$ was crystallized with a lipidic sponge phase (LSP) batch method and injected into the beam using a gas-driven virtual nozzle (GDVN) system. **Paper I** describes how a low-resolution structure at 8.2 Å was obtained from 265 diffraction patterns.

As the LCLS developed and increased the X-ray energies deliverable by the machine, a high-resolution data set could be collected at the Coherent X-ray Imaging (CXI) beamline. The LSP batch method was optimized and production and purification of RC_{vir} scaled up to meet the requirements of a high-resolution 12-hour shift experiment. The crystals yielded were shown to diffract to 2.8 Å and a structure solved to 3.5 Å is presented in **Paper II**, the first side-chain resolution structure of a membrane protein solved with XFEL radiation.

The next goal was to perform a time-resolved experiment on RC_{vir} and for the first time probe sub-picosecond structural movements using the time-resolved wide-angle X-ray scattering (WAXS) method. These experiments also extended the technique from synchrotrons to show that it could be used at an XFEL. Data were collected at the CXI beamline using the GDVN microjet system to inject solubilized RC_{vir} non-crystalline samples to the XFEL beam aligned with a Ti:Sa femtosecond laser system. In **Paper III**, an ultrafast global conformational change arising within picoseconds after multi-photon laser excitation and preceding the propagation of heat through the protein is presented, providing direct structural evidence of a protein quake as a way for RC_{vir} to rapidly dissipate the energy deposited. This also proved that time-resolved WAXS could be used as a technique to probe ultra-fast structural changes not observable at a synchrotron facility.

To provide new insight on the topic of whether or not ultra-fast structural movements in RC_{vir} occur upon single-photon absorption, a second time-resolved WAXS experiment was performed at the CXI beamline, again using the GDVN microjet system for sample delivery. The XFEL beam was aligned on a Ti:Sa femtosecond laser system and power titration data on solubilized RC_{vir} samples were collected as well as low-fluence datasets. The protein and heating signals were shown to decay differently and a single-photon light-induced structural change could be observed. This is presented in **Paper IV** together with a novel data rejection routine developed to correct for the X-ray shot-to-shot fluctuations observed at the XFEL experiments.

2. Methodology

2.1 Production and purification of reaction centre from *Bl. viridis*

After the initial detailed characterization in 1966[108], extensive studies on the purple, non-sulphur bacterium *Blastochloris viridis* showed that the photosynthetic membranes were tightly packed and lamellar in nature with each cell containing one large stack of membranes, similar to what has been observed in higher plants[109]. Due to the high abundance of photosynthetic membranes in *Bl. viridis* and reaction centre in each membrane, RC_{vir} can be harvested directly from the natural source without any need for genetically modified overexpression.

2.1.1 Cell growth

Bl. viridis can be grown under both anaerobic and microaerophilic conditions, provided that the cells are exposed to light in the former and darkness in the latter case[110]. Under microaerophilic growth of *Bl. viridis* synthesis of bacteriochlorophyll *b*, cytochrome c_2, cytochrome bc_1 complex and the reaction centre is suppressed but respiratory function is still maintained. Under anaerobic, photosynthetic growth, the bacteriochlorophyll *b* concentration (sum of the LH_1 complex and RC_{vir}) is inversely proportional to the intensity of the light exposed to the cells[110].

2.1.2 Protein purification

The first step towards producing a pure membrane protein sample from a cell is to disrupt the cell wall in order to expose the inner, protein containing membrane. This can be done mechanically with a sonicator, X-press or French press[111] or by enzymatic means, typically using lysozyme[112, 113]. After centrifugation the cell wall fragments are pelleted and removed and the membrane containing supernatant can be collected and subsequently pelleted with ultracentrifugation. Since membrane proteins have both a hydrophilic part extending from, and a hydrophobic part integrated into the biological membrane, detergents must be used to keep the protein folded. A detergent is a small, amphiphilic molecule with a polar head group connected to a hydrophobic tail, which arrange in micelles that can surround the target membrane protein. The critical micelle concentration, the CMC, is a measurement of how many detergent molecules that are needed to form a micelle and it is crucial that all buffers used during the protein purification are above this CMC to prevent the membrane protein from precipitating.

There is a large number of commercially available detergents with varying polar groups and hydrocarbon tail length. To determine which detergent is suitable for a certain target, a detergent screen can be performed. For RC_{vir} lauryldimethylamine-N-oxide (LDAO), which is a zwitterionic compound, has been used for both purification and crystallization. In order to remove lipids and other proteins solubilized after detergent addition a number of chromatography steps are undertaken. To produce pure RC_{vir} protein the solubilized sample is passed over an ionic exchange column followed by size exclusion chromatography after which the purity of RC_{vir} is assessed by calculating the absorbance ratio between 280 nm and 830 nm. For WAXS samples, fractions from the size exclusion column with a ratio below 3 are pooled, concentrated and flash-frozen whereas the ratio limit for crystallization grade RC_{vir} needs to be below 2.6.

2.2 Protein crystallography

2.2.1 Crystallization

To obtain data to solve the structure using X-ray sources the protein must first be crystallized. When crystallized, proteins arrange in a repetitive manner, forming a three-dimensional lattice where the smallest building block, the unit cell, is repeated by translation throughout the entire crystal. The unit cell can be divided further into the asymmetric unit, which through rotational and translational movement according to the symmetry operators can describe the unit cell.

Finding optimal conditions under which the target protein crystallizes is an iterative process with crystallization trials performed, evaluated and optimized until crystals are obtained. Further optimization to improve the resolution of the crystal diffraction is usually needed. Going from protein production to a well-diffracting crystal can take a very long time and is usually referred to as the bottleneck in protein crystallography.

For the most common method of protein crystal growth, the vapour diffusion approach, precipitant and protein drops are mixed and left to equilibrate over a reservoir solution containing the same precipitant. Water will evaporate from the protein drop as the reservoir has higher osmolarity causing an increase in protein and precipitant concentration in the drop. Small nuclei in the drop will form as it reaches the nucleation zone and this reduces the free protein concentration and the drop goes into the metastable zone where no more nucleation sites will form but crystals will grow (Figure 5). An important aspect is to have the correct concentration of protein and precipitant since too low concentration will make the drop enter the under saturated zone where the protein will remain in solution yet if the concentration of the protein or precipitate is too high the precipitation zone

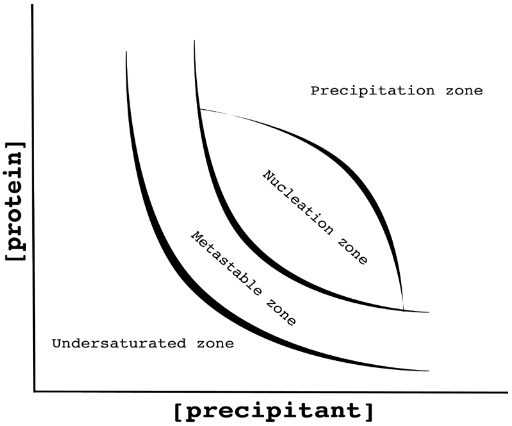

Figure 5. Protein crystallography solubility diagram. For crystal nucleation to occur the nucleation zone must be reached upon which the free protein concentration will decrease and the drop will move into the metastable zone where protein crystals will grow but no additional nucleation sites will form.

will be reached where the protein forms an amorphous precipitate lacking the order needed for X-ray crystallography experiments. Key factors for successful protein crystal growth are purity and concentration of the target protein, pH, temperature, salt, precipitant choice and the ratio between the protein and precipitant.

For membrane proteins other techniques have been developed for crystallization that attempt to mimic the membrane environment that the proteins are stable in. A method that has been successful for a wide variety of membrane proteins such as reaction centre[114], oxidases[115, 116], different rhodopsins[117-123], G-protein coupled receptors from seventeen families[124-140] and several other proteins[141-145] is the lipidic cubic phase (LCP) crystallization method first described in 1996[117]. The membrane protein is reconstituted into monoolein in a 60:40 w/w ratio. By mixing with a pair of syringes a transparent, highly viscous LCP is formed after which it is transferred to a sandwich plate or vial. Precipitate solution is added and the LCP is left to equilibrate. LCP crystallization can only be performed at room temperature since the LCP below 20°C can turn into a solid lamellar crystal phase (L_C) not suitable for crystallization[146]. The proposed mechanism for crystal growth is that it proceeds *via* a salt-induced lamellar phase in which the protein can diffuse laterally, nucleate and form a crystal[147, 148]. A major breakthrough for LCP crystallization was the introduction of LCP-robots for high-throughput crystallization trials and the development of micro focus beamlines at synchrotrons enabling diffraction data collection from the increasingly small crystals[149].

The crystallization method used in **Paper I** and **II** presented in this thesis is the lipidic sponge phase (LSP) method that is derived from the LCP crystallization technique. It has been showed to be successful for a number of membrane proteins such as light-harvesting complex II[148], cobalamin transporter BtuB[150] and reaction centres[77, 151, 152]. It has also been proposed that several reported LCP structures with crystals obtained from conditions with ~40 % PEG-400 proceed *via* an LSP[153]. Swelling an LCP forms the LSP, which is done by adding an extra component to the LCP. The component added can be polyethylene glycol (PEG), 2-Methyl-2,4-pentanediol (MPD), Jeffamine M600, dimethyl sulfoxide (DMSO), N-methyl-α-pyrrolidone (NMP), propylene glycol (PG) or ethanol[154] and depending on which one is used the degree of swelling varies. From a phase diagram one can deduce how much is needed to form an LSP[154]. Since the aqueous pore diameter increases in an LSP it allows for membrane proteins with large hydrophilic domains to be incorporated. As for LCP the LSP crystallization method is limited to room temperature since the LSP will form a lamellar crystalline phase below 20 °C[155]. Additionally, pH must be below 9.0 due to hydrolysis of the monoolein ester bonds occurring at higher pH[156].

A practical difference between LSP and LCP is that the protein is not reconstituted into monoolein prior to crystallization. Preparation of the LSP is done by mixing buffer, monoolein and precipitant solution after which the LSP is left to equilibrate at ~37 °C until a homogenous, non-birefringent transparent phase forms. LSP is a viscous liquid and can thus be used for standard vapour diffusion setups, used in vials or sandwich plates. LSP-grown crystals can be harvested without addition of a lipase to dissolve the lipid phase, which is sometimes needed in LCP crystallization[157].

2.2.2 X-ray diffraction from a protein crystal

As X-rays interact with an electron cloud around an atom the electrons begin to oscillate causing them to emit new radiation, a diffracted X-ray photon. For a crystal to diffract X-rays the so-called Laue conditions must be met which state that reciprocal lattice indices *h, k, l* must be an integer for constructive interference to occur, often referred to as Bragg's law

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

where n is the order of diffraction, λ the X-ray wavelength, d the spacing between lattice planes and θ the angle between the lattice plane and the incoming beam (Figure 6).

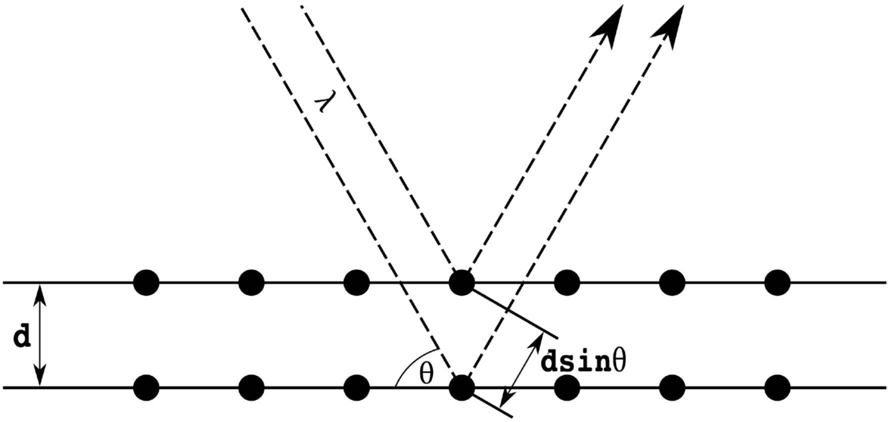

Figure 6. *Bragg's law illustrated by two incident beams reflected as Bragg's law is fulfilled.*

Bragg reflections in the diffraction pattern are reciprocally related to the unit cell symmetry and dimensions and not to the actual content of the unit cell. At a certain angle only a small number of Bragg reflections are produced, hence different rotational angles of the crystal are collected to cover the entire reciprocal space.

The measured intensities from the Bragg reflections contain information about the unit cell content but lack information about the phases required for calculating the electron density of the unit cell and thus the individual atomic positions. Fourier transformation of the structure factor describes the electron density as follows:

$$\rho(x,y,z) = \frac{1}{v}\sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx+ky+lz)+i\alpha(hkl)} \qquad (2)$$

where $\rho(x,y,z)$ is the electron density at position x,y,z in the unit cell, v the volume of the unit cell, *h, k, l* the Miller indices, F_{hkl} the structure factor and $\alpha(hkl)$ the phase information not extractable from the diffraction pattern. Since $\alpha(hkl)$ cannot be measured directly this is referred to as the phase problem of X-ray crystallography.

2.2.3 Structural determination in protein crystallography

The first step in determining the structure of the protein that the diffraction data is collected from is to index the reflections yielding their positions and *h, k, l* indices, followed by determination of the crystallographic space group and integration of spot intensities. To correct for parameters such as radiation damage, diffraction power, crystal size and difference in X-ray path through the crystal, the data is scaled after which the integrated intensities are combined to structure factors. These actions are done computationally with various software, e.g. XDS[158] or CCP4[159].

As mentioned in section 2.2.1 the phase information is not available in a diffraction pattern, and to calculate the electron density in the unit cell this information must be retrieved. There are various methods for this, such as single- and multiple anomalous diffraction (SAD and MAD)[160] or single- and multiple isomorphous replacement (SIR and MIR)[161]. For high-resolution data of smaller proteins *ab initio* methods for determining the phases can be adopted (Shake-and-bake algorithm)[162, 163]. The most common method used for obtaining the phases and solving the structure is the molecular replacement method where phases from a similar structure are used to solve the phases of the target protein using the Patterson function

$$\mathrm{P}(u,v,w) = \frac{1}{v}\sum_h \sum_k \sum_l |F_{hkl}|^2 \, e^{-2\pi i(hu+kv+lw)} \qquad (3)$$

Correlations between the Patterson maps of the model and target protein guides the superpositioning of the phase model on the target structure in the unit cell and by rotational and translational search the initial phases can be obtained. Solving the phases for a protein macromolecule can be computationally done with Phaser, a program in the CCP4 software suite. An electron density map is retrieved by Fourier transformation of the experimental data and the solved phases into which a model of the target protein can be built with software such as COOT[164]. The REFMAC[165] or PHENIX[166] program can subsequently be utilized to refine and rebuild the model until it converges. A measurement of how well the model agrees with the experimental data is given by the crystallographic R-factor

$$R = \frac{\sum_{hkl}\left||F_{obs}(hkl)| - k|F_{calc}(hkl)|\right|}{\sum_{hkl}|F_{obs}(hkl)|} \qquad (4)$$

where a perfect agreement would give an R-factor value of 0 since F_{obs} and F_{calc} would be identical. Typically the R-factor is presented as a percentage and should not exceed ten times the resolution the data is refined to. To reduce model bias and cross-validate the data the R_{free} value was introduced where 5-10 % of all observed reflections are excluded from the refinement. If the R_{free}-value shows a major deviation from the R-value the quality of the model is poor or the data is being over- interpreted.

Using molecular replacement to solve the phase problem has the inherent risk of introducing model bias. To analyse to what extent this is true an omit-map is calculated. Here, parts of the model are removed prior to calculating the electron density map and if the resulting map does not show the whole model, including those omitted, model bias is likely. An automated approach is to use CNS[167], which omits different parts of the structure for each calculation and then combines them into a composite omit map.

Another validation of the structure is the Ramachandran plot where the dihedral angles θ and φ of the backbone are plotted against each other. Only certain combinations of θ and φ are allowed so a Ramachandran plot can detect amino acids with disfavoured conformations.

With the advent of serial femtosecond crystallography (SFX) data can be collected on crystals with a size in the micro- to nanometre range since the X-ray intensities provided by an X-ray free-electron laser are exceptionally high. In SFX, data is collected from each crystal only once, giving rise to partial reflections since no rotation of the crystal can be done. To deal with this new approach of collecting crystallographic data novel software such as CrystFEL[168] and Cheetah[169] have been developed, where the former is designed to handle partial reflection recordings for indexing and the latter coping with the enormous amount of data generated during an SFX experiment.

Indexing in CrystFEL is done with the routine Indexamajig, which calls upon other indexing programs (DirAx[170] and MOSFLM[171]), using a list with files where a peak search is performed and sent for auto-indexing. From the predicted peak locations given by CrystFEL intensities are integrated and merged using a Monte Carlo approach[172]. The mean value of all measurements is taken and over the 3D reflection profile a value proportional to the integrated intensity is given by a Monte Carlo integration.

Quality assessments of an SFX data set differ from those applied to traditional crystallography. In SFX the data is separated into two sets (odds and evens), each merged independently, and the figure of merit (R_{split}) is calculated which measures the agreement between the two sets of intensities from the overall data set

$$R_{split} = \frac{1}{\sqrt{2}} \frac{\Sigma |I_{even} - I_{odd}|}{{}^{1}/_{2}\, \Sigma |I_{even} + I_{odd}|} \qquad (5)$$

where the factor $\sqrt{2}$ is introduced since the quality of the data is underestimated by that factor due to splitting the data in half. Signal to noise assessment (I/σ) for each *hkl*-index is calculated by

$$I_{hkl} = \langle I_{spot} \rangle \; ; \; \sigma_{hkl} = \frac{\sqrt{\sum_N \left(I_{spot} - I_{hkl}\right)^2}}{N} \qquad (6)$$

where I_{hkl} is defined as the mutual mean intensity of that reflection and σ as the root-mean-square deviation of spot intensity values from the average divided by the number of spots.

2.3 Time-resolved wide-angle X-ray scattering

2.3.1 X-ray diffraction from a protein solution

Unlike diffraction from an ordered lattice of protein molecules in X-ray crystallography, diffraction from a protein in solution does not give rise to any Bragg reflections containing the sought information. Instead the randomly oriented molecules gives rise to a diffuse diffraction pattern as can be seen in Figure 2.

The scattered beam can be described to be the sum of all scattered waves from all atom pairs in the solution

$$\mathrm{F}(q) = \sum_{\boldsymbol{j}} f_j e^{\boldsymbol{iq \cdot r_j}} \qquad (7)$$

where $\mathrm{F}(q)$ is the amplitude of the scattered wave as a function of q, f_j the form factor and r_j the position vector of an atom j. As the scattering from a protein solution is the sum of all atoms in the system, including solvent molecules, it is useful to separate the contributions by defining a scattering density, ρ_s, for the solvent and describing the scattering vector r_j and form factor f_j of the protein separately. The scattering contribution can then be described with three parts: the protein, the bulk solvent and a third contribution representing the scattering from the excluded volume that the protein occupies. The bulk scattering will be dominant only at very low angles (q~0) with little contribution at low to intermediate angles. This gives the following description of the scattering contributions

$$F(\boldsymbol{q}) = \sum_{\boldsymbol{j}} \left(f_j - \rho_s v_j\right) e^{\boldsymbol{iq \cdot r_j}} \qquad (8)$$

where v_j is the volume of atom j and the term $\left(f_j - \rho_s v_j\right)$ is the contrast amplitude between the bulk solvent and atom j in the protein. The bulk solvent scattering is usually experimentally subtracted to negate the contribution the bulk solvent causes.

