Thesis for the degree of Doctor of Philosophy
in the Natural Sciences

Structural studies of the photosynthetic reaction center of *Blastochloris viridis* using XFEL and Synchrotron radiation

Analia Banacore

UNIVERSITY OF GOTHENBURG

Department of Chemistry and Molecular Biology

Gothenburg, 2025

Thesis for the degree of Doctor of Philosophy
in the Natural Sciences

Structural studies of the photosynthetic reaction center of *Blastochloris viridis* using XFEL and Synchrotron radiation

Analia Banacore

Cover: Artistic representation of elements involved in the completion of this work.

ISBN 978-91-8115-114-5 (Print)
ISBN 978-91-8115-115-2 (PDF)
Available online at http://hdl.handle.net

Department of Chemistry and Molecular Biology
Division of Biochemistry and Structural Biology
University of Gothenburg
SE-405 30, Göteborg, Sweden
Printed by Stema Specialtryck AB
Borås, Sweden, 2025

Abstract

Life on Earth as we know it has been possible because of photosynthesis, a process by which hydrocarbons and chemical energy are generated using the energy extracted from sunlight. As with any other biological process, there is a protein responsible and it is called the reaction center. It is found in various photosynthetic organisms, including bacteria, algae, and plants. A photosynthetic reaction center has a conserved structure and function, where light absorption triggers an electron transfer process leading to energy production. This thesis is focused on the bacterial reaction center from *Blastochloris viridis*, analogous to photosystem II in plants. This study is focused on characterization of protein structural changes upon light absorption.

For decades, X-ray crystallography has been a key method for determining protein atomic structures. X-ray free-electron lasers have facilitated new approaches, called serial femtosecond X-ray crystallography (SFX), which enable new experimental possibilities with smaller crystals. This approach requires new crystal preparation methods for specialized delivery systems. In this thesis, a technique for growing membrane protein micro-crystals using lipidic cubic phase crystallization and adding crystal seeds is presented. With this approach, a 2.3 Å resolution SFX structure of the reaction center was obtained. Detergent grown micro-crystals were used to validate a fixed target device, yielding a 3.3 Å structure from data collected using serial synchrotron X-ray crystallography. LCP grown micro-crystals were used in time resolved SFX studies, where structural movements of the co-factors in the protein on a sub-picosecond timescale after photon absorption were assessed. Further, time resolved SFX results show multi-photon absorption induces a thermally driven expansion of the protein structure surrounding the co-factors. These findings contribute to a sequence of studies that highlight the intricate interplay between light absorption and structural changes within the protein.

Sammanfattning

Livet på jorden, som vi känner det, har varit möjligt tack vare fotosyntesen, en process genom vilken kolväten och kemisk energi genereras med hjälp av den energi som extraheras från solljus. Som med alla andra biologiska processer finns det ett protein som är ansvarigt, och det kallas reaktionscentrum. Det finns i olika fotosyntetiska organismer, inklusive bakterier, alger och växter. Ett fotosyntetiskt reaktionscentrum har en bevarad struktur och funktion, där ljusabsorption triggar en elektronöverföringsprocess som leder till energiproduktion. Denna avhandling fokuserar på det bakteriella reaktionscentrumet från Blastochloris viridis, som är analogt med fotosystem II i växter. Denna studie fokuserar på karaktärisering av förändringar i proteinstrukturer vid ljusabsorption.

Under flera decennier har röntgenkristallografi varit en nyckelmetod för att bestämma proteiners atomstrukturer. Fri-elektronlasrar har underlättat nya metoder, kallade seriell femtosekunds röntgenkristallografi (SFX), som möjliggör nya experimentella möjligheter med mindre kristaller. Denna metod kräver nya kristallförberedelsemetoder för specialiserade leveranssystem. I denna avhandling presenteras en teknik för att växa membranproteinkristaller i mikroformat med hjälp av lipidisk kubisk fas (LCP) och sådda kristaller. Med denna metod erhölls en SFX-struktur av reaktionscentrumet med en upplösning på 2,3 Å. Detergentväxta mikrokristaller användes för att validera en anordning för fixering av provet, vilket gav en struktur med 3,3 Å upplösning vid seriell synkrotronsljus kristallografi. LCP-växta mikrokristaller användes i tidsupplösta SFX-studier, där strukturella rörelser av kofaktorerna i proteinet på en sub-pikosekundsk tidskala efter fotonabsorption bedömdes. Vidare visar tidsupplösta SFX-resultat att multipel-fotonabsorption inducerar en termiskt driven expansion av proteinstrukturen runt kofaktorerna. Dessa resultat belyser det intrikata samspelet mellan ljusabsorption och strukturella förändringar inom proteinet.

Acknowledgements

I would like to start by thanking my supervisor, **Richard Neutze** for the opportunity to take part in his group and particularly in this project. You literally changed my life and took me out of my "biology comfort zone" into the "unknown magic data processing realm". I still have volumes to learn but have pushed to improve and learn every step of the way. **Gisela Branden**, my co-supervisor, thank you for always having a word of advice and for being kind and approachable. **Allen M. Orville**, thank you for taking on the task of being my opponent, hope it results on a great experience and that you enjoy Sweden. **Gergely**, we have not talked a lot but you always were there for me and my questions, and every time with great suggestions so thank you. **Julia**, thank you for being so driven for science and such a fantastic human being. You are an inspiration, you always have a kind word and good advice, always willing to help. Special mention to **Swagata Ghosh**, my guide and friend during the PHD journey. You were a helping hand when I was literally lost and you were always around the lab, every day at any hour. I missed your company when you left. And your food. Even after moving to Japan we managed to meet and work and keep contact. You are one of the jewels I found in this journey. **Giorgia**, you were in the group before I started and will be there after I finish. I am so thankful for that! You have the best energy to be around, I love your optimism and your thirst for life and adventures! Emotionally, I totally agree but then my body would rather have a fika! Keep doing what you do how you do it, with love and passion! And extra thanks for your help proof-reading this work. **Lucija**, it was always a pleasure to talk to you, we usually ended up laughing out loud. How to forget our two first beamtimes together: XFEL and SLS. The first night shift with the weird animal sound in the middle of the night and Giorgia driving us around are two experiences I will never forget! The office was indeed quieter after you left. **Arpitha**, I am so glad we met even before being fellow PHDs. You are such a kind soul and have the best life advice, always so pragmatic and practical, down to earth. I have been very lucky to share this time with you. I was very sorry to miss your wedding, but am happy everything worked out and so many people could attend. I lived the experience through pictures! **Adams**, I am so very sorry for all the times I asked you the same thing. And all the times I assured you my line of code was not working... and it was missing a space... or it had an uppercase ... or a dot. I am eternally grateful for your patience and I hope at least I managed to make it up with food. **Greger**, the only swede left of the group, thank you for all the advice, the fika times and all my last minute corrections to this document. Thank you also for the lovely book you gave Sol, she loves it! **Bernadette** I liked you from the first day, even if you managed to unplug my computer and stop all the processing hahaha nice first day. You survived and I am sure you will do great in your PHD. **Yanyan** I love having you around in the office, a pity we could not manage to do my last beamtime together! I am sure it would have been fun! **Monika**, thank you for always making time to help, you made me feel supported regardless of the distance. I was always looking forward to see you when in MAXIV, I enjoyed every one of our coffee breaks, lunch times and dinners! **Leona**, thank you for making me feel at home in Lundberg, always nice to have good company for lunch or fika and having someone to talk! The moving resulted in us having far apart office space and even lab space, hope that resulted in increased productivity since I was not around to distract you. **Andrea**, you have successfully managed to relax a bit. You have been in perpetual stress since I met you, though the levels decreased when you finished the PHD. **Andreas**, I am glad I met you even before getting started in Lundberg, a pity we were not often working at the same time, I always enjoy your sense of humor! **Gabrielle**, I enjoyed very much that we took that course together, hopefully you remember way more of that than me... **Johan** You are a great lab assistant and I think any student will be lucky to have you as their teacher. **Emelie**, one of the ones that was pregnant almost at the same time... I cannot believe you already have two now. I am still struggling with number one! **Davide**, I was very glad to see you return to GU. I still remember when you trained me for my master thesis project, thank you for that and for always being kind. **Torbjörn, Valerio, Charles** Natrium managed to put our offices apart, I hoped I can know you better in this time to come but in case it does not happen, good luck in your PHD. **Filippo**, you probably had a lot of work related