In a monodisperse protein solution system without particle interference between the protein molecules, the intensity of all the scattered waves can be described by summing up the intensities over all individual molecules such that,

$$S(q) = \sum_{\boldsymbol{n=1}}^{\boldsymbol{N}} |F_n(q)|^{\boldsymbol{2}} \qquad (9)$$

Different orientations of the individual protein molecules gives rise to different scattering properties, so to account for the rotational average in space the scattering intensity can be described by using the Debye formula,

$$S(q) = N\langle|F(q)|^2\rangle = N\sum_i\sum_j (f_i - \rho_s v_i)\left(f_j - \rho_s v_j\right)\frac{\sin(qr_{ij})}{qr_{ij}} \qquad (10)$$

where N is the number of molecules in the system and r_{ij} denotes the distance between atoms *i* and *j*. As can be seen from the Debye formula, in order to calculate the expected scattering from a protein in solution the three-dimensional coordinates of the target protein must be known. An experiment where a fit between the experimental and theoretical scattering data is a necessity is thus limited to proteins with solved structures.

2.3.2 Pump-probe data collection

To study a protein as it performs it biological function the resting state is perturbed and subsequently measured as a function of time. For X-ray diffraction methods the time-resolution of the measurement is limited by the integration and read-out time of the detector but can be circumvented by using the pump-probe technique where a system is excited with an activating pulse (pump) and after a defined time-delay data is collected with a measuring pulse (probe). The pulse length of the probe then determines the time-resolution achievable. By careful experiment design with regards to the time-delays chosen the reaction process induced by the pump pulse can in principle be sampled over several orders of magnitudes.

The time-resolved wide-angle X-ray scattering (TR-WAXS) data presented in this thesis were collected using the pump-probe approach at the Coherent X–ray Imaging (CXI) beamline[173] of the Linac Coherent Light Source (LCLS). A titanium-sapphire (Ti:Sa) femtosecond laser was used to pump the photoactive $\mathrm{RC_{vir}}$ injected into the vacuum shroud, described in Figure 7.

Since each X-ray pulse delivered by the LCLS is only 40 fs, sub-picosecond timepoints could be sampled for the first time in a protein TR-WAXS experiment (**Paper III**). One major advantage with XFEL TR-WAXS compared to synchrotron experiments is that each pump-probe data collection measurement is performed on a fresh sample, eliminating the risk of collecting data on photo bleached or X-ray damaged protein samples.

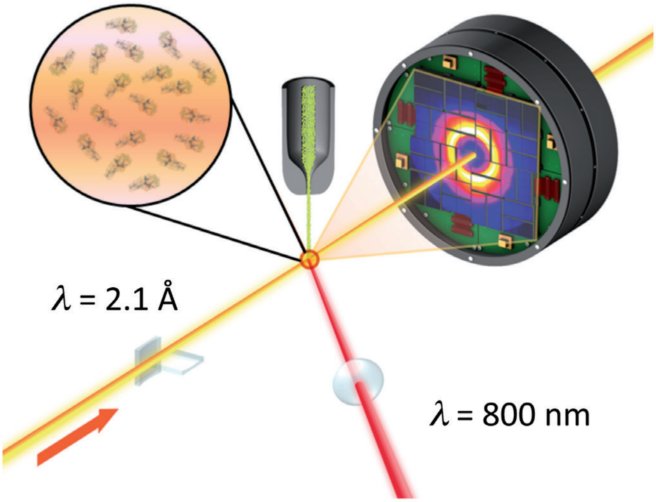

Figure 7. *Schematic representation of pump-probe TR-WAXS data collection at the Coherent X-ray Imaging (CXI) beamline of the LCLS. Solubilized* RC_{vir} *is injected into the sample environment using a liquid microjet. Protein is pumped with a 800 nm femtosecond laser pulse and data is collected on a CSPAD detector.*

Data were collected on a Cornell-SLAC Pixel Array Detector (CSPAD)[174] read out at 120 Hz (same as the XFEL) with the optical laser operating at a 60 Hz repetition rate. To reduce background scattering in a TR-WAXS experiment a helium cone is traditionally placed between the sample position and detector whereas background minimization at the CXI is done by having the sample environment in an enclosed vacuum chamber.

Each recording from the detector was angularly integrated to yield one-dimensional scattering curves, with the scattering intensity (S) plotted as a function of q (Figure 8). Interleaving dark (laser off) and light (laser on) images enabled difference scattering analysis while minimizing the risk of hardware drifts over time.

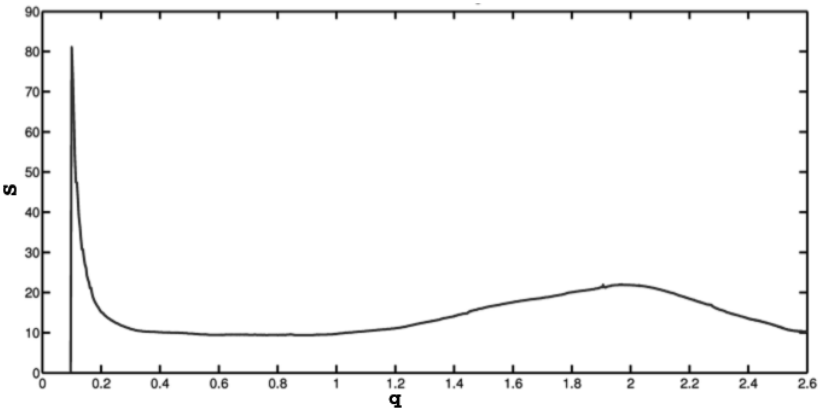

Figure 8. *Angularly integrated diffuse patterns yields a one-dimensional scattering curve.*

2.3.3 Difference scattering

Time-resolved solution scattering has proven to be a powerful method for studying reaction dynamics in the solution phase despite the fact that the relative sensitivity of the scattering to the protein is low since the solvent molecules outnumber the protein molecules at the concentrations typically used. The structural sensitivity is increased by subtracting the raw scattering curve of a negative or dark time-point from that of a positive time-delay, creating a difference scattering curve[175].

Initially the absolute scattering curves are normalized in a region (q_0) where scattering changes between the light and dark are unlikely to occur

$$S_{norm}(q, \Delta t) = \frac{S(q,\Delta t)}{\sum_{q_0} S(q,\Delta t)} \qquad (11)$$

Difference scattering can then be calculated by either subtracting a negative time point where the laser arrives after the X-rays (used for synchrotron pump-probe experiments) or by subtracting the dark scattering (no laser) from the light scattering curve (positive time-delays),

$$\Delta S(q, \Delta t) = S_{norm}(q, \Delta t) - S_{norm}(q, \Delta t < 0) \qquad (12)$$

It is common to use two or more flanking dark images as reference when calculating the difference scattering, which minimizes errors due to measurement variations. Despite the inherently low signal to noise ratio in WAXS there are some advantages with the difference curve method. Systematic errors in the experimental setup and hardware are reduced by he interleaved dark/light approach, background scattering cancels out when calculating difference curves since they are identical in both dark and light curves, and the relative scattering contribution from the perturbed protein is enhanced compared to those in the resting-state. Cross-terms between protein-micelle, micelle-solvent and solute-solvent are not cancelled out though and must be corrected for when interpreting the data.

2.3.4 Solvent thermal response

Using a laser to induce structural changes in a protein will also cause excess energy to be deposited to the surrounding bulk solvent. The average distance between water molecules is ~3 Å giving a distinct ~2 $Å^{-1}$ peak in the absolute scattering curves (Figure 8). As heat is transferred to the surrounding water the solvent expands causing the water peak around 2 $Å^{-1}$ to change, which can be seen in Figure 9b. The thermal-induced changes in the solvent primarily contribute in the 1.2-2.5 $Å^{-1}$ region, but scattering differences throughout the entire *q*-region can be seen, albeit not as strong as around the water peak.

Practically, this has the effect that each difference curve is a combination of the protein structural signal and the thermal response signal, hence separating them is crucial for correct interpretation of the data.

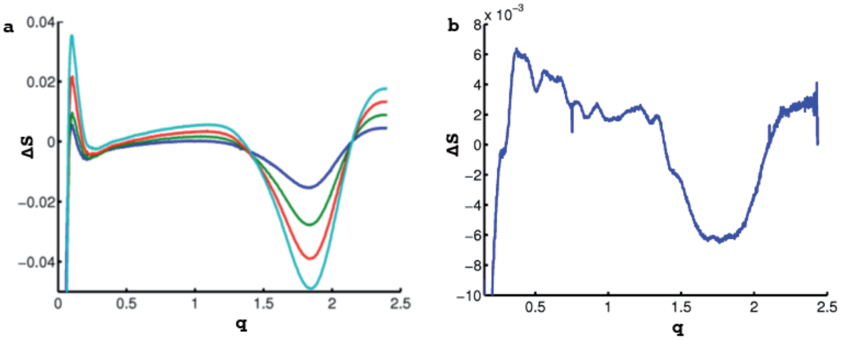

Figure 9. *Heating and light-induced differences in* RC_{vir}. *(a) Pure heating data collected on* RC_{vir} *using a cryojet system at the BioCARS beamline of the APS. (b) Difference scattering data collected on* RC_{vir} *on the Coherent X-ray Imaging beamline of the LCLS using a 800 nm Ti:Sa femtosecond laser to pump the protein. As can be seen no signs of structural re-arrangements is visible in the q = 0.2 – 1.0 region for the pure heating (a) whereas several oscillation is observed in the WAXS difference scattering (b).*

Usually this is done by collecting a pure thermal signal with the use of an infrared laser, which only deposits heat and does not give rise to structural changes in the protein, or by applying a temperature shift at the sample position and recording X-ray scattering data at different temperatures.

To reliably extract the difference signal, one or more outlier rejection criteria are usually needed. For data with fluctuating scattering intensities, such as the data collected at an XFEL where there is a shot-to-shot variation in the X-ray intensity, a rejection based on the standard deviation compared to the average scattering can be adopted. Individual curves exceeding the criteria in Equation 13 at any q within a given q-range are rejected

$$|S(q,\Delta t) - \langle S(q,\Delta t)\rangle| > \sigma S_{stdev}(q,\Delta t) \qquad (13)$$

where σ denotes how many allowed standard deviations are tolerated.

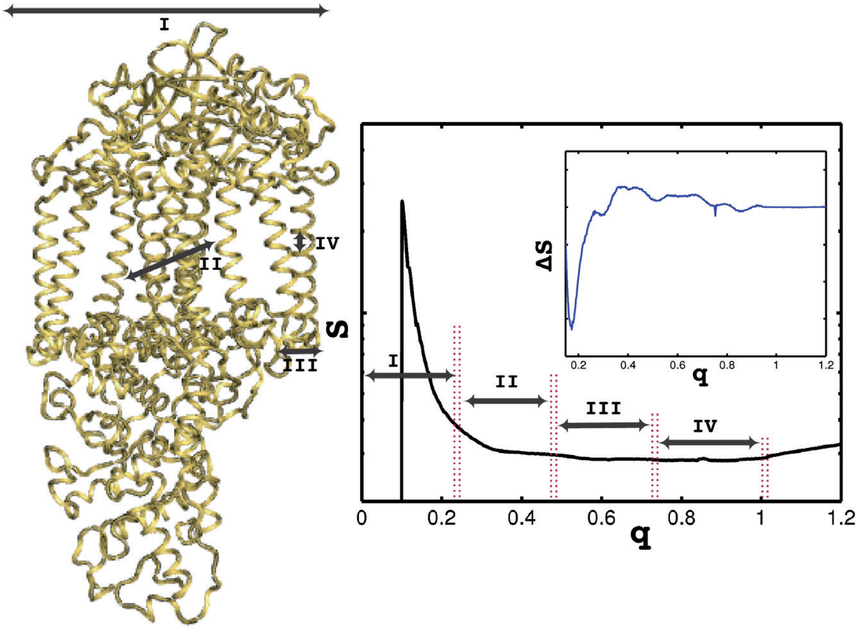

Figure 10. Illustration of the distances that can be seen in WAXS and the corresponding q-region on the scattering curve where the light-induced changes are observed. Inset shows a WAXS difference curve of RC$_{vir}$.

2.3.5 Structural interpretation of time-resolved WAXS data

Different *q*-regions in the scattering curve obtained from TR-WAXS contain information about different length distributions in the protein (Figure 10). In the low-q region, the overall size and shape of the sample molecule as well as properties such as the maximum distance and radius of gyration of the particle can be obtained[83]. At intermediate to wide angles distances between secondary structural elements and individual domains can be probed and at even wider angles distances within individual side chains add to the scattered intensity. Beyond $q > 1.2$ Å^{-1} solvent scattering dominates and the contributed scattering intensity at these angles from the protein is very weak, limiting the resolution range. In the difference scattering data oscillations indicate structural changes in the studied system and are related to the formation and depletion of distances.

The main challenge in TR-WAXS is interpreting the one-dimensional difference scattering curves as a three-dimensional conformational change. There are several approaches to solving this challenge, but fitting theoretical scattering to experimental data is a common denominator. Producing the excited structures for the theoretical difference calculation can be done by rigid-body movement of secondary structure elements[96] or by producing an ensemble of PDB structures sampling the conformational space around a known intermediate by molecular dynamics simulations[104]. However, both these methods require a solved three-dimensional structure of an intermediate of the target protein. Another way of producing perturbed model structures is to induce structural changes in the resting state protein of a molecular dynamics (MD) trajectory by modifying the chromophore or simulating an absorbed laser pulse by deploying heat into a cofactor (**Paper III**).

Irrespective of how the theoretical difference scattering curves are generated, a scoring criteria to the experimental data needs to be applied, seeking to minimize the differences and yield a three-dimensional solution to the conformational change observed. This can be done by an error-weighted χ^2-scoring function[101]

$$\chi^2 = \sum_q \frac{\left[\Delta S_{theory}(q) - \Delta S_{exp}(q)\right]^2}{\sigma_q^2(N-M-1)} \qquad (14)$$

where σ_q is the standard deviation in each *q*-point of the data, *N* is the number of data points in the *q*-range and *M* the number of parameters used for fitting.

For 3D-structural interpretation of the data presented in **Paper III** an internal distance matrix approach was used to identify conserved movements in RC_{vir} in the best fits ensemble, further explained in section 3.4.6.

2.3.6 Molecular dynamics

Molecular dynamics (MD) simulations have been employed for studies of proteins for nearly thirty years[176] and they are a very useful tool for investigating the dynamics of protein structural rearrangements. An MD simulation attempts to mimic the experimental conditions by constructing an atomic *in silico* system comprising of protein, solvent, ions and detergents (for membrane proteins). From this starting point, atom movements can be simulated in small time-steps (fs) based on Newton's law of motions. With the increased computational power in modern day simulations can now be followed for several nanoseconds[177]. From the simulation, an MD trajectory is produced containing the atom coordinates as a function of time, and by averaging over a trajectory that has reached equilibrium macroscopic properties can be extracted. For the work presented in **Paper III**, MD was used to simulate an RC_{vir}-LDAO micelle and by depositing heat into the RC_{vir} cofactors an excited trajectory was created. From this trajectory, PDB snapshots were extracted and used as input for theoretical difference scattering, further explained in section 3.4.4.

3. XFEL studies on Reaction Centre

3.1 Microcrystallization for SFX experiments

Unlike conventional crystallography where one single, well-diffracting crystal is used for collecting a dataset, SFX requires a large number of smaller crystals that each gives rise to partial reflections. Since the orientation of each crystal is random, sufficient amounts of crystal containing solution must be injected into the XFEL beam in order to get a complete dataset. One way of approaching the need for a large quantity of small diffracting microcrystals is to mechanically destroy larger crystals grown in traditional ways. One disadvantage with mechanical force as a way to produces microcrystals is that it is hard to control the size distribution and you run the risk of destroying the crystals altogether. The method used for the SFX experiments presented here is the batch crystallization in LSP where conditions were tweaked to produce large number of smaller crystals in large volume batches to ensure that enough crystal material would be produced.

3.1.1 Scaling up production and purification

An SFX experiment at the LCLS using a liquid microjet to inject the samples requires approximately 7 ml of sample per 12-hour shift. With a protein concentration of 50 mg/ml this corresponds to a need for 350 mg of pure membrane protein per day, hence the production and purification of RC_{vir} had to be scaled accordingly. This was done by running several *Bl. viridis* cultivations simultaneously and exchanging both the anionic exchange and the size exclusion columns. The former was switched for a larger diameter column, enabling a fivefold increase in loading volume, and the size exclusion column increased with a factor of three. Using this scaled up purification protocol a yield of 100 mg RC_{vir} pure enough for crystallization was extracted from 50 g of *Bl. viridis* cells.

3.1.2 Batch microcrystallization of RC_{vir} in lipidic sponge phase

LSP crystallization conditions resulting in well diffracting macrocrystals had previously been established in the lab[77, 152] and were used as a starting point for growing microcrystals. Initially vapour diffusion sitting drop experiments were conducted where the Jeffamine M600 content and the LSP to protein ratio were varied. Although small crystals were identified in these trials they were not sufficiently small for SFX. This was followed by growing the crystals in batch using glass vials with a volume of 60 µl per setup with varying ratios of protein, additives and LSP. Small crystals similar to those found in the first round of crystallization trials were found. A subsequent step was to omit all non-protein containing solution filtration steps to provide more nucleation sites and increasing the batch size to 250 µl. The final conditions were as follows: 100 µl of RC_{vir} at 20-30 mg/ml, 100 µl LSP (12 % monoolein, 17.5 %

Jeffamine-M600, 1.0 M Hepes at pH 8.0, 0.7 M $(NH_4)_2SO_4$, 2.5 % 1,2,3-heptanetriol) and 50 µl of 1.0-1.2 M tri-sodium citrate as additive.

Several methods can be adopted to confirm the existence of microcrystals. Light spectroscopy with a polarizer or ultraviolet source can be used to detect crystals in the medium to high micron region, though for crystals in the lower to sub micron region other methods must be applied for visualizing the crystals. Small-angle X-ray scattering (SAXS), Scanning Electron Microscopy (SEM) and Second Order Non-linear Imaging of Chiral Crystals (SONICC)[178] have all been shown to work and were applied to the RC_{vir} microcrystal sample where both SAXS and SEM were successful (Figure 11). Wavelengths used for identifying the crystals in SONICC conflicted with RC_{vir} absorption in that region, which may be the cause for not detecting microcrystals with that method.

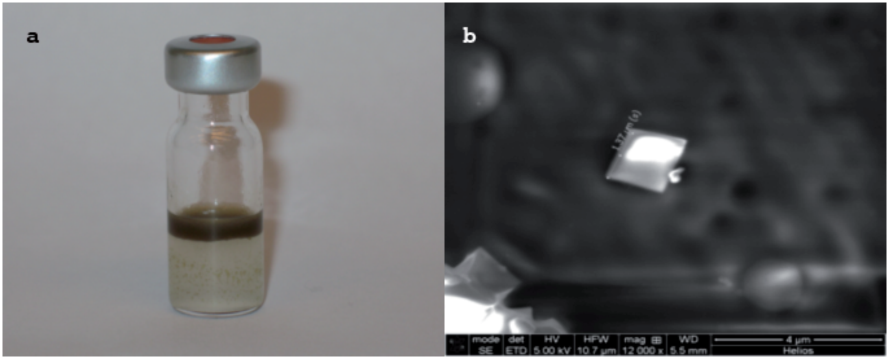

Figure 11. *(a) LSP batch crystallization setup of RC_{vir} in a glass vial and (b) Scanning Electron Microscopy image of a batch-setup RC_{vir} crystal with a length of 1.37 µm.*

3.1.3 Sample delivery methods

Over the last couple of years a lot of research has been focused on developing tools for injecting SFX or WAXS samples into the XFEL beam. Some of the fundamental requirements of a delivery system are that they can be used at room temperature, produce thin jets to reduce background scattering, do not harm the protein, and that each XFEL shot hits a fresh sample[179]. One method is the Rayleigh jet that was shown to produce a water jet in vacuum in 1988[180]. A major disadvantage with the Rayleigh jet is that a very high flow rate is required. Moreover if the orifice diameter is reduced to the same dimension as the XFEL beam (0.2-10 µm) to minimize background scattering, the jet clogs[181, 182]. Another technique which has reported flow rates of 0.17-3.1 µl/min is the electrospinning delivery method[183] which has been successfully used in several SFX experiments[183-185]. This method requires that the protein crystal sample is stable with the addition of antifreeze solutions and is not affected by the electrostatic charging. The lowest achieved flow rate to date has been produced by the LCP injector[186] with a

300 pl/min sample consumption[179]. The LCP injector has been used at the LCLS for solving two human G-protein coupled receptor structures[186, 187] and recent developments of the LCP injector enables a flow speed of 2-3 mm/s which is sufficient for collecting data at the full repetition rate of 120 Hz currently provided at the LCLS[179]. An obvious limitation of the LCP injector is that it requires crystals to be grown in LCP, so for the LSP based SFX and WAXS experiments presented here the gas dynamic virtual nozzle (GDVN)[188] injection system was used (Figure 12). GDVN uses a co-flowing helium gas stream to focus a liquid jet to a factor of ten smaller diameter than the orifice it originates from. Since a capillary of 50 µm can then be used, the risk of clogging is greatly reduced compared to the standard Rayleigh jet, while still enabling a ~4 µm jet providing low background and with a sample flow rate of around 10 µl/min[179]. This is the most commonly used method for delivering samples to the XFEL beam and has been successfully used for a number of biological experiments[189-193]. A major flaw with the GDVN is that the flow rate of the microjet needs to be high enough to produce a stable jet, which renders most of the sample injected without ever interacting with the XFEL beam and a sample volume of at least 10 ml is required for a full dataset to be collected. To address this protein waste problem, pulsed injection systems synchronized with the XFEL repetition rate have been proposed as a solution, but are still in the development stage and have not been successfully used in an SFX experiment to date[179].

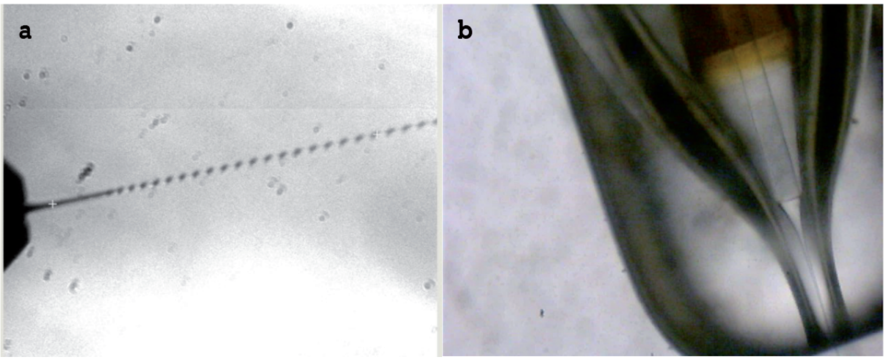

Figure 12. *The gas-driven virtual nozzle (GDVN) microjet system. (a) A microjet formed by a GDVN. Data is collected before the Rayleigh break-up region where the uniform jet is turned to droplets and (b) a close-up of the GDVN microjet tip.*

To investigate the compatibility of RC_{vir} grown in LSP and the GDVN microjet, both Jeffamine and PEG based LSPs were tested. The PEG-based LSPs have a higher monoolein content and required a 100 µm nozzle for a stable microjet to be produced whereas the Jeffamine LSPs flowed well in a 50 µm inner diameter nozzle.

3.2 Proof-of-principle low resolution structure (Paper I)

The initial proof-of-principle experiment that RC_{vir} grown in batch LSP could produce diffraction data was performed at the Atomic, Molecular and Optical Science (AMO) beamline[194] at the LCLS in June, 2010. The highest resolution achievable at that time was 7.4 Å due to limits in experimental geometry and X-ray wavelength. Though the crystal hit-rate was low and the resolution was limited it was still possible to obtain a molecular replacement solution, produce an electron density map and discover a new space group and crystal packing for RC_{vir}.

3.2.1 Data collection

Using the GDVN microjet, hydrated RC_{vir} crystals were injected with a flow rate of 10 μl/min producing a 4 μm thick microjet aligned vertically relative to the XFEL beam. Diffraction data were collected at the AMO beamline of the LCLS using the Center for Free-Electron Laser Science Advanced Study Group multipurpose chamber (CAMP) pn-junction charge-coupled devices (pnCCDs)[195] two-panel detector (Figure 13) read out with a repetition rate of 60 Hz, the same as the XFEL rate at that time. X-ray photon energy was 2 keV, corresponding to a wavelength of 6.17 Å, with 10^{12} photons per pulse, which had a duration of 70 fs and spot size of 10 μm^2.

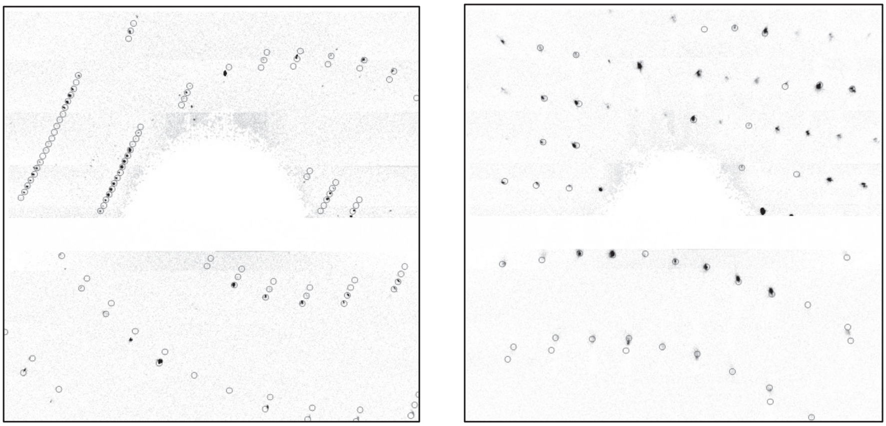

***Figure 13**. Two diffraction patterns collected on the pnCCD detector panels at the AMO beamline of the LCLS with indexing rings from CrystFEL.*

3.2.2 Data processing

During the ~100 minutes of data collection a total of 365,035 images were recorded out of which 1,542 were identified as hits using the Cheetah software[169]. Manual inspection of the recorded data, where diffraction patterns were arranged based on the number of Bragg peaks using in-house

MatLab code, yielded a list with the highest quality data that were then analysed with the CrystFEL software suite[168] that uses DirAx[170] and MOSFLM[171] for indexing. Initially cell parameters from a previously solved RC_{vir} crystal structure grown in LSP[152] were used as input, but the peaks in the SFX diffraction patterns did not agree with the predicted peak locations. The CrystFEL software suite allows for *ab initio* indexing without previously known cell parameters and space group. Applying this to the collected dataset gave, after several iterations, axes a=57.6 Å, b=84.6 Å, c=375.8 Å and α=β=γ=90° in the novel space group $P2_12_12_1$. In total, 265 diffraction images were indexed and no other crystal forms were observed in the cell parameter histogram (Figure 14).

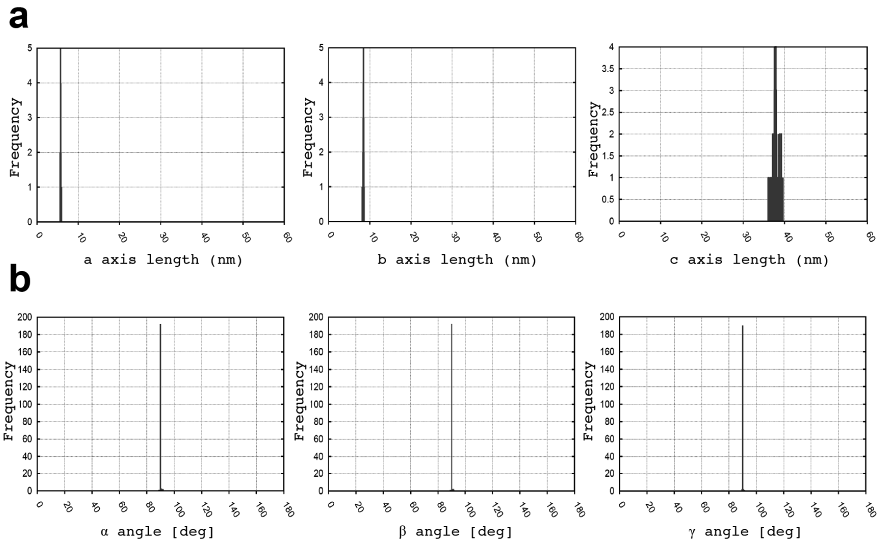

Figure 14. *Histograms of cell axes (a) and cell angles (b) from 265 indexed RC_{vir} diffraction patterns.*

Diffraction data were processed using a Monte Carlo method[172] since each recorded diffraction image only contains partial reflections from randomly oriented microcrystals. NZ- and L-test plots were produced with the xtriage program of the PHENIX[166] package and showed distributions as expected for a non-twinned space group such as the $P2_12_12_1$ space group.

3.2.3 Molecular replacement and data refinement

Phases were solved with Phaser 2.3.0[196] by molecular replacement using PDB entry 2WJN[152] as a search model. For a definite solution a translation function Z score (TFZ) of 8 and log likelihood gain (LLG) of 40 are required in Phaser and the best solution had TFZ = 8.5 and LLG = 81 compared to the second best solution with 4.6 and 39 scores for TFZ and LLG respectively. Refinement was performed using twenty cycles of rigid body and restrained refinement in REFMAC5[165] producing R_{factor} and R_{free} values of 0.35 and 0.38 and an overall figure of merit of 0.62. Wilson scaling was used during refinement and the solvent mask was calculated using default parameters. Diffraction data were cut at 8.2 Å (Figure 15) where the multiplicity >4 and completeness >85 %. The crystallographic data are summarized in Table 1.

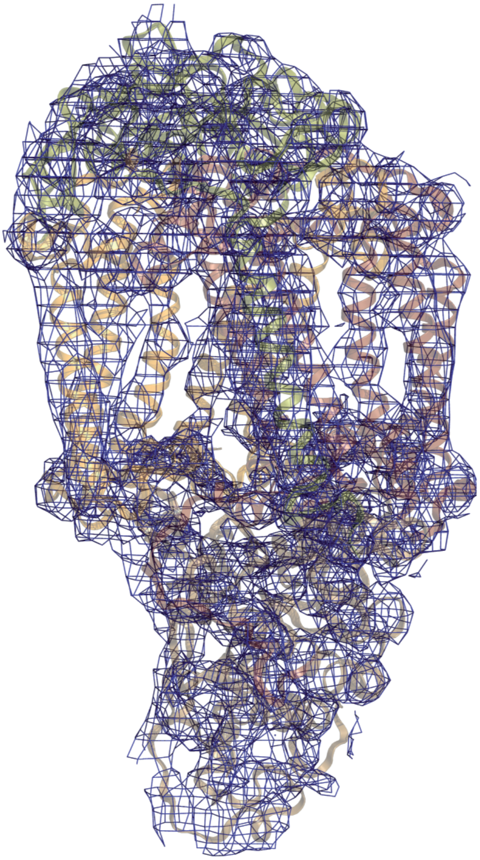

***Figure 15**. $2F_{obs}$-F_{calc} electron density of the RC_{vir} structure at 8.2 Å recovered from 265 processed SFX images. Contoured at 1.0 σ.*

Crystallographic data summary	
Total number of images	365,035
Images with diffraction	1,542
Number of indexed images	265
X-ray wavelength	6.17 Å
Data processing resolution limits	46.1 – 7.4 Å
Number of used reflections	2,247
Space group	$P2_12_12_1$
Unit cell parameters	a=57.6 Å, b=84.6 Å, c=375.8 Å
	$\alpha=\beta=\gamma=90°$
Completeness (%)	85
Multiplicity	8.0
I/σ	1.9
Refinement resolution limits	46.1 – 8.2 Å
R-factor (%)	35
R_{free} (%)	38
Overall figure of merit (%)	62.0

Table 1. *Summary of crystallographic data and refinement statistics for the SFX data collected in 2010.*

3.2.4 Control maps

At 8.2 Å the detailed information content that can be observed is quite low, hence a number of control maps were calculated on the low-resolution SFX data as well as on monochromatic[152] and a Laue[77] data sets for comparison. Additionally, to verify the molecular replacement solution, control calculations were performed using two control datasets: one with randomly shuffled *h*, *k*, *l* indices of the processed SFX data and one with uniform intensities. Both these control datasets failed to produce a molecular replacement solution. Using these control datasets together with the phases recovered from the experimental data resulted in electron density maps that did not show any α-helical structure. The conclusion was thus drawn that uniform or random data could not produce a structurally correct electron density map and that the surprisingly convincing electron density map obtained from the experimental data was not purely due to model bias.

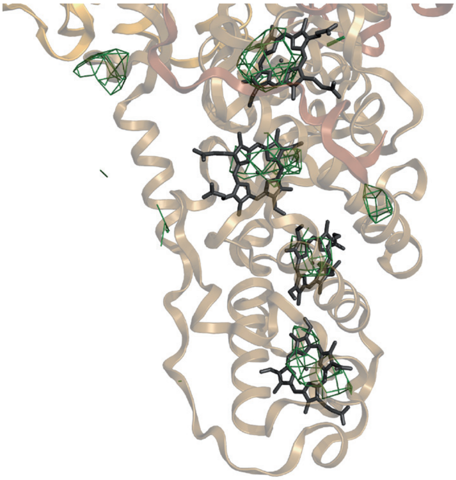

Figure 16. Omit electron density map calculated with all four heme groups removed. Contoured at 2σ.

For the omit map calculation all four heme cofactors of the cytochrome *c* subunit were removed and a positive F_{obs}-F_{calc} electron density correlated with each of the heme groups were observed in the resulting map (Figure 16). Finally, a composite omit map was calculated on the experimental data as well as on the monochromatic and Laue data both cut at 8.2 Å using CNS[167, 197]. All three datasets produced electron density map of similar quality (Figure 17).

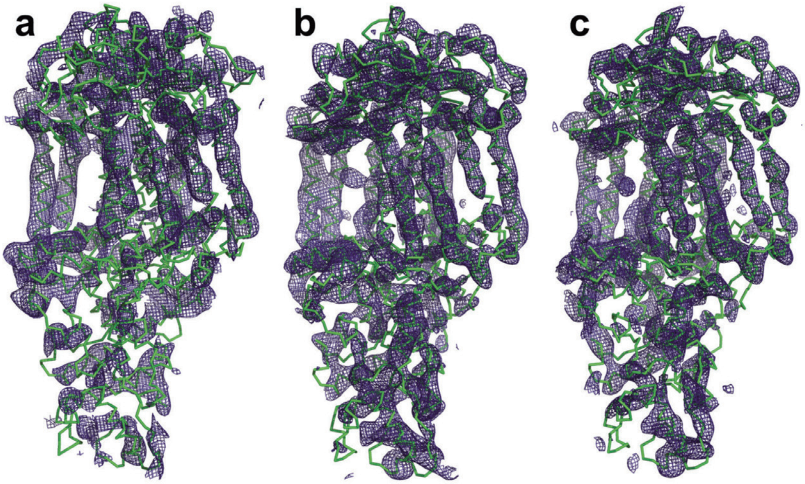

***Figure 17**. Composite omit electron density maps calculated with CNS 1.3 where 5 % of the model was omitted in each step. (a) SFX data, (b) data extracted from PDB entry 2WJN and (c) data extracted from PDB entry 2X5U. All datasets were cut at 8.2 Å and contoured at 1.5 σ.*

3.2.5 Summary Paper I

Using 265 diffraction images collected at the LCLS on LSP batch grown microcrystals of RC_{vir}, an 8.2 Å resolution electron density map could be produced and the first ever novel space group from a structure solved with an XFEL was presented. This novel space group indicates that the way the crystals grow in large batches differs to the crystallization route taken in vapour diffusion setups. Several different LSPs based on both PEGs and Jeffamine M600 were shown to produce a stable jet using the GDVN injection system making LSP a viable option for SFX protein targets.

Data were collected in 2010 at the AMO beamline at the LCLS providing X-rays with a pulse energy of 2 keV and an experimental setup limiting resolution to 7.4 Å. Using 100 minutes worth of data collection at 60 Hz an RC_{vir} structure processed to 8.2 Å was obtained with R_{factor} and R_{free} values of 0.35 and 0.38 respectively. The $2F_{obs}$-F_{calc} and composite omit electron density map clearly shows transmembrane α-helical structures despite the low resolution. Furthermore, positive F_{obs}-F_{calc} electron densities were observed, correlated to the heme groups of the cytochrome *c* subunit in the omit map calculated on a structural model with these cofactors removed, indicating acceptable levels of model bias.

3.3 High resolution structure of RC_{vir} (Paper II)

In December 2010 the Coherent X-ray Imaging (CXI) beamline[173] was commissioned, enabling crystallographic data to be collected at 120 Hz with X-ray pulse energies of up to 11 keV focused on a 1.3 x 1.3 μm^2 spot. This enabled high-resolution SFX experiments to be performed. An experiment with RC_{vir} as one of the high-resolution protein targets was conducted in February 2011.