to my sample in the lab, sorry for that! And thank you for the advice and always being willing to help. **Petra**, thank you for many things, but mainly for adopting my cat for 2 months and our lunch times, I always enjoy spending time with you, lets keep the tradition every now and then. **Majo**, I have the best memories of you, and every time we meet you have the best advice and it is like time has not passed. I truly hope you manage to get all you wish for, you deserve it. And I really hope some day we will meet for that beer. **Maja**, thank you for all your help, kind words and all the times we met. I really enjoy spending time with you and I love that when we meet we talk about everything there is! I also love all your pictures of your cats, I can never see enough. **Viktor**, I still remember when you brought your baby to Lundberg. . . she must be huge now! She was as chill as you. . . that is probably not the case anymore. . . **Florian**, It is great to still have you around in the lab every now and then, I will tell the stories of you lunch for years to come. **Jessica**, also pregnant at the same time! It was fantastic to have someone to talk about everything and it is so nice every time we meet. You are easygoing and relaxed, sweet and welcoming, one of a kind. And you have a beautiful family! My best wishes always for you guys! **Jonatan**, the laser expert. . . only you know where everything is in the laser lab. I still remember when you were directing us over the phone while sick. . . no rest for you. Hope my defense goes as smooth as yours! Although if it is a bit shorter I would not mind... **Emil** Always next to Jonatan. Your beer is great, you do know you have a career there should you get tired of chemistry! **Andrew** The lab will miss you dearly, especially during beerclub! You had the best stories! **Bozidar** It was always fun to talk to you, we still have time so that you can practice your spanish. **Taru** I honestly think it is a pity you are so far away, you have the best style and vibe and are so easy to talk to. **Ylber, Jens** The best bro-mance ever. A pity you now work apart but the story will live on. **Irena** Thank you for your stories and advice. And for always being willing to listen to the lab issues and suggesting solutions or alternatives. **Björn** take care of her, she is truly one of a kind. **Ulrika** You are great company in the lab and always willing to help. **Emma**, being in charge of the course together was nice, but sharing a dance floor was amazing. You rocked the dance-floor. **Catarina** I love your style, I might have said so but in case I didn't it needs to be stated. **Johanna**

We have not talked much but I admire you, congratulations on becoming a professor. **Yosh** Thank you for the cell strains and for always being willing to help and talk. It is great to have you around. **Weixiao** You are an asset wherever you go, I hope the NMR center acknowledges all your hard work and your willing to help. **Dimitra**, also a fellow pregnant at the same time, it is a pity we got to share office space for such a short time. You had the best advice and it was always nice to talk to you. **Doris**, amiga mia, we had the best of times every time we met. Even that day you came to the park with Sol and we could not talk even one word, we still managed with our faces. I hope we stay in touch forever. **Lidja**, I must say I always had fun with you and I enjoyed that you also like beer. It was incredibly handy to have a tall person in the lab, my height is not great when the goods are stored on top of shelves. **Per**, hope you are enjoying fatherhood. It is a tough hood. **Tinna**, also pregnant at the same time, not sure how we managed to coordinate this! **Ann**, I don't remember we had a time to talk one on one but I always enjoyed having you around. **Rob, Cecilia S., Rebecca, Elin, Leo, Cecilia W., Darius, Daniel, Amke**, our time was short but I carry the best memories. **Laras**, I might some day take you up on that trip, would love to visit you. **Damasus**, I was glad to find out we are neighbors as well as colleagues. **Ayaan**, you were my first student and rocked it! Had lots of fun working in the lab with you, wish you the best for your PHD. **Beatrice**, I have no clue what you are up to, sorry I had to leave in the middle of our project but I am pretty sure Andrea took good care of you and KcsA. **Bruno**, I am not sure how everything will ever work without you around. **Valida**, thank you for always being kind and making me practice my Swedish. **Forough** thank you for helping me always, particularly with looking for space for my huge flasks and the massive amounts of light bulbs! **Peter** thank your for always having the willing to help, even though you did not necessarily have the time. **Lars**, you are always there to the rescue! Can fix from a common PC to whatever equipment in the lab. I have never met anyone as handy! **Marie**, I am so glad I met you. Hope we keep in touch forever and ever. **Anne**, Thank you for having been a part of this journey and for taking me as a master student. Your classes are the best. **Kristina**, a very special thanks to you, who encouraged me to keep applying for positions. You were right. To everyone at former Lundberg,

it was a pleasure having the opportunity to share this time with you. You have enriched this journey with priceless memories.