3.3.1 Data collection

For high-resolution SFX data collection, the GDVN microjet delivery system already proven to work with LSP batch grown microcrystals of RC_{vir} was used for sample injection at the CXI beamline. The XFEL beam delivered X-rays with 9 keV pulse energy and with a focal spot of 10 μm^2. Data were recorded on a Cornell-SLAC Pixel Array Detector (CSPAD)[174] read out at 120 Hz. The CSPAD detector consists of 64 independent panels, 192 x 185 pixels each with a pixel size of 110 x 110 μm, which forms a 1516 x 1516 pixel array with small gaps between each panel and a larger gap in the detector centre through which the XFEL beam passes. The detector distance during data collection was 142 mm, corresponding to 2.62 Å resolution at the detector edges. RC_{vir} crystals diffracted up to 2.8 Å (Figure 18) and during the five-hour data collection 2,744,614 images were recorded.

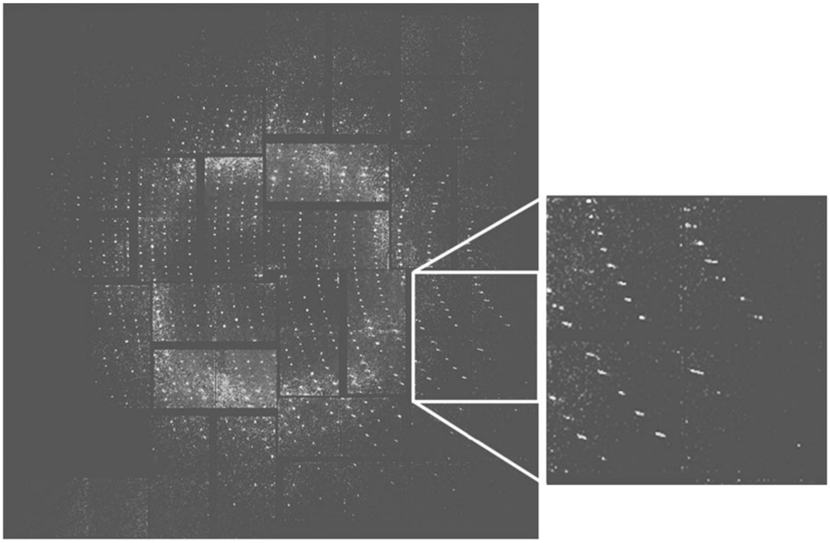

Figure 18. *Diffraction pattern of RC_{vir} collected on a CSPAD detector at CXI in 2011 showing diffraction spots up to 2.8 Å.*

3.3.2 Data processing

Images with more than ten Bragg peaks were extracted from the data stream using the Cheetah[169] program in which background and coherent scattering from the microjet were also removed. 88,924 images were classified as hits by the Cheetah software and subsequently ordered into lists based on the diffraction quality using an in-house MatLab script. The list with the best diffraction images were then iteratively processed with CrystFEL[168] where MOSFLM and DirAx were used by the CrystFEL indexing routine. To account for the c-axis variability, high tolerances were used. Scaling and merging of the data were also done with CrystFEL, where the merging was performed according to the Monte Carlo method[172]. Data collection statistics are summarized in Table 2 and were calculated by splitting the experimental data into two halves and comparing them relative to each other with CrystFEL.

Data collection and refinement statistics	
Total number of recorded images	2,744,614
Number of images > 10 spots	88,924
Confirmed diffraction patterns	5,767
Number of indexed images	1,175
X-ray energy (keV)	9.34
Space group	$P2_12_12_1$
Unit cell parameters	a=57.9 Å, b=84.8 Å, c=384.3 Å
	$\alpha=\beta=\gamma=90°$
Completeness (%)	99.1 (93.4)
Multiplicity	27.0 (27.6)
Overall R_{split} on I (%)	36.5 (52.7)
Mean $I/\sigma(I)$	3.50 (2.0)
$CC_{1/2}$	0.54 (0.32)
Refinement resolution limits	49.6 – 3.50 (3.66 – 3.50)
Number of unique reflections	24,721
R_{work} / R_{free} (%)	29.4 / 32.7
Overall figure of merit (%)	78
Number of atoms	10,039
Protein residues	1,191
Root mean square (bonds, Å)	0.02
Root mean square (angles, °)	1.517
Wilson B factor (Å^2)	79.5

***Table 2**. Summary of crystallographic data and refinement statistics on SFX data collected in 2011. Values in parenthesis refer to the highest resolution shell (3.65-3.5 Å)*

3.3.3 Molecular replacement and data refinement

Phases were obtained by molecular replacement using the PhaserMR 2.5.2[196] software using PDB entry 2WJN[152] as a search model. A translation function Z score (TFZ) = 32.8 and log likelihood gain (LLG) = 2,879 were obtained, both clearly above the threshold for a definite solution. REFMAC5[165] was used for refining the structure with rigid body and jelly body refinement together with manual rebuilding in COOT[164]. The R_{free} value was improved by ~1 % using TLS[198] in combination with the jelly body refinement. The final R and R_{free} values converged to 0.29 and 0.33 respectively with an overall figure of merit of 0.78. PROCHECK[199] and MOLPROBITY[200] were used for validating the structure. Table 2 shows a summary of the refinement statistics.

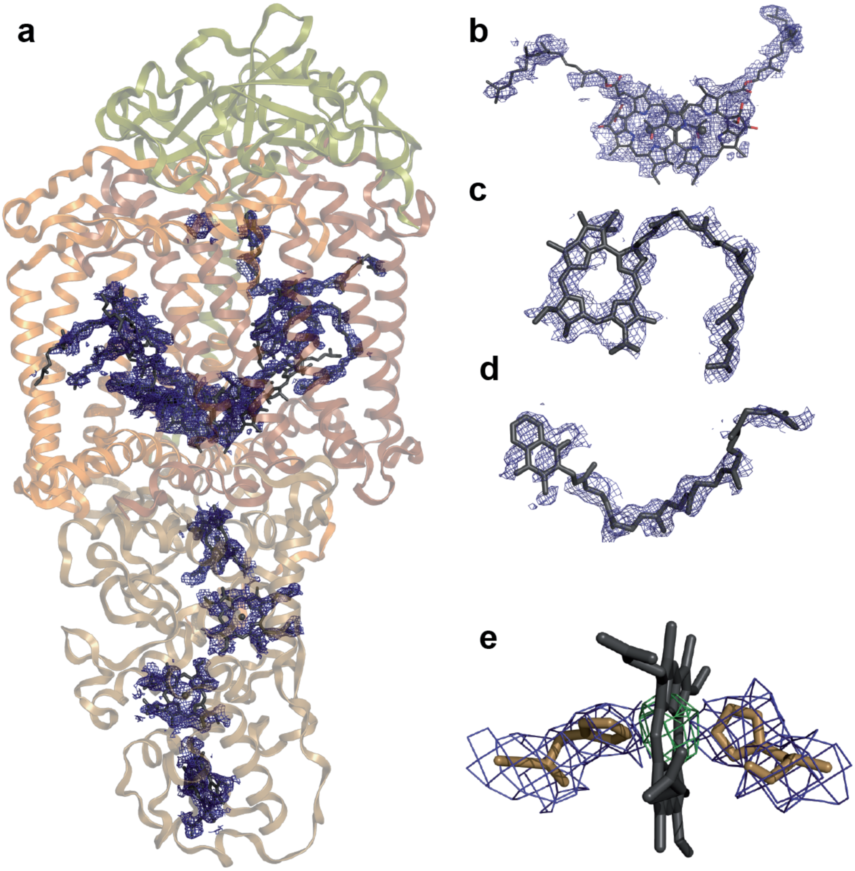

Figure 19. *$2mF_{obs}$-DF_{calc} electron density maps of (a) RC_{vir} cofactors with close-up on (b) the special pair, (c) bacteriopheophytin and (d) the menaquinone in the Q_A pocket. (e) F_{obs}-F_{calc} electron density constructed with the iron atom of heme 4 removed. Panels (a)-(d) are contoured at 1.0 σ and panel (e) at 3.0 σ.*

3.3.4 The 3.5 Å resolution structure

Using 1,175 indexed diffraction images, data were refined to 3.5 Å resolution where I/σ was 2.0 and with a completeness of 93.4 %. The resulting electron density map clearly shows the bacteriochlorophyll tail of the special pair as well as visible density for all the lipids found in the structure (Figure 19). A monoolein molecule was observed in the Q_B binding pocket instead of the native ubiquinone-9. Using an internal matrix approach[201], the 3.5 Å RC_{vir} structure solved with SFX was compared to all other RC_{vir} structures deposited in the protein data bank. Since internal distances are calculated between the C_α atoms of the backbone, the structures do not have to be super-positioned. The structure from LSP batch grown microcrystals were shown to cluster with the room temperature Laue structure[77], having an average internal distance matrix change on the C_α atoms of 0.27 Å, similar to the value given when comparing the room temperature Laue structure to all cryo-cooled structures of RC_{vir}. The major change is revealed to be in the C-subunit with relatively small changes in the transmembrane domains (less than 0.5 Å). A slight twist of the C-subunit is observed when overlaying the room temperature Laue with the SFX structure with movements of up to 1.9 Å at the C-subunit edges. Since subunit C is involved in the crystal packing contacts the conclusion was drawn that these were not systematic differences between the SFX and Laue data collection but rather an effect of packing differences.

All iron and magnesium atoms were omitted from the model for calculation of the control maps, resulting in a positive F_{obs}-F_{calc} electron density associated with each of the iron-containing heme cofactors and for the special pair coupled magnesium (omit map obtained for heme 4 displayed in Figure 19e). There are no signs of radiation damage in the SFX structure seen in the F_{obs}-F_{calc} map. With an X-ray fluence of 4×10^{11} 9 keV focused on a 10 μm^2 spot, a dose corresponding to 33 MGy is deposited to the sample. This is a similar exposure to that of cryo-cooled crystals and two orders of magnitude higher than useful for standard room temperature crystallography[38], demonstrating that SFX is an advantageous method for radiation damage-free structure solving.

3.3.5 Summary Paper II

Building on the results from **Paper I** where lipidic sponge phase batch crystallization of RC_{vir} was shown to be compatible with the GDVN injection system the LSP batch method was optimized and shown to yield crystals diffracting up to 2.8 Å when conducting a high-resolution SFX experiment at the CXI beamline of the LCLS. From five hours of beamtime more than 2.7 million diffraction images were recorded, with 88,924 crystal hits and 1,175 diffraction patterns used for processing the dataset to 3.5 Å resolution, making RC_{vir} the first high-resolution membrane protein structure to be solved at an X-ray free-electron laser.

An internal distance matrix calculation between all deposited RC_{vir} structures revealed that the SFX and the Laue diffraction ground state structure solved to 2.95 Å were the most similar. Upon overlaying the SFX and Laue structure a shift in the C-subunit with movements up to 1.9 Å of the C_α atoms was observed. This is proposed to be a result of crystal contacts in the SFX structure rather than differences in data collection approaches.

Signs of radiation damage were also investigated, with the iron and magnesium atoms omitted from the model for calculation of a F_{obs}-F_{calc} map. Well-ordered positive electron densities can be seen without any indication of radiation damage around each of the metal atoms of the cofactors.

3.4 Time-resolved WAXS study of RC_{vir}: visualizing a protein quake (Paper III)

In February 2012 data were collected on solubilized samples of RC_{vir} at the CXI beamline[173] of the LCLS. Using a Ti:Sa femtosecond laser to excite the protein molecules at 800 nm wavelength the first sub picosecond time-resolved WAXS experiment on protein structural dynamics could be conducted.

3.4.1 Data collection

RC_{vir} samples solubilized in lauryldimethylamine-*N*-oxide (LDAO) detergent at a concentration of 30 mg/ml were injected to the XFEL beam using the GDVN microjet system. A 50 µm nozzle was used to produce a stable jet with a diameter of 10 µm with a sample flow rate of approximately 10 µl/min. The fundamental 800 nm wavelength of a titanium-sapphire (Ti:Sa) femtosecond laser was used to pump RC_{vir} using a pulse power of 260 µJ delivered with a 500 fs pulse and a 250 µm circular spot size. The beam profile of the pump laser was Gaussian, with an average energy corresponding to a fluence of 5.3 mJ/mm^2 passing through the circular focus.

Data were recorded on a Cornell-SLAC pixel array detector (CSPAD)[174] composed of 64 individual tiles each consisting of 192 x 185 pixels with a pixel size of 110 x 110 µm^2. The detector and XFEL repetition rate were 120 Hz with the Ti:Sa laser operating at 60 Hz such that dark and light images were interleaved. The XFEL beam was tuned to 6 keV, corresponding to a wavelength of 2.07 Å, and was focused to a 10 µm^2 spot size intersecting the liquid microjet before the Rayleigh breakup region. The flow speed of the sample jet was 10 m/s, thus ensuring a fresh sample for each XFEL shot.

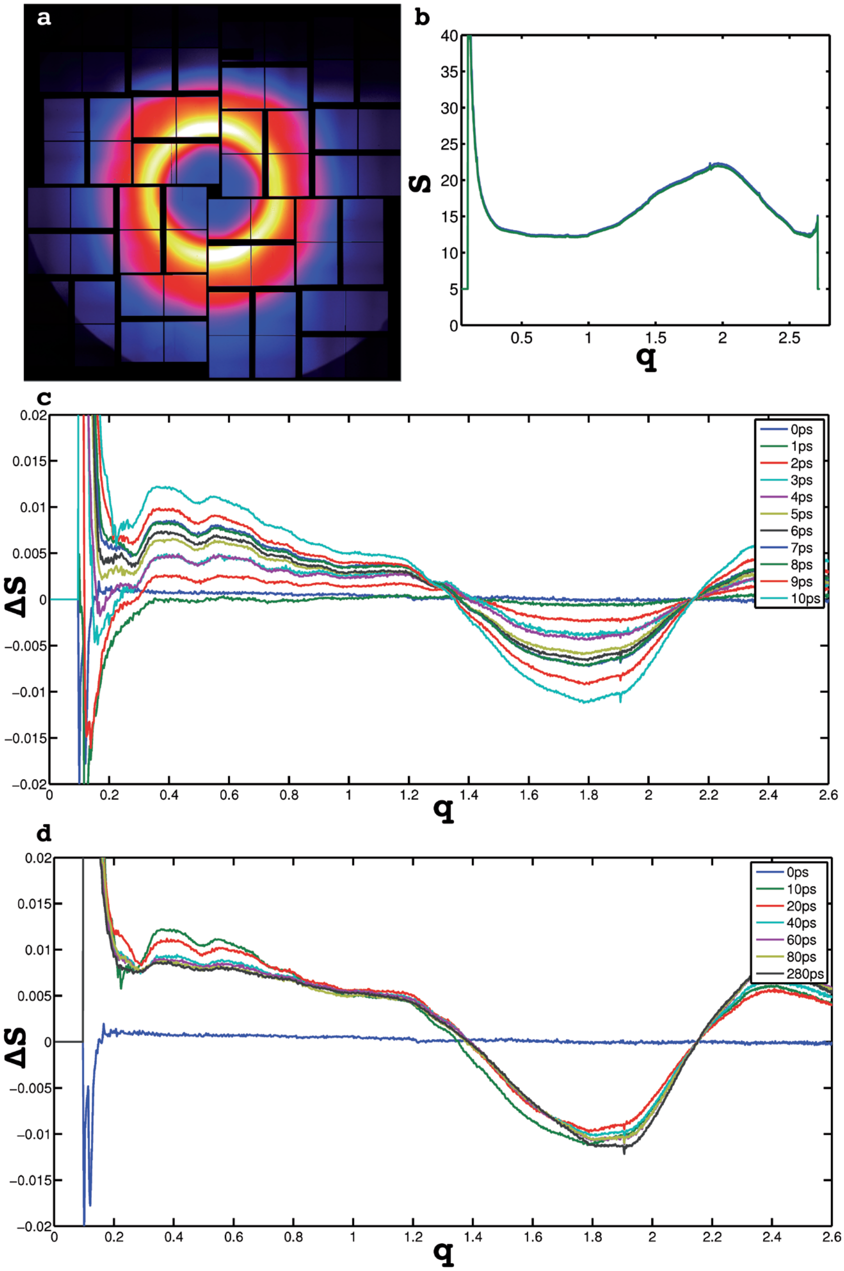

Figure 20. TR-WAXS data collected on RC$_{vir}$ at CXI in 2012. (a) Diffuse diffraction pattern recorded on a CSPAD detector after a single FEL X-ray exposure, (b) radially integrated absolute scattering with laser on (blue) and dark (green) virtually identical in the absolute scattering, (c) TR-WAXS difference scattering illustrating the rise of heating and protein signal within the first few picoseconds and (d) showing the decay of the protein signal with heating signal remaining at t=280 ps.

The setup at CXI when conducting the experiment limited the timing between laser and XFEL to 4 ns. To improve the simultaneous arrival of the Ti:Sa laser and XFEL pulses an online preliminary analysis script in MatLab was used. Using a high laser power we were able to follow the heating response from the RC_{vir} sample and determine time zero down to a sub picosecond level. Improvement in radial integration speed of the diffraction images and simplistic analysis of the data enabled feedback of the data collection within a few minutes after data were recorded.

3.4.2 Data processing

Each run was radially integrated using Cheetah[169], producing radial stack files with laser on and laser off radials separate. MatLab was used to produce difference curves by normalizing the raw data between q = 2.10 and 2.20 Å^{-1} and by three individual rejection routines, each with 3σ-criteria. These were an initial rejection on the raw absolute scattering data followed by rejection based on the normalized curves and a final rejection on the difference curves. Approximately 60 % of the data from each run passed these rejection routines, leaving ~9,000 difference measurements for merging at each time point. In total, 41 difference WAXS curves with unique time delays (Δt) between the arrival of the 800 nm optical laser and XFEL beam were recorded and used for subsequent analysis.

3.4.3 Extracting basis spectra using spectral decomposition

Spectral decomposition was used to process the 41 difference WAXS curves. Spectral decomposition is similar to the principal-component analysis approach, but in spectral decomposition the time-dependence of the states to be extracted is specified according to a kinetics scheme. From the data four components could be extracted: an ultrafast component (C1), a component due to protein conformational changes (C2) and two heating components (C3 and C4) where C3 describes the non-equilibrated and C4 the equilibrated heating. Simple first-order kinetics with exponential rise and decay were used to model components C1-C3. With Δt_{max} = 280 ps no decay for C4 was modelled and the exponential growth of the C4 component was coupled to the decay of C3 (i.e. C3→C4) and a time offset of 2 ps was used to fit the rise of C3.

The strategy for extracting the individual basis spectra was to focus on the q = 1.35 - 2.5 Å^{-1} domain to extract the C3 and C4 time-dependencies after which the C1 component was defined as the $\Delta S(q)$ averaged over the three first time-points. Iterative cycles of subtracting the C1, C3 and C4 components from the difference WAXS data were used to characterize the C2 basis spectra. Removal of the C2 component from the data was used to better characterize the C3 and C4 components. Final basis spectra yielded are shown in Figure 21c.

A linear sum of each of the final four components was used to reproduce the experimental data and validate the spectral decomposition kinetic model (Figure 21a).

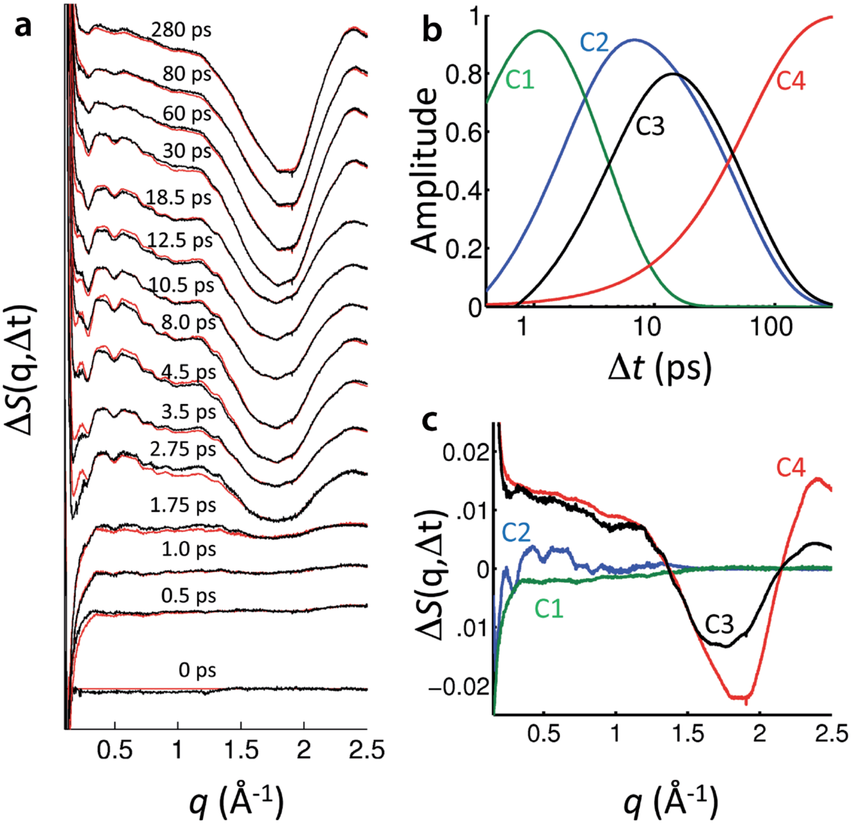

Figure 21. *(a) Time-resolved WAXS difference data (red) overlaid with linear sums of the four basis spectra shown in (c). (b) Time-dependent amplitudes of the four components C1-C4 used to extract the basis spectra. (c) Basis spectra extracted from the time-resolved WAXS difference data by spectral decomposition: An ultrafast component (C1), a protein component (C2), a non-equilibrated heating component (C3) and an equilibrated heating component (C4).*

3.4.4 Molecular dynamics simulations

Molecular dynamics trajectories with reaction centre coordinates taken from the 2.3 Å structure[202] of RC_{vir} were launched in Gromacs[203] using the Amber03 force field[204] to model all interactions. Force fields for the RC_{vir} cofactors were constructed based on the atom type information given by Ceccarelli *et al.*[205] Infinite cut-offs for both Lennard-Jones and Coulomb interactions were used to minimize the RC_{vir} structure after which it was embedded into an LDAO detergent micelle consisting of 260 LDAO molecules (Figure 22). Sodium and chloride ions were added to yield a 0.2 M salt solution. The system was equilibrated for 20 ns with restraints on the heavy cofactor and protein atoms and further equilibrated for 14 ns after constraints were lifted.

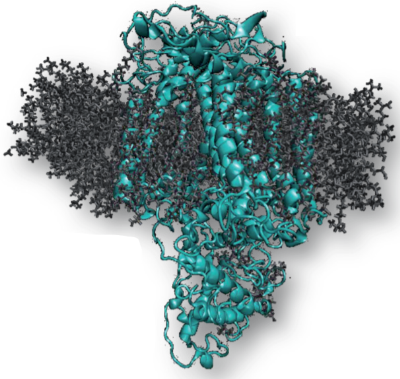

Figure 22. *RC_{vir} surrounded by an LDAO micelle used in the molecular dynamics trajectory simulation.*

After the system was equilibrated, a 5 ns simulation was run with no constraints on either protein or detergents. From the 5-ns simulation new trajectories were branched off every 100 ps. To model the effect of multi-photon absorption in the cofactors every branch was split with one trajectory continuing unaltered (reference) and the other with the energy of 100 photons deposited in the auxiliary chlorophylls and special pair as kinetic energy (excited). Heat flow was monitored by dividing the contents of the trajectory into four subsystems (co-factors, protein, micelle and solvent) and the excess heat in a subsystem was defined as the difference between total kinetic energy and the average total kinetic energy in the subsystem in the 300 K equilibrium simulation. Figure 23a shows the heat flow averaged over 38 trajectories. The timescale given from the molecular dynamics simulation for when the heat content peaks in the protein agrees remarkably well with the non-equilibrated heat component C3, both having t_{max} = 14 ps. Additionally, the timescale of the heat flow from the protein into the solvent predicted from the simulations fits the time dependence of the equilibrated heat component C4 (Figure 23b), concluding that the flow of heat through the protein into the surrounding medium can be experimentally distinguished and are here defined by C3 and C4.

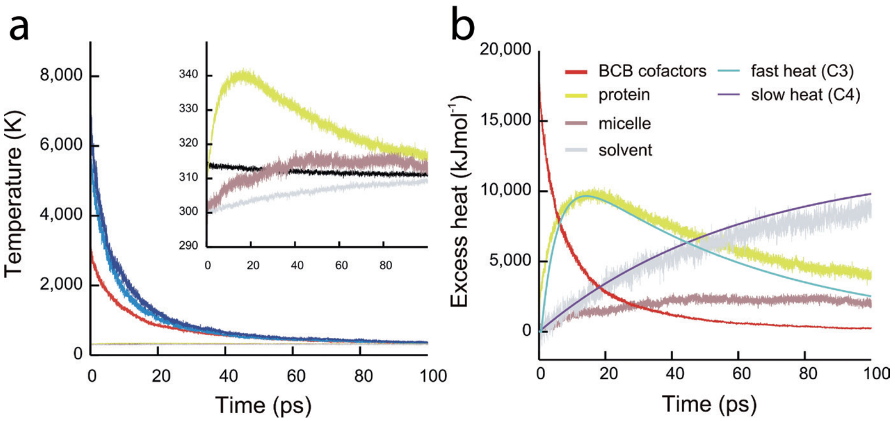

Figure 23. (a) Temperature change over time in the MD trajectories shown for the bacteriochlorophylls (dark and light blue), special pair P_{960} (red), detergent micelle (brown), solvent (gray) and the total system (black). (b) Heat content in the cofactors, protein, micelle and solvent determined from MD trajectories where the time-dependencies for heating components C3 and C4 from the spectral decomposition agree with the MD simulation.

3.4.5 Structural refinement

A set of 1,000 PDB snapshots at 100-fs intervals were generated from 72 reference and excited trajectories (36 each) and theoretical difference WAXS spectra $\Delta S_{\mathrm{theory}}(q)$ between reference and excited PDB structures were calculated using the Debye formula. For evaluation of agreement between $\Delta S_{\mathrm{theory}}(q)$ and $\Delta S_{\mathrm{experiment}}(q)$ a χ^2-scoring function was used. Fits between experimental data and theoretical scattering shown in Figure 24 used an ensemble average of the 20 best predictions from all 36 trajectories (in total 720 $\Delta S_{\mathrm{theory}}(q)$ curves) for describing the theoretical difference curve.

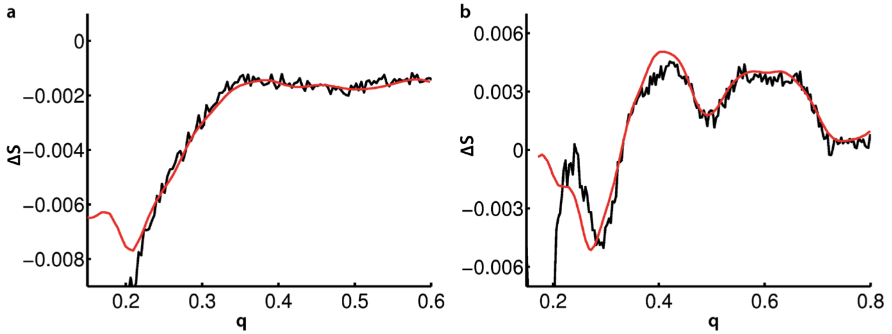

***Figure 24**. (a) Theoretical WAXS difference scattering where movements of only the RC_{vir} cofactors were considered (red) predicted from χ^2 fitting to the ultrafast basis component C1 (black). (b) Theoretical WAXS difference scattering predicted from χ^2 fitting to the protein component C2.*

For the ultrafast C1 component fit, atomic positions for the bacteriochlorophyll cofactors were extracted from the excited and reference trajectories whereas all other atoms were frozen at their t = 0 conformation. Hence, only movements in the cofactors were used to model the experimental data supporting the hypothesis that the C1 component is an effect of ultrafast laser-induced heating of the cofactors.

Structural fitting of the protein C2 component used trajectories with no restraints, using the atomic position of all atoms in the protein to calculate the $\Delta S_{\mathrm{theory}}(q)$ from the reference and photo excited molecular dynamics trajectories. The convincing agreement between $\Delta S_{\mathrm{theory}}(q)$ and $\Delta S_{\mathrm{experiment}}(q)$ for the C2 component shown in Figure 24b strongly makes the case that the structural changes observed in the C2 component represents an ultrafast global conformational change in $\mathrm{RC}_{\mathrm{vir}}$.

3.4.6 3D-interpretation from conserved movements

For a three-dimensional representation of the light-induced structural changes observed in the C2 component, all the 720 pairs for theoretical fitting were used. The individual excited PDB snapshots were not identical, as expected since all atoms vary in the molecular dynamics simulation, so an internal difference matrix calculation approach was adopted to identify recurring structural changes in the 720 pair ensemble. Distance matrices between all the C_α-atoms of the reference and excited structures were calculated as

$$[\Delta\mathrm{C}_\alpha]_{ij} = \left[\mathrm{C}_\alpha^{photexcited} - \mathrm{C}_\alpha^{reference}\right]_{ij} \qquad (15)$$

where *i,j* label all C_α-atoms of $\mathrm{RC}_{\mathrm{vir}}$. The average $\left(\langle[\Delta\mathrm{C}_\alpha]_{ij}\rangle\right)$ and the standard deviation $\left([\sigma(\Delta\mathrm{C}_\alpha)]_{ij}\right)$ from these 720 matrices were calculated and a threshold for characterization of a recurring structural movement was defined by

$$\langle[\Delta\mathrm{C}_\alpha]_{ij}\rangle \geq 0.9 \times [\sigma(\Delta\mathrm{C}_\alpha)]_{ij} \qquad (16)$$

with the additional requirement of at least 20 internal distance changes above this threshold for the C_α atom change to be considered as a conserved movement. This analysis concluded that only C_α atoms moving apart (i.e., $\langle[\Delta\mathrm{C}_\alpha]_{ij}\rangle > 0$) were recurrent in the ensemble and that almost all identified C_α atom conserved movements were identified in the transmembrane domains of subunit H, L and M of $\mathrm{RC}_{\mathrm{vir}}$ as can be seen in Figure 25.

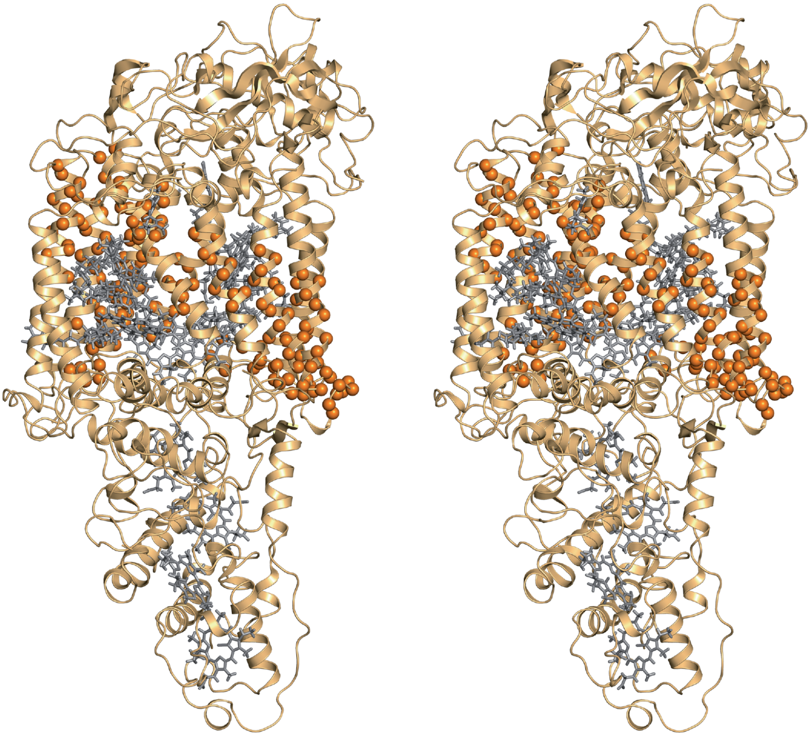

Figure 25. Stereo view of the light-induced movements in RC$_{vir}$ where spheres indicate recurrent movements of C_α atoms within the 720 pair ensemble yielded from χ^2 fitting to the protein component C2. All recurrent movements identified were from C_α atoms moving apart and were primarily located in the transmembrane domains of subunit H, L and M.

3.4.7 Summary Paper III

Time-resolved wide-angle X-ray scattering data were collected at the CXI beamline of the LCLS using a 500 fs laser pulse of 800 nm wavelength from a Ti:Sa femtosecond laser system to pump solubilized RC$_{vir}$ protein sample prior to probing with the XFEL beam. A laser power fluence of 5.3 mJ/mm^2 was used to collect the data, corresponding to hundreds of photons absorbed per RC$_{vir}$ molecule. The use of multi-photon absorption guided the setting of time zero by following the heat signal in the difference WAXS traces with online analysis tools designed and developed for the experiment.

Difference WAXS spectra for a total of 41 time delays were recorded during the ~5 hours of beamtime and spectral decomposition of the data enabled the extraction of four components: an ultrafast component C1, a component C2 describing the light-induced protein structural movement, a non-equilibrated heating component C3 and an equilibrated heating component C4. Molecular dynamics simulations were used to generate 36 trajectories, each split into a reference and excited part, and from these trajectories PDB snapshots were taken at a 100-fs interval. χ^2-scoring between the theoretical difference scattering from reference and excited snapshots and the experimental data yielded 720 pairs (20 from each trajectory) with best agreement and the ensemble average of the 720 pairs fit strikingly well with the experimental data.

A three-dimensional representation of the changes observed was calculated by extracting conserved structural movements from a C_α difference matrix. This shows that the recurrent changes in the C_α-atoms are focused within the transmembrane part of RC_{vir} and that the C_α-atoms moves away from each other, demonstrating an ultrafast light-induced structural response in RC_{vir}.

Paper III presents the first ever TR-WAXS experiment on ultrafast protein structural changes performed at an XFEL and also the first TR—WAXS study to probe low picosecond time-resolution, enabled by the femtosecond X-ray pulse duration from the LCLS XFEL. These X-ray snapshots demonstrate how the energy input into the RC_{vir} cofactors moves from the protein to the surrounding solvent while including a structural deformation that propagates through the protein, enabling a visualization of an ultrafast protein quake in a photosynthetic reaction centre.

3.5 Time-resolved WAXS study of RC_{vir}: ultrafast structural changes in photosynthesis (Paper IV)

A follow-up experiment using time-resolved WAXS were conducted in May 2014 at the CXI[173] beamline of the LCLS in order to elucidate if structural changes were observable in RC_{vir} at a single-photon absorption level.

3.5.1 Data collection

Data were collected on LDAO-solubilized RC_{vir} at a concentration of 30 mg/ml. Protein sample was injected into the XFEL beam using the GDVN microjet system[188] with a 50 µm nozzle, forming a jet with a diameter of ~10 µm. Sample consumption rate was approximately 20 µl/min. A Titanium-sapphire (Ti:Sa) femtosecond laser was used to induce light-activated structural changes in RC_{vir}. The fundamental 800 nm wavelength of the Ti:Sa femtosecond laser pumped the protein sample with a pulse length of 150 fs and was focused to a Gaussian spatial profile with a spot size of 375 µm. The XFEL beam energy was 6 keV (corresponding to a wavelength of 2.07 Å) and was focused on a 10 µm^2 spot aligned on the liquid microjet before the

Rayleigh breakup region. A CSPAD detector comprised of 64 individual tiles with 192 x 185 pixels (pixel size = 110 µm^2) on each tile was used for data recording. Data were recorded with interleaved laser off and laser on images with a repetition rate of 120 Hz. The temporal overlap of the XFEL beam and the optical laser ($\Delta t = 0$) was set with sub-picosecond precision by combining the timing tool at the CXI with online analysis of multi-photon time-resolved WAXS difference data collected around $\Delta t = 0$.

After timing had been established power titration measurements were performed by stepwise decreasing the laser pulse energy from 100 µJ to 1.7 µJ at $\Delta t = 5.0$ ps. Data were collected for approximately two minutes at each of the eleven different powers (100, 90, 80, 60, 50, 40, 30, 20. 10, 5 and 1.7 µJ/pulse) and this cycle was repeated three times. Additional low-fluence data with laser powers < 40 µJ/pulse were then recorded in order to improve the signal-to-noise in the data with weaker difference WAXS signal.

3.5.2 Data analysis

Each diffuse diffraction pattern was radially integrated using the Cheetah software[169] producing laser on and laser off radial stacks for each and every data run. These stack files were then used for further analysis in MatLab. For calculating the $\Delta S(q,\Delta t)$ difference scattering curves an optimized approach was used developed for XFEL WAXS data, implemented to improve the sensitivity to minor changes in the X-ray scattering. By representing each diffraction measurement on a scatter plot (Figure 26) where the sum of the counts in the $q = 0.2$ to $q = 2.4$ Å^{-1} region was plotted against the ratio of the

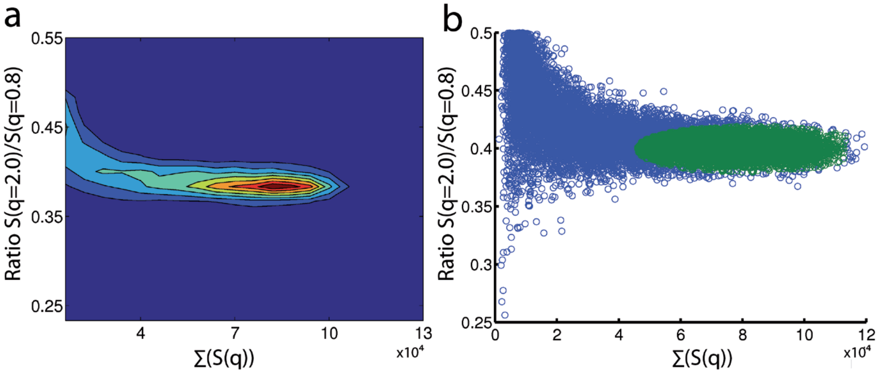

Figure 26. Pulse-by-pulse variations in the XFEL TR-WAXS data. (a) Contour plot describing the relative frequency of the radially integrated measurements sorted according to the sum of all counts between q=0.2–2.4 Å^{-1} (x-axis) versus the ratio between S(q=2.0)/S(q=1.0) (y-axis). (b) To increase the sensitivity in the difference scattering calculations data were selected elliptically around the contour plot maxima for further analysis.

scattering around the water peak (q = 2.0 Å^{-1}) to that of the plateau (q = 0.8 Å^{-1}). The scatter plot effectively illustrated the fluctuations associated with the XFEL beam and microjet changes. Measurements falling outside of an ellipse around the maxima of the 2D distribution were discarded. After normalization over the 2.05 Å^{-1} ≤ q ≤ 2.15 Å-1 domain the difference scattering ($\Delta S_i(q,\Delta t)$) calculation was performed by subtracting the average of the three dark images closest to each and every laser on image on this scatter plot.

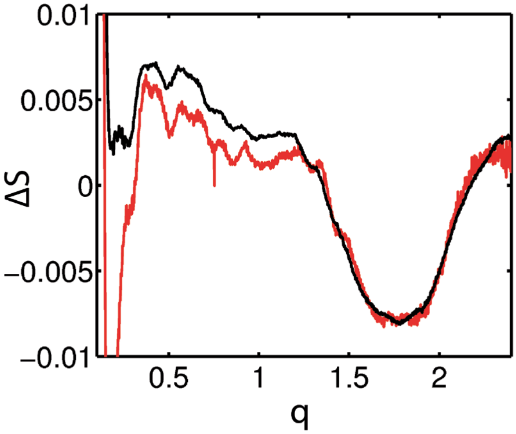

Figure 27. Calibration of the heat deposited in the sample at a laser fluence of 100 μJ/pulse (red). Scaling the difference scattering curve of the Δt=5ps time-point from the data in ***Paper III*** *by 0.58 (black) gives a strong agreement in the q = 1.2-2.4 Å^{-1} domain which is associated with the heating. Since the 2012 data had a thermal jump of 7 K we conclude that the temperature increase in the 100 μJ/pulse data is 4 K.*

3.5.3 Following the heating signal down to a single-photon level

To determine how much heat was deposited to the RC$_{vir}$ samples at the different laser pulse powers, the 100 μJ difference data were calibrated to the 5 ps time-point in the data from **Paper III**, which corresponded to a thermal jump of 7 K. The high-resolution change around the water peak in the X-ray scattering is a direct measurement of the total heat deposited in the sample. Mapping the difference signal in that q-region onto the 5 ps data in **Paper III** showed that the temperature increase was 58 % of that in 2012 (Figure 27). The heating signal was defined as the difference of the average $\Delta S(q)$ over 1.0 Å^{-1} ≤ q ≤ 1.1 Å^{-1} and 1.7 Å^{-1} ≤ q ≤ 1.85 Å^{-1} and the decay of this signal with decreasing laser powers is shown in Figure 28. As can be seen, for laser pump fluences of 30 μJ/pulse or higher the deposited heat increases quadratically with the laser power whereas the dependence is linear for lower laser fluences. At 30 μJ/pulse 15 absorbed photons/molecule gives a temperature change ΔT = 0.13 K and by extrapolation using a linear fit a ΔT = 0.007 K is given for the 1.7 μJ/pulse, corresponding to 0.8 absorbed photons per molecule.

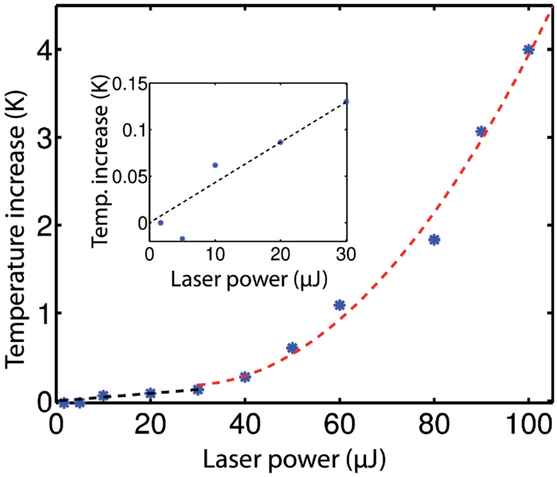

Figure 28. Temperature change determined from the TR-WAXS difference scattering data as a function of the pump laser pulse energy. Below 30 μJ/pulse the heating deposited in the sample increases linearly with the laser power (dashed black line) whereas it above this value becomes quadratic (dashed red line). Inset shows a blow-up of the linear regime.

3.5.4 The structural signal from single-photon absorption

Difference TR-WAXS scattering data from three different laser pulse powers (100, 40 and 1.7 μJ/pulse) are shown in Figure 29. To further highlight the features in the experimental data a high-frequency filter was applied to remove components associated with the statistical noise (black line, Figure 29). By comparing the filtered WAXS difference curves from the data several trends becomes evident (Figure 30). Two peaks centred at *q* = 0.4 Å^{-1} and *q* = 0.6 Å^{-1} clearly dominates the difference data for the highest laser fluence with lower amplitude oscillatory features at higher *q* values also observable. As the laser fluence decreases the dominant peaks are shown to have sub-structure as the peak at *q* = 0.4 Å^{-1} splits into two and the relative heights of the two peaks around *q* = 0.6 Å^{-1} change in relation to each other. At the 20 μJ/pulse and 1.7 μJ/pulse power the oscillatory features

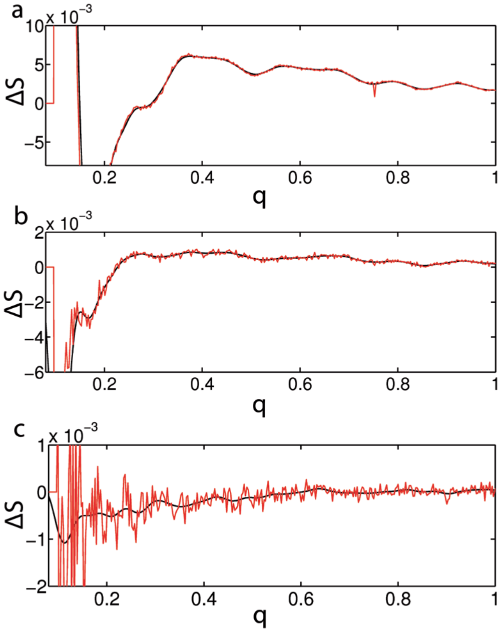

Figure 29. *Difference TR-WAXS data (Δt = 5 ps) of RC$_{vir}$ collected at the CXI. Experimental data (red lines) collected with laser pulse energy at (a) 100 μJ, (b) 40 μJ and (c) 1.7 μJ are shown. Black line shows the data after high-frequency filtering.*

are still present in the filtered data and the oscillations seen correlate in *q* with those observed for higher pump fluence.

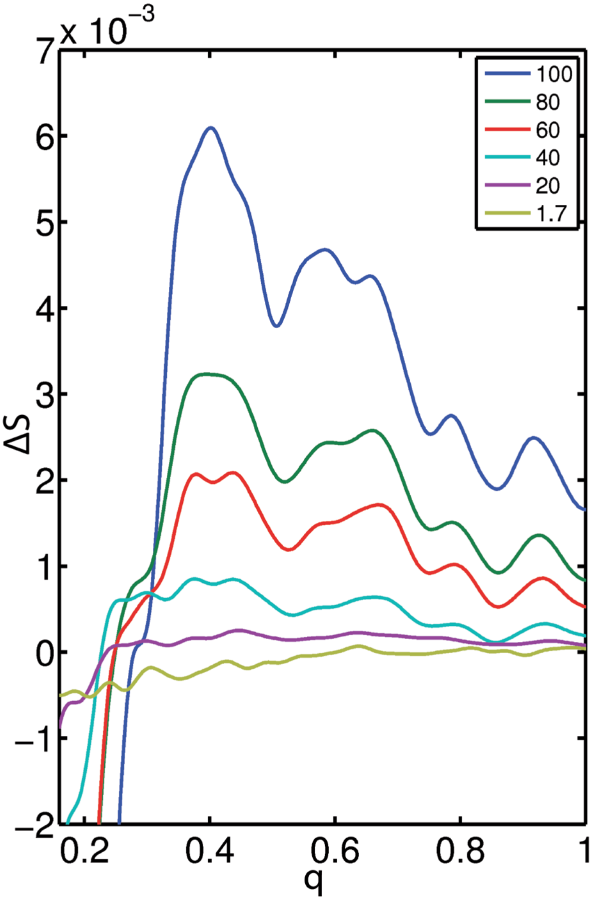

Figure 30. *Filtered difference TR-WAXS data for pump laser pulse energies of 100 μJ (blue), 80 μJ (green), 60 μJ (red), 40 μJ (cyan), 20 μJ (purple) and 1.7 μJ (yellow). The two dominate peaks centred at q = 0.4 Å^{-1} and q = 0.6 Å^{-1} respectively are revealed to have substructure as the pump-laser fluence is decreased. For the lowest pump fluence oscillatory features appear in the filtered data that correlate in q with those observed at higher pump fluence.*

An analysis of the Fourier components of the difference WAXS data was subsequently performed to further investigate the structural signal. By systematically removing each and every component before reconstructing the data by inverse Fourier transform the Fourier components that are most rich in information on protein structural changes could be identified. These were concluded to be components 5 to 7 and the amplitude and phase of these components as a function of laser power are shown in Figure 31. The amplitude clearly increases linearly up to 80 μJ/pulse and the average phase of the three Fourier components varies smoothly with increased laser power, thus providing further indication that the oscillatory features in the time-resolved WAXS signal observable in Figure 30 are correlated for all pump-laser fluences in these data.

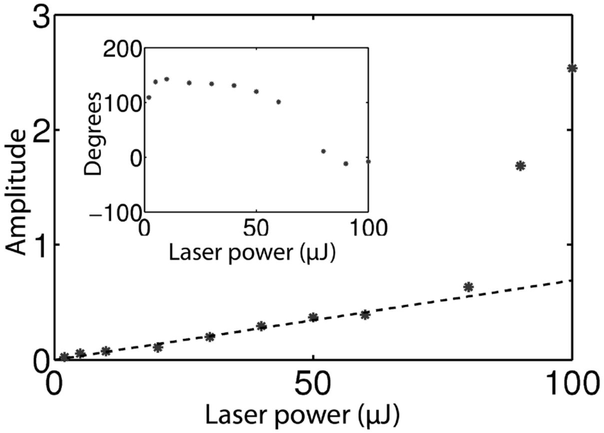

Figure 31. Amplitude of the three Fourier components most strongly correlated with the protein signal visible in Figure 30. An approximately linear dependence of these amplitudes on the laser fluence is seen up to 80 μJ/pulse whereas the average phase of the three components (inset) begins to deviate above 60 μJ/pulse.

In **Paper III**, molecular dynamics was used to produce 720 pairs of excited and ground state RC_{vir} PDB snapshots. The theoretical difference scattering of this ensemble reconstructed the experimental data and by an internal distance matrix approach the structural changes associated with the multi-photon absorption could be identified. To investigate if this description of the protein motion could be applied to the low-fluence data, new 720-pair ensembles were generated with decreasing amplitude of the relative movement between each pair. The difference WAXS scattering of these ensembles goes down linearly with lower amplitude motions (Figure 32), and while the experimental data in **Paper IV** show that the amplitudes of oscillatory features seen in the time-resolved WAXS data does decrease as the laser fluence is reduced, the relative amplitudes of some of the fingerprint features and the *q*-values at which the peaks are visible changes. These changes cannot be seen with direct extrapolation of the protein motion from **Paper III**, thus do not completely describe the observed data. A full three-dimensional interpretation of the structural signal observed at single-photon absorption must therefore await the analysis of the time-resolved serial femtosecond crystallography data collected on RC_{vir} microcrystals during the same experimental run at the LCLS. Nevertheless, the observation of a structural signal from TR-WAXS at a single-photon absorption level is a potentially far-reaching result since it implies that ultrafast global protein structural rearrangements arise under physiological conditions and may therefore play a crucial part in the primary charge separation events of photosynthesis.

3.5.5 Summary Paper IV

Time-resolved WAXS data were collected on solubilized RC_{vir} samples at the CXI beamline of the LCLS in May 2014. A Ti:Sa femtosecond laser system was used to pump the auxiliary chlorophylls of RC_{vir} at 800 nm with a 150 fs laser pulse. Power titration measurements were collected at eleven different laser pulse energies, ranging from 100 μJ/pulse down to 1.7 μJ/pulse. By following the heating signal in the difference WAXS data from these eleven energies it was concluded that a laser power of 1.7 μJ/pulse corresponded to 0.8 photons absorbed per RC_{vir} molecule. Furthermore, the heating signal was shown to increase linearly with the laser fluence up to a laser power of 30 μJ/pulse whereas it above this threshold increased quadratically with the applied laser intensity.

Oscillatory features in the $0.2 \leq q \leq 1.0$ region in the difference WAXS data in **Paper IV** was shown to decay with decreasing laser power, but were still visible for the lowest laser power where 0.8 photons/molecule was absorbed by the RC_{vir} cofactors. The relative amplitude of some of the oscillatory features and, to some extent, the *q*-values at which these peaks are visible were shown to change. Lowering the amplitudes of the protein motion observed for multi-photon absorption could not reproduce these changes, thus the protein quake motion seen in **Paper III** does not completely describe these power-titration data. Though complete analysis of the time-resolved SFX data collected on RC_{vir} microcrystals is needed for a three-dimensional interpretation of the signal observed after single-photon absorption in the RC_{vir}, the result that a fingerprint signal for structural changes is observed in the TR-WAXS difference scattering implies that ultrafast protein structural rearrangements may influence the primary charge separation events of photosynthesis.

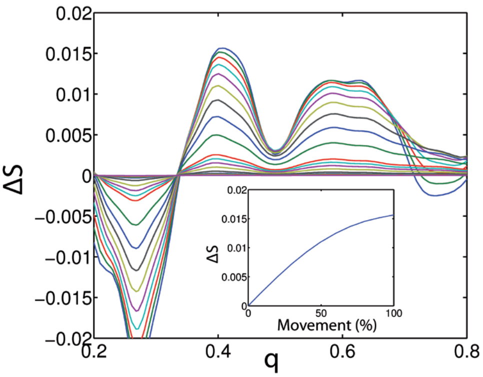

Figure 32. *Theoretical difference WAXS scattering as a function of the extent to which the RC_{vir} protein motion modelled for the protein quake in* ***Paper III*** *is allowed to move. As the amplitude of this expansive motion decreases the predicted peaks in the WAXS scattering becomes smaller but the nature and position of the main features remain unperturbed. Inset shows the dependence of the amplitude of the difference WAXS peak around q = 0.4. A clear linear dependence is observed until the motion becomes relatively large.*

4. Concluding remarks and future perspectives

This thesis has contributed with structural insight on the photosynthetic reaction centre from the purple bacterium *Blastochloris viridis* (RC_{vir}) using the recently developed X-ray free-electron laser (XFEL) at the Linac Coherent Light Source (LCLS). The method of serial femtosecond crystallography (SFX) has been developed for XFEL data collection and was utilized in **Paper I** and **Paper II** for crystallographic studies of RC_{vir}. It was shown that batch crystallization of RC_{vir} in lipidic sponge phase (LSP) could produce micrometre-sized crystals that diffracted, and that the viscous LSP could be injected to the XFEL beam using a gas-driven virtual nozzle microjet system, making LSP a useful tool for SFX. As LCLS commissioned the Coherent X-ray Imaging (CXI) beamline, high-resolution studies with SFX could be achieved and in **Paper II** the first high-resolution structure of a membrane protein is presented, a 3.5 Å structure of RC_{vir}.

In addition to data collection on nano- to micron-sized crystals, the XFEL enables time-resolved studies with temporal resolution not previously achievable at synchrotron sources. Due to the short X-ray pulse length of tens of femtoseconds, time-points in the sub-picosecond regime can be probed. Two such experiments, using time-resolved wide-angle X-ray scattering (TR-WAXS), were conducted on solubilized RC_{vir} at the LCLS. In **Paper III**, a structural deformation propagating through the RC_{vir} protein upon multi-photon absorption of the cofactors could be followed. This visualization of a protein quake is the first direct structural study of such a motion in photosynthesis.

In **Paper IV**, the pump laser intensity was decreased in order to probe for single-photon structural changes in RC_{vir}. At the lowest laser fluence, corresponding to less then one photon absorbed per RC_{vir} molecule, a structural signal could be seen in the TR-WAXS difference scattering data. This result implicates that ultrafast structural changes within the photosynthetic reaction centre may influence the primary charge separation events of photosynthesis.

The advent of the XFEL has provided new possibilities for structural studies of proteins. The SFX method has been successfully used for solving a number of protein structures, with data collected on crystals too small for conventional protein crystallography. A limitation in the first SFX studies was the need for large amounts of sample, not achievable for many membrane protein targets. As the microjet injection systems develop and improve, the sample consumption required for an SFX experiment will reach more feasible levels. For time-resolved methods, the XFEL opens up the possibility to study protein dynamics with significantly improved temporal resolution, both on crystallized and solubilized protein. The X-ray brilliance and high repetition rate of the XFEL makes methods such as TR-WAXS a very good tool for such studies, since data on individual time-points are collected very rapidly with better signal-to-noise compared to TR-WAXS data recorded at synchrotrons. Additionally, time-resolved studies on both reversible and irreversible systems are made possible by the use of the microjet injection system, which deliver sample to the XFEL beam at a high velocity, ensuring that each XFEL pulse interacts with fresh sample.

In the early stages of the XFEL, many proof-of-principle experiments were performed to validate the use of XFEL radiation for structural studies. With several new XFELs being commissioned in the near future, the accessibility of XFEL beamtime will increase and the focus of the life science experiments will move towards biological impact, with the potential to dramatically improve our understanding of protein structure and function.

5. Populärvetenskaplig sammanfattning

Proteiner är molekyler som möjliggör många av de essentiella reaktioner som sker i alla levande celler. Membranproteiner är proteiner som sitter i eller knutna till det yttre skiktet som omger cellen och kan ha varierande uppgifter såsom transport av ämnen genom membranet, kommunicera med omgivande celler eller generera ett cellsvar på yttre stimuli. Då membranproteiner är för små för att kunna visualiseras i ljusmikroskop använder man röntgenstrålning som har kortare våglängd än vanligt ljus för att få en tredimensionell bild. Den vanligaste metoden för att skapa en sådan bild är att kristallisera proteinet och sedan bestråla kristallen med röntgenljus. I en kristall ordnas proteiner i ordnade mönster och på grund av denna regelbundenhet så sprids röntgenstrålning som träffar kristallen i specifika mönster. Utifrån dessa mönster kan sedan en tredimensionell struktur av proteinet erhållas med hjälp av matematiska metoder.

En röntgen fri-elektronslaser (XFEL) är den kraftfullaste röntgenkällan som finns tillgänglig idag och tack vare att strålningen som genereras vid en sådan är mer än en miljard gånger mer intensiv än tidigare utvecklade röntgenkällor så kan data samlas in på betydligt mindre kristaller än vad som tidigare krävts. Tack vare att röntgenpulsen är extremt kort vid en XFEL möjliggörs även tidsupplösta studier av membranproteiner där reaktioner som sker på femtosekundsskalan kan undersökas (fs = 10^{-15} s).

Denna avhandling har syftat att med hjälp av XFEL-strålning undersöka det fotosyntetiska reaktionscentret i *Bl. viridis* (RC_{vir}) genom att utveckla nya metoder för att få stora mängder små kristaller för datainsamling. Ur dessa data har sedan den tredimensionella strukturen av RC_{vir} kunnat bestämmas med hög detaljrikedom.

Dessutom har studier på RC_{vir} i löslig form (ej kristaller) gjorts där en laserpuls används för att aktivera RC_{vir} och sedan med hjälp av tidsupplöst vidvinkelsröntgenspridning (WAXS) följa hur proteinet genomgår strukturella förändringar efter dess aktivering. Genom att samla in spridningsmönster med olika tidsavstånd mellan den optiska lasern och röntgenstrålningen erhålls stillbilder som tillsammans bygger upp en filmsekvens. Resultaten visar på hur proteinet genomgår en ”proteinbävning” som likt vågorna från en jordbävnings epicentrum sprider sig utåt då det belysts med intensiv laserpuls.

För att undersöka huruvida RC_{vir}-proteinet genomgår några strukturella förändringar under biologiskt relevanta förhållanden (1 foton absorberad per proteinmolekyl) utfördes ett andra experiment där styrkan på lasern som används för fotoaktivering gradvis sänktes ner till singel-fotonsnivå. Resultaten visar att även på den nivå som motsvarar 0.8 absorberade fotoner/molekyl så går det att detektera en signal i WAXS som indikerar att ultrasnabba strukturella förändringar i ett fotosyntesprotein kan vara förklaringen till den extrema effektivitet dessa proteiner uppvisar vid omvandling av solenergi till kemisk lagrad energi.

6. Acknowledgements

It's rather hard to grasp that half a decade has passed since I first started working in the Lundberg lab. It's been a period where I've learnt a tremendous deal on many exciting topics and writing this thesis has sort of put it all in perspective. All the hard days of work actually paid off in the end. I could never have done it alone and so would like to thank the following people:

Probably best to start with **Richard**, my supervisor. The interview you held with me during my first beamtime trip to the SLS was short and concise: "You want it?", "Yep", "Good". I've always felt you've believed in me and the work we've been doing even at times I myself was doubtful. Calling you a hands-on supervisor with detailed plans for what's to be done the upcoming weeks couldn't be further from the truth, and that is one of the many reasons why I've enjoyed working in your group. Your vision sets the target and then it's up to us to reach the goal and your enthusiasm about the science you do does rub off, which is a big plus. It's easy to become a beamtime junkie in your group, and the trips to X-ray facilities around the globe are actually one of the things I'll miss the most (and the thousands of frequent flyer miles that comes along…).

Sebastian, my co-supervisor. I started at the Lundberg lab as a master student for you when you were "only" a post-doc. Five years later and now you're group is bigger than Richard's, so you're obviously doing many things right. I'm sure you'll continue to be successful and soon can add professor as a title. Thank you for taking me on as a master student and all your help on the WAXS data analysis.

Martin, thanks for being my examiner. You've always answered any mail swiftly, be it about course points or defence bureaucracy. Also, thank you for chairing the defence.

Jan Davidsson, thank you for always finding time to join us with all your laser know-how for our beamtime travels. Though the food is not always to your liking, you're always great to have along, both professionally and socially.

My fellow beamtime junkies **Linda** and **Erik**. Linda, though you did your best to scare me away from the office at first you eventually accepted that you had to share the space. After that major breakthrough our friendship has grown and we've shared many fun (and somewhat cloudy…) times on our after-beamtime vacations. You've also been very good to work with as you always complete whatever task you set at hand. Erik, I went in your footsteps with the WAXS and you taught me a great deal about data analysis in general and MatLab in particular. You've always been open for grabbing a beer and some accompanying nice conservation, first at BC then in the Bay Area. And your bongo drum skills are out of this world! We should definitely tour again sometime.

Cecilia, Jennie and **Rob**. Cecilia, it's been great working with you on the RC data. You're very clever and always manage to come up with scripts that both work and are somewhat pedagogical ☺. Jennie, it's impressive that you manage to work hard both at the lab and at Friskis and also found time to run the Ph.D. student council for three years. Good luck with the WAXS data! Rob, you've been a great addition to the Neutze group. Your loyal love for me (equal only to the love you have for Mrs. Thatcher) is something I've truly enjoyed. For being a posh, boarding school Englishman you're actually quite all right. I'm sure you'll be able to bring the RC project home with your XFEL SFX data. And thanks for toasting at the party.

Rajiv, Peter, Gisela and **Amit**. It's been great getting to know you and working with you at the LCLS. Peter, I hope you'll enjoy your new adventures down under. Gisela, now that I'm leaving you'll probably going to have to start paying for the coffee. Rajiv, we'll always have Hamburg.

~~Petra~~ Pia, finally you got a Ph.D. student position after years of slaving for Linda *et al*. Good for you, and good for us since you're always nice company. I promise I'll attend a Roller Derby game soon (if you promise to toss the tiny ones). Thank you also for being my toast madam.

The (ever increasing) Westenhoff group: **Alex B, Sebastian P, Maria, Stephan, Rafal, Heikki, Oskar, Madde.** Good luck with your projects and thanks for many pleasant coffee breaks. Oskar, you know how you'll look in a couple of year's time. Enjoy your youth!

Kristina, Gergely, Örjan, Ida, Annette, Weixiao, Rosie, Elin, Rhawnie, Susanna, Mike, Maria S, Karin R. Some of you have left for new adventures but to me you'll always be part of the Lundberg core. Keep up the good work. Karin, I hope your new project(s) in Denmark turn out fruitful. Thanks for the nice (and angry) BC company.

All the new(ish) people at the lab: **Petra, Alex J, Rebecka, Maria-José, Emil, Ashley, Matthijs, Parveen.** Welcome to the Lundberg Lab, I hope you have a pleasant journey. Emil, you've been a good office mate replacement.

The “old gang”: **Mikael, Gerhard, Anna, Fredrik.** Thanks for introducing me to the Ph.D. student world way back when.

Linnéa and **Johan**. You held the NMR flag high in X-ray land and were always fun to have on BC or for a small chitchat in the corridor.

Lars and **Bruno**, without you the Lundberg lab would not get much work done.

Also I’d like to thank all the master students that have come and gone during my years at the Lundberg lab. You all contributed to the working atmosphere at the lab and some of you were (very) frequent guests at the BC table. **Vlad**, good luck with your Ph.D. studies in Stockholm.

Alla **vänner**, tack för att ni gett mig många roliga stunder utanför jobbets murar.

Min **familj**, för att ni alltid finns där för mig och kan erbjuda surströmming, snaps och skratt när så stunden kräver det.

Åsa. Tack för att du stått ut med allt mitt flängande jorden runt och mina ibland udda arbetstider. Nu när detta är över så kanske du får se mig i dagsljus snart igen! Ska bli kul att börja nya äventyr i januari när vi blir tre. Älskar dig!

7. References

1. von Heijne, G. The membrane protein universe: what's out there and why bother? *Journal of Internal Medicine* **261**, 543-557 (2007).
2. Drews, J. Drug discovery: A historical perspective. *Science* **287**, 1960-1964 (2000).
3. Lundstrom, K. Structural genomics. *Molecular Biotechnology* **34**, 205-212 (2006).
4. Muller, D.J., Wu, N. & Palczewski, K. Vertebrate membrane proteins: Structure, function, and insights from biophysical approaches. *Pharmacological Reviews* **60**, 43-78 (2008).
5. Krauss, N. Mechanisms for photosystems I and II. *Current Opinion in Chemical Biology* **7**, 540-550 (2003).
6. Xiong, J., Fischer, W.M., Inoue, K., Nakahara, M. & Bauer, C.E. Molecular evidence for the early evolution of photosynthesis. *Science* **289**, 1724-1730 (2000).
7. Xiong, J. & Bauer, C.E. A cytochrome b origin of photosynthetic reaction centers: an evolutionary link between respiration and photosynthesis. *Journal of Molecular Biology* **322**, 1025-1037 (2002).
8. Xiong, J. & Bauer, C.E. Complex evolution of photosynthesis. *Annual Review of Plant Biology* **53**, 503-521 (2002).
9. Gray, M.W. The endosymbiont hypothesis revisited. *International Review of Cytology-a Survey of Cell Biology* **141**, 233-357 (1992).
10. Morden, C.W., Delwiche, C.F., Kuhsel, M. & Palmer, J.D. Gene phylogenies and the endosymbiotic origin of plastids. *Biosystems* **28**, 75-90 (1992).
11. Nelson, N. & Ben-Shem, A. The complex architecture of oxygenic photosynthesis. *Nature Reviews Molecular Cell Biology* **5**, 971-982 (2004).
12. Drews, G. & Niederman, R.A. Membrane biogenesis in anoxygenic photosynthetic prokaryotes. *Photosynthesis Research* **73**, 87-94 (2002).
13. Chitnis, P.R. Photosystem I: Function and physiology. *Annual Review of Plant Physiology and Plant Molecular Biology* **52**, 593-626 (2001).
14. Jordan, P. et al. Three-dimensional structure of cyanobacterial photosystem I at 2.5 angstrom resolution. *Nature* **411**, 909-917 (2001).
15. Buttner, M. et al. Photosynthetic reaction center genes in green sulfur bacteria and in photosystem-1 are related. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 8135-8139 (1992).
16. Michel, H. & Deisenhofer, J. Relevance of the photosynthetic reaction center from purple bacteria to the structure of photosystem-II *Biochemistry* **27**, 1-7 (1988).
17. Michel, H. et al. The light and medium subunits of the photosynthetic reaction center from *Rhodopseudomonas viridis* - isolation of the genes, nucleotide and amino-acid-sequence. *Embo Journal* **5**, 1149-1158 (1986).
18. Trebst, A. The topology of the plastoquinone and herbicide binding peptides of photosystem-II in the thylakoid membrane. *Zeitschrift Fur Naturforschung C-a Journal of Biosciences* **41**, 240-245 (1986).
19. Yano, J. et al. Where water is oxidized to dioxygen: Structure of the photosynthetic Mn4Ca cluster. *Science* **314**, 821-825 (2006).
20. Yocum, C.F. The calcium and chloride requirements of the O-2 evolving complex. *Coordination Chemistry Reviews* **252**, 296-305 (2008).
21. Hillier, W. & Babcock, G.T. Photosynthetic reaction centers. *Plant Physiology* **125**, 33-37 (2001).
22. Kok, B., Forbush, B. & McGloin, M. Cooperation of charges in photosynthetic O2 evolution .1. A linear 4-step mechanism. *Photochem. Photobiol.* **11**, 457-& (1970).

23. Lupinkova, L., Metz, J.G., Diner, B.A., Vass, I. & Komenda, J. Histidine residue 252 of the Photosystem II D1 polypeptide is involved in a light-induced cross-linking of the polypeptide with the alpha subunit of cytochrome b-559: study of a site-directed mutant of Synechocystis PCC 6803. *Biochimica Et Biophysica Acta-Bioenergetics* **1554**, 192-201 (2002).
24. Schubert, W.D. et al. A common ancestor for oxygenic and anoxygenic photosynthetic systems: A comparison based on the structural model of photosystem I. *Journal of Molecular Biology* **280**, 297-314 (1998).
25. Svensson, B. et al. A model for the photosystem II reaction center core including the structure of the primary donor P-680. *Biochemistry* **35**, 14486-14502 (1996).
26. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. X-ray structure analysis of a membrane protein complex - electron density map at 3A resolution and a model of the chromophores of the photosynthetic reaction center from *Rhodopseudomonas viridis*. *Journal of Molecular Biology* **180**, 385-398 (1984).
27. Garfield, E. The 1988 Nobel-prize in chemistry goes to Deisenhofer, Johann, Huber, Robert, and Michel, Hartmut for elucidating photosynthetic processes. *Current Contents* **22**, 3-8 (1989).
28. Raman, P., Cherezov, V. & Caffrey, M. The membrane protein data bank. *Cellular and Molecular Life Sciences* **63**, 36-51 (2006).
29. Coolidge, W.D. The development of modern roentgen-ray generating apparatus. *Am. J. Roentgenology* **24**, 605-620 (1930).
30. Skarzynski, T. Collecting data in the home laboratory: evolution of X-ray sources, detectors and working practices. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **69**, 1283-1288 (2013).
31. Arndt, U.W. Focusing optics for laboratory sources in X-ray crystallography. *Journal of Applied Crystallography* **23**, 161-168 (1990).
32. Powers, R.A. Rotating Anode X-ray Tube. *California and western medicine* **48**, 339-341 (1938).
33. Veksler, V.I. A new method of accelerating relativistic particles. *Comptes Rendus (Dokaldy) de l'Academie Sciences de l'URSS* **43**, 329-331 (1944).
34. McMillan, E.M. The synchrotron - a proposed high energy particle accelerator. *Physical Review* **68**, 143-144 (1945).
35. Rosenbau.G, Holmes, K.C. & Witz, J. Synchrotron radiation as a source for X-ray diffraction. *Nature* **230**, 434-& (1971).
36. Helliwell, J.R. et al. Central data-collection facility for protein crystallography, small-angle diffraction and scattering at the Daresbury Laboratory synchrotron radiation source (SRS), England. *Journal of Physics E-Scientific Instruments* **15**, 1363-1372 (1982).
37. Haensel, R. European synchrotron radiation facility (ESRF). *Review of Scientific Instruments* **63**, 1571-1574 (1992).
38. Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic-resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics* **28**, 171-193 (1995).
39. Henderson, R. Cryoprotection of protein crystals against radiation-damage in electron and X-ray diffraction. *Proceedings of the Royal Society B-Biological Sciences* **241**, 6-8 (1990).
40. Owen, R.L., Rudino-Pinera, E. & Garman, E.F. Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4912-4917 (2006).
41. Solem, J.C. Imaging biological specimens with high-intensity soft X-rays. *Journal of the Optical Society of America B-Optical Physics* **3**, 1551-1565 (1986).

42. Doniach, S. Studies of the structure of matter with photons from an X-ray free-electron laser. *Journal of Synchrotron Radiation* **3**, 260-267 (1996).
43. 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).
44. Emma, P. et al. First lasing and operation of an angstrom-wavelength free-electron laser. *Nature Photonics* **4**, 641-647 (2010).
45. McNeil, B.W.J. & Thompson, N.R. X-ray free-electron lasers. *Nature Photonics* **4**, 814-821 (2010).
46. Barty, A. et al. Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nature Photonics* **6**, 35-40 (2012).
47. Siewert, F. et al. Ultra-precise characterization of LCLS hard X-ray focusing mirrors by high resolution slope measuring deflectometry. *Optics Express* **20**, 4525-4536 (2012).
48. Ihee, H., Wulff, M., Kim, J. & Adachi, S.-i. Ultrafast X-ray scattering: structural dynamics from diatomic to protein molecules. *International Reviews in Physical Chemistry* **29**, 453-520 (2010).
49. Pile, D. X-rays first light from SACLA. *Nature Photonics* **5**, 456-457 (2011).
50. Decking, W. The European XFEL project. (2007).
51. Oberta, P., Flechsig, U. & Abela, R. The SwissFEL facility and its preliminary optics beamline layout. *Advances in X-Ray Free-Electron Lasers: Radiation Schemes, X-Ray Optics, and Instrumentation* **8078** (2011).
52. Drenth, J. Principles of Protein X-Ray Crystallography, 2nd edition. *(Springer Advanced Texts In Chemistry)* (1999).
53. Berman, H.M. et al. The Protein Data Bank. *Nucleic Acids Research* **28**, 235-242 (2000).
54. Schlichting, I., Berendzen, J., Phillips, G.N. & Sweet, R.M. Crystal-structure of photolyzed carbonmonoxy-myoglobin. *Nature* **371**, 808-812 (1994).
55. Kack, H., Gibson, K.J., Lindqvist, Y. & Schneider, G. Snapshot of a phosphorylated substrate intermediate by kinetic crystallography. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5495-5500 (1998).
56. Edman, K. et al. High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature* **401**, 822-826 (1999).
57. Chu, K. et al. Structure of a ligand-binding intermediate in wild-type carbonmonoxy myoglobin. *Nature* **403**, 921-923 (2000).
58. Wilmouth, R.C. et al. X-ray snapshots of serine protease catalysis reveal a tetrahedral intermediate. *Nature Structural Biology* **8**, 689-694 (2001).
59. Berglund, G.I. et al. The catalytic pathway of horseradish peroxidase at high resolution. *Nature* **417**, 463-468 (2002).
60. Katona, G. et al. Conformational regulation of charge recombination reactions in a photosynthetic bacterial reaction center. *Nature Structural & Molecular Biology* **12**, 630-631 (2005).
61. Neelon, K., Roberts, M.F. & Stec, B. Crystal Structure of a Trapped Catalytic Intermediate Suggests that Forced Atomic Proximity Drives the Catalysis of mIPS. *Biophysical Journal* **101**, 2816-2824 (2011).
62. Genick, U.K. et al. Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science* **275**, 1471-1475 (1997).
63. Perman, B., Wulff, M., Hellingwerf, K. & Moffat, K. Crystallography of a photocycle intermediate - Response. *Science* **281**, 1964-1964 (1998).
64. Ren, Z. et al. A molecular movie at 1.8 angstrom resolution displays the photocycle of photoactive yellow protein, a eubacterial blue-light receptor, from nanoseconds to seconds. *Biochemistry* **40**, 13788-13801 (2001).

65. Anderson, S. et al. Chromophore conformation and the evolution of tertiary structural changes in photoactive yellow protein. *Structure* **12**, 1039-1045 (2004).
66. Schmidt, M. et al. Protein kinetics: Structures of intermediates and reaction mechanism from time-resolved x-ray data. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4799-4804 (2004).
67. 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).
68. 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 of the United States of America* **109**, 19256-19261 (2012).
69. Srajer, V. et al. Photolysis of the carbon monoxide complex of myoglobin: Nanosecond time-resolved crystallography. *Science* **274**, 1726-1729 (1996).
70. Srajer, V. et al. Protein conformational relaxation and ligand migration in myoglobin: A nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction. *Biochemistry* **40**, 13802-13815 (2001).
71. Schotte, F. et al. Watching a protein as it functions with 150-ps time-resolved X-ray crystallography. *Science* **300**, 1944-1947 (2003).
72. Bourgeois, D. et al. Complex landscape of protein structural dynamics unveiled by nanosecond Laue crystallography. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8704-8709 (2003).
73. Bourgeois, D. et al. Extended subnanosecond structural dynamics of myoglobin revealed by Laue crystallography. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4924-4929 (2006).
74. Schmidt, M. et al. Ligand migration pathway and protein dynamics in myoglobin: A time-resolved crystallographic study on L29W MbCO. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11704-11709 (2005).
75. Schotte, F. et al. Real-time tracking of CO migration and binding in the alpha and beta subunits of human hemoglobin via 150-ps time-resolved Laue crystallography. *Chemical Physics* **422**, 98-106 (2013).
76. 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).
77. Wohri, A.B. et al. Light-Induced Structural Changes in a Photosynthetic Reaction Center Caught by Laue Diffraction. *Science* **328**, 630-633 (2010).
78. Katona, G. et al. Raman-assisted crystallography reveals end-on peroxide intermediates in a nonheme iron enzyme. *Science* **316**, 449-453 (2007).
79. McGeehan, J.E., Carpentier, P., Royant, A., Bourgeois, D. & Ravelli, R.B.G. X-ray radiation-induced damage in DNA monitored by online Raman. *Journal of Synchrotron Radiation* **14**, 99-108 (2007).
80. Bourgeois, D., Katona, G., de Rosny, E. & Carpentier, P. Raman-assisted X-ray crystallography for the analysis of biomolecules. *Methods in molecular biology (Clifton, N.J.)* **544**, 253-267 (2009).
81. McGeehan, J. et al. Colouring cryo-cooled crystals: online microspectrophotometry. *Journal of Synchrotron Radiation* **16**, 163-172 (2009).
82. Schlichting, I. & Goody, R.S. in Macromolecular Crystallography, Pt B, Vol. 277. (eds. C.W. Carter & R.M. Sweet) 467-490 (1997).
83. Svergun, D.I. & Koch, M.H.J. Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics* **66**, 1735-1782 (2003).
84. Svergun, D.I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* **76**, 2879-2886 (1999).

85. Konarev, P.V., Petoukhov, M.V. & Svergun, D.I. MASSHA - a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. *Journal of Applied Crystallography* **34**, 527-532 (2001).
86. Petoukhov, M.V., Konarev, P.V., Kikhney, A.G. & Svergun, D.I. ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. *Journal of Applied Crystallography* **40**, S223-S228 (2007).
87. Hirai, M., Iwase, H., Hayakawa, T., Miura, K. & Inoue, K. Structural hierarchy of several proteins observed by wide-angle solution scattering. *Journal of Synchrotron Radiation* **9**, 202-205 (2002).
88. Makowski, L. Characterization of proteins with wide-angle X-ray solution scattering (WAXS). *Journal of structural and functional genomics* **11**, 9-19 (2010).
89. Neutze, R. et al. Visualizing photochemical dynamics in solution through picosecond x-ray scattering. *Physical Review Letters* **87** (2001).
90. Plech, A. et al. Visualizing chemical reactions in solution by picosecond x-ray diffraction. *Physical Review Letters* **92** (2004).
91. Davidsson, J. et al. Structural determination of a transient isomer of CH2I2 by picosecond x-ray diffraction. *Physical Review Letters* **94** (2005).
92. Ihee, H. et al. Ultrafast x-ray diffraction of transient molecular structures in solution. *Science* **309**, 1223-1227 (2005).
93. Cammarata, M. et al. Impulsive solvent heating probed by picosecond x-ray diffraction. *Journal of Chemical Physics* **124** (2006).
94. Georgiou, P. et al. Picosecond calorimetry: Time-resolved x-ray diffraction studies of liquid CH2Cl2. *Journal of Chemical Physics* **124** (2006).
95. Cammarata, M. et al. Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. *Nature methods* **5**, 881-886 (2008).
96. Andersson, M. et al. Structural Dynamics of Light-Driven Proton Pumps. *Structure* **17**, 1265-1275 (2009).
97. Ahn, S., Kim, K.H., Kim, Y., Kim, J. & Ihee, H. Protein Tertiary Structural Changes Visualized by Time-Resolved X-ray Solution Scattering. *Journal of Physical Chemistry B* **113**, 13131-13133 (2009).
98. 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 of the United States of America* **107**, 7281-7286 (2010).
99. 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).
100. Kim, K.H. et al. Direct observation of myoglobin structural dynamics from 100 picoseconds to 1 microsecond with picosecond X-ray solution scattering. *Chemical Communications* **47**, 289-291 (2011).
101. Malmerberg, E. et al. Time-Resolved WAXS Reveals Accelerated Conformational Changes in Iodoretinal-Substituted Proteorhodopsin. *Biophysical Journal* **101**, 1345-1353 (2011).
102. Kim, K.H. et al. Direct Observation of Cooperative Protein Structural Dynamics of Homodimeric Hemoglobin from 100 ps to 10 ms with Pump-Probe X-ray Solution Scattering. *Journal of the American Chemical Society* **134**, 7001-7008 (2012).
103. Kim, T.W. et al. Protein Structural Dynamics of Photoactive Yellow Protein in Solution Revealed by Pump-Probe X-ray Solution Scattering. *Journal of the American Chemical Society* **134**, 3145-3153 (2012).
104. Takala, H. et al. Signal amplification and transduction in phytochrome photosensors. *Nature* **509**, 245-+ (2014).

105. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. Structure of the protein subunits in the photosynthetic reacion center of *Rhodopseudomonas viridis* at 3A resolution. *Nature* **318**, 618-624 (1985).
106. Leonova, M.M., Fufina, T.Y., Vasilieva, L.G. & Shuvalov, V.A. Structure-function investigations of bacterial photosynthetic reaction centers. *Biochemistry-Moscow* **76**, 1465-1483 (2011).
107. Feniouk, B.A., Cherepanov, D.A., Junge, W. & Mulkidjanian, A.Y. Coupling of proton flow to ATP synthesis in Rhodobacter capsulatus: F0F1-ATP synthase is absent from about half of chromatophores. *Biochimica Et Biophysica Acta-Bioenergetics* **1506**, 189-203 (2001).
108. Drews, G. & Giesbrec.P *Rhodopseudomonas viridis* nov spec ein neu isoliertes obligat phototrophes bakterium. *Archiv Fur Mikrobiologie* **53**, 255-& (1966).
109. Miller, K.R. & Jacob, J.S. The *Rhodopseudomonas viridis* photosynthetic membrane - arrangement *in situ*. *Archives of Microbiology* **142**, 333-339 (1985).
110. Lang, F.S. & Oesterhelt, D. Microaerophilic growth and induction of the photosynthetic reaction center in *Rhodopseudomonas viridis Journal of Bacteriology* **171**, 2827-2834 (1989).
111. Edebo, L. A new press for the disruption of micro-organisms and other cells. *Journal of Biochemical and Microbiological Technology and Engineering* **2**, 453-479 (1960).
112. Alderton, G., Ward, W.H. & Fevold, H.L. Isolation of lysozyme from egg white. *Journal of Biological Chemistry* **157**, 43-58 (1945).
113. Alderton, G., Lewis, J.C. & Fevold, H.L. The releationship of lysozyme, biotin and avidin. *Science* **101**, 151-152 (1945).
114. Katona, G., Andreasson, U., Landau, E.M., Andreasson, L.E. & Neutze, R. Lipidic cubic phase crystal structure of the photosynthetic reaction centre from Rhodobacter sphaeroides at 2.35 angstrom resolution. *Journal of Molecular Biology* **331**, 681-692 (2003).
115. Tiefenbrunn, T. et al. High Resolution Structure of the ba3 Cytochrome c Oxidase from Thermus thermophilus in a Lipidic Environment. *Plos One* **6** (2011).
116. Lyons, J.A. et al. Structural insights into electron transfer in caa(3)-type cytochrome oxidase. *Nature* **487**, 514-518 (2012).
117. Landau, E.M. & Rosenbusch, J.P. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 14532-14535 (1996).
118. Kato, H.E. et al. Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* **482**, 369-U115 (2012).
119. Kolbe, M., Besir, H., Essen, L.O. & Oesterhelt, D. Structure of the light-driven chloride pump halorhodopsin at 1.8 angstrom resolution. *Science* **288**, 1390-1396 (2000).
120. Ran, T. et al. Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **69**, 1965-1980 (2013).
121. Vogeley, L. et al. Anabaena sensory rhodopsin: A photochromic color 0 sensor at 2.0 angstrom. *Science* **306**, 1390-1393 (2004).
122. Royant, A. et al. X-ray structure of sensory rhodopsin II at 2.1-angstrom resolution. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10131-10136 (2001).
123. Wada, T. et al. Crystal Structure of the Eukaryotic Light-Driven Proton-Pumping Rhodopsin, Acetabularia Rhodopsin II, from Marine Alga. *Journal of Molecular Biology* **411**, 986-998 (2011).

124. Jaakola, V.-P. et al. The 2.6 Angstrom Crystal Structure of a Human A(2A) Adenosine Receptor Bound to an Antagonist. *Science* **322**, 1211-1217 (2008).
125. Cherezov, V. et al. High-resolution crystal structure of an engineered human beta(2)-adrenergic G protein-coupled receptor. *Science* **318**, 1258-1265 (2007).
126. Wu, B. et al. Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science* **330**, 1066-1071 (2010).
127. Chien, E.Y.T. et al. Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. *Science* **330**, 1091-1095 (2010).
128. Srivastava, A. et al. High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. *Nature* **513**, 124-+ (2014).
129. Shimamura, T. et al. Structure of the human histamine H-1 receptor complex with doxepin. *Nature* **475**, 65-U82 (2011).
130. Kruse, A.C. et al. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* **482**, 552-556 (2012).
131. White, J.F. et al. Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508-+ (2012).
132. Wu, H. et al. Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* **485**, 327-U369 (2012).
133. Zhang, C. et al. High-resolution crystal structure of human protease-activated receptor 1. *Nature* **492**, 387-+ (2012).
134. Zhang, J. et al. Agonist-bound structure of the human P2Y(12) receptor. *Nature* **509**, 119-+ (2014).
135. Wang, C. et al. Structural Basis for Molecular Recognition at Serotonin Receptors. *Science* **340**, 610-614 (2013).
136. Hanson, M.A. et al. Crystal Structure of a Lipid G Protein-Coupled Receptor. *Science* **335**, 851-855 (2012).
137. Hollenstein, K. et al. Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature* **499**, 438-+ (2013).
138. Siu, F.Y. et al. Structure of the human glucagon class B G-protein-coupled receptor. *Nature* **499**, 444-+ (2013).
139. Dore, A.S. et al. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature* **511**, 557-+ (2014).
140. Wang, C. et al. Structural basis for Smoothened receptor modulation and chemoresistance to anticancer drugs. *Nature Communications* **5** (2014).
141. Zhou, X. et al. Structural basis of the alternating-access mechanism in a bile acid transporter. *Nature* **505**, 569-573 (2014).
142. 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).
143. Suzuki, H. et al. Crystal Structure of a Claudin Provides Insight into the Architecture of Tight Junctions. *Science* **344**, 304-307 (2014).
144. Li, D. et al. Crystal structure of the integral membrane diacylglycerol kinase. *Nature* **497**, 521-+ (2013).
145. Efremov, R.G. & Sazanov, L.A. Structure of Escherichia coli OmpF porin from lipidic mesophase. *Journal of Structural Biology* **178**, 311-318 (2012).
146. Cherezov, V., Fersi, H. & Caffrey, M. Crystallization screens: Compatibility with the lipidic cubic phase for in meso crystallization of membrane proteins. *Biophysical Journal* **81**, 225-242 (2001).
147. Caffrey, M. A lipid's eye view of membrane protein crystallization in mesophases. *Current Opinion in Structural Biology* **10**, 486-497 (2000).

148. Cherezov, V., Clogston, J., Papiz, M.Z. & Caffrey, M. Room to move: Crystallizing membrane proteins in swollen lipidic mesophases. *Journal of Molecular Biology* **357**, 1605-1618 (2006).
149. Caffrey, M. in Annual Review of Biophysics, Vol. 38 29-51 (2009).
150. Cherezov, V. et al. In meso structure of the cobalamin transporter, BtuB, at 1.95 angstrom resolution. *Journal of Molecular Biology* **364**, 716-734 (2006).
151. Wadsten, P. et al. Lipidic sponge phase crystallization of membrane proteins. *Journal of Molecular Biology* **364**, 44-53 (2006).
152. Wohri, 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).
153. Wohri, A.B. et al. A lipidic-sponge phase screen for membrane protein crystallization. *Structure* **16**, 1003-1009 (2008).
154. Engstrom, S., Alfons, K., Rasmusson, M. & Ljusberg-Wahren, H. in Colloid Science of Lipids: New Paradigms for Self-Assembly in Science and Technology, Vol. 108. (eds. B. Lindman & B.W. Ninham) 93-98 (1998).
155. Qiu, H. & Caffrey, M. The phase diagram of the monoolein/water system: metastability and equilibrium aspects. *Biomaterials* **21**, 223-234 (2000).
156. Engstrom, S., Wadsten-Hindrichsen, P. & Hernius, B. Cubic, sponge, and lamellar phases in the glyceryl monooleyl ether-propylene glycol-water system. *Langmuir* **23**, 10020-10025 (2007).
157. Nollert, P. & Landau, E.M. Enzymic release of crystals from lipidic cubic phases. *Biochemical Society Transactions* **26**, 709-713 (1998).
158. Kabsch, W. XDS. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **66**, 125-132 (2010).
159. Bailey, S. The CCP4 suite - programs for protein crystallography. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **50**, 760-763 (1994).
160. Hendrickson, W.A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51-58 (1991).
161. Blow, D.M. & Rossmann, M.G. Single isomorphous replacement method. *Acta Crystallographica* **14**, 1195-& (1961).
162. Weeks, C.M., Detitta, G.T., Miller, R. & Hauptman, H.A. Applications of the minimal principle to peptide structures. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **49**, 179-181 (1993).
163. Miller, R. et al. On the application of the minimal principle to solve unknown structures. *Science* **259**, 1430-1433 (1993).
164. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **60**, 2126-2132 (2004).
165. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **53**, 240-255 (1997).
166. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **66**, 213-221 (2010).
167. Brunger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **54**, 905-921 (1998).
168. White, T.A. et al. CrystFEL: a software suite for snapshot serial crystallography. *Journal of Applied Crystallography* **45**, 335-341 (2012).
169. 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).

170. Duisenberg, A.J.M. Indexing in single-crystal diffractometry with an obstinate list of reflections. *Journal of Applied Crystallography* **25**, 92-96 (1992).

171. Leslie, A.G.W. The integration of macromolecular diffraction data. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **62**, 48-57 (2006).

172. Kirian, R.A. et al. Femtosecond protein nanocrystallography-data analysis methods. *Optics Express* **18**, 5713-5723 (2010).

173. Boutet, S. & Williams, G.J. The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). *New Journal of Physics* **12** (2010).

174. Hart, P. et al. The CSPAD megapixel x-ray camera at LCLS. *X-Ray Free-Electron Lasers: Beam Diagnostics, Beamline Instrumentation, and Applications* **8504** (2012).

175. Lee, J.H. & Ihee, H. Advantages of time-resolved difference X-ray solution scattering curves in analyzing solute molecular structure. *Structural Chemistry* **21**, 37-42 (2010).

176. McCammon, J.A., Gelin, B.R. & Karplus, M. Dynamics of folded proteins. *Nature* **267**, 585-590 (1977).

177. Karplus, M. & McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology* **9**, 646-652 (2002).

178. Wampler, R.D. et al. Selective Detection of Protein Crystals by Second Harmonic Microscopy. *Journal of the American Chemical Society* **130**, 14076-+ (2008).

179. Weierstall, U. Liquid sample delivery techniques for serial femtosecond crystallography. *Philosophical Transactions of the Royal Society B-Biological Sciences* **369** (2014).

180. Faubel, M., Schlemmer, S. & Toennies, J.P. A molecular-beam study of the evaporation of water from a liquid jet. *Zeitschrift Fur Physik D-Atoms Molecules and Clusters* **10**, 269-277 (1988).

181. Kunnus, K. et al. A setup for resonant inelastic soft x-ray scattering on liquids at free electron laser light sources. *Review of Scientific Instruments* **83** (2012).

182. Weierstall, U. et al. Droplet streams for serial crystallography of proteins. *Experiments in Fluids* **44**, 675-689 (2008).

183. Sierra, R.G. et al. Nanoflow electrospinning serial femtosecond crystallography. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **68**, 1584-1587 (2012).

184. Kern, J. et al. Room temperature femtosecond X-ray diffraction of photosystem II microcrystals. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 9721-9726 (2012).

185. Kern, J. et al. Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature. *Science* **340**, 491-495 (2013).

186. Weierstall, U. et al. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature Communications* **5** (2014).

187. Liu, W. et al. Serial Femtosecond Crystallography of G Protein-Coupled Receptors. *Science* **342**, 1521-1524 (2013).

188. DePonte, D.P. et al. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D-Applied Physics* **41** (2008).

189. Chapman, H.N. et al. Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73-U81 (2011).

190. Boutet, S. et al. High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* **337**, 362-364 (2012).

191. Redecke, L. et al. Natively Inhibited Trypanosoma brucei Cathepsin B Structure Determined by Using an X-ray Laser. *Science* **339**, 227-230 (2013).

192. Barends, T.R.M. et al. De novo protein crystal structure determination from X-ray free-electron laser data. *Nature* **505**, 244-+ (2014).

193. Barends, T.R.M. et al. Anomalous signal from S atoms in protein crystallographic data from an X-ray free-electron laser. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **69**, 838-842 (2013).
194. Bozek, J.D. AMO instrumentation for the LCLS X-ray FEL. *European Physical Journal-Special Topics* **169**, 129-132 (2009).
195. Strueder, L. et al. Large-format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources. *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment* **614**, 483-496 (2010).
196. McCoy, A.J. et al. Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).
197. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728-2733 (2007).
198. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **57**, 122-133 (2001).
199. 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).
200. Lovell, S.C. et al. Structure validation by C alpha geometry: phi,psi and C beta deviation. *Proteins-Structure Function and Genetics* **50**, 437-450 (2003).
201. Schneider, T.R. Objective comparison of protein structures: error-scaled difference distance matrices. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **56**, 714-721 (2000).
202. Deisenhofer, J., Epp, O., Sinning, I. & Michel, H. Crystallographic refinement at 2.3 Angstrom resolution and refined model of the photosynthetic reaction center from *Rhodopseudomonas viridis*. *Journal of Molecular Biology* **246**, 429-457 (1995).
203. Lindahl, E., Hess, B. & van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling* **7**, 306-317 (2001).
204. Duan, Y. et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry* **24**, 1999-2012 (2003).
205. Ceccarelli, M., Procacci, P. & Marchi, M. An ab initio force field for the cofactors of bacterial photosynthesis. *Journal of Computational Chemistry* **24**, 129-142 (2003).

ISBN 978-91-628-9235-7
Printed by Ale Tryckteam AB