I will also take the opportunity to thank people that has not been here with me physically but helped from a distance and even before my journey begun. One can say they even made it possible. **Lau**, always there, a pity we can no longer quietly talk when we meet. The time will come when we will be able to have our tea times again. And yes, I also miss you. **Mari**, I am over the moon that you moved so close to me. It is encouraging that you are so close, though I yet have to visit you. I promise I will. **Natus**, we always enjoyed to plan things and were overly optimistic of our capabilities during a finite time to make things happen. This resulted in both being swamped during these years, but we kept in touch. **Jacke**, thank you for coming to visit! And for always making the time to meet when I go to Uruguay. And for gathering the courage to present your work here. **Charly**, keep it up! You are close to finishing your PHD. **Chech**, **Lu**, even though we live in 3 different countries we keep in touch and updated in a daily basis. I love it. **Perro, Lop, Tau,** we need to celebrate this. And some other things. Though I am not sure when the three of us will manage to be in the same geographic position. **Maca**, sis, I can't wait to see you again. Hope we can manage that soon. **Moni**, my oldest friend. Hope we keep it up. To infinity and beyond. **Pato**, a vos te salteo el ingles sino no vas a entender nada. Gracias por todo, en especial por ser vos y por siempre estar ahi para todo. Por ultimo voy a agradecer a mi familia, particularmente a mis padres, que aunque no tengan idea de lo que estoy haciendo me apoyan incondicionalmente para que consiga lo que sea que me proponga. Y una mención doble y especial para mi mamá que se cruzó el mundo varias veces no solo para estar con su nieta sino para ayudarme a terminar el doctorado.

Publications

This thesis consists of the following research papers:

PAPER I: Petra Båth, **Analia Banacore**, Per Börjesson, Robert Bosman, Cecilia Wickstrand, Cecilia Safari, Robert Dods, Swagatha Ghosh, Peter Dahl, Giorgia Ortolani, Tinna Björg Ulfarsdottir, Greger Hammarin, a María-José García Bonete, Adams Vallejos, Lucija Ostojic, Petra Edlund, Johanna-Barbara Linse, Rebecka Andersson, Eriko Nango, Shigeki Owada, Rie Tanaka, Kensuke Tono, Yasumasa Joti, Osamu Nureki, Fangjia Luo, Daniel James, Karol Nass, Philip J. M. Johnson, Gregor Knopp, Dmitry Ozerov, Claudio Cirelli, Christopher Milne, So Iwata, Gisela Brändén and Richard Neutze.

"Lipidic cubic phase serial femtosecond crystallography structure of a photosynthetic reaction centre"
(2022) Acta Crystallographica
`https://dx.doi.org/10.1107/s2059798322004144`

PAPER II: **Manuscript**
Swagatha Ghosh, **Analia Banacore**, Per Börjesson, Monika Bjelcic, Arpitha Kabbinale, Padmini Nileshwar, Gabrielle Wellhander, Danielle D Sanctis, Shibom Basu, Julien Orlans, Leonard M. G. Chavas, Richard Neutze and Gisela Brändén
"A low barrier cost effective fixed-target platform for synchrotron based serial crystallography"

PAPER III: Manuscript

Analia Banacore, Petra Båth, Adams Vallejos, Giorgia Ortolani, Swagatha Ghosh, Greger Hammarin, Per Börjesson, Peter Dahl, Lucija Ostojic, Petra Edlund, Tinna Björg Ulfarsdottir, Daniel James, Karol Nass, Philip J. M. Johnson, Gregor Knopp, Dmitry Ozerov, Claudio Cirelli, Christopher Milne, Gisela Brändén, Richard Neutze.

"No observable Sub-picosecond structural impulse within a photosynthetic reaction centre"

PAPER IV: Manuscript

David Arnlund, Adams Vallejos, Petra Båth, **Analia Banacore**, Robert Dods, Joachim Kübel, Michał Maj, Despina Milathianaki, Kenneth Beyerlein, Peter Berntsen, Chelsie Conrad, Garret Nelson, Erik Malmerberg, Cecilia Wickstrand, Linda Johansson, Rajiv Harimoorthy, Amit Sharma, Chufeng Li, Stella Lisova, Garth Williams, Gerrit Groenhof, Sebastien Boutet, Daniel P. DePonte, Sebastian Westenhoff, Anton Barty, Gisela Brändén, Jan Davidsson, Richard Neutze.

"Multiphoton induced conformational changes within a photosynthetic reaction centre"

Related papers that I have co-authored but that are not included in this thesis:

PAPER VI: Swagatha Ghosh, Doris Zoric, Peter Dahl, Monika Bjelcic, Jonatan Johannesson, Emil Sandelin, Per Borjesson, Alexander Björling, **Analia Banacore**, Petra Edlund, Oskar Aurelius, Mirko Milas, Jie Nan, Anastasya Shilova, Ana Gonzalez, Uwe Mueller, Gisela Brändén and Richard Neutze.

"A simple goniometer-compatible flow cell for serial synchrotron X-ray crystallography"
(2023) Journal of applied crystallography
`https://doi.org/10.1107/S1600576723001036`

Contribution report

PAPER I: I grew and harvested the *Blastochloris viridis* cells, purified RC and prepared the crystals for the 2019 XFEL experiment. I attended the experimental session in Paul Scherrer Institute, participated in the data processing, image production and uploaded the structures to PDB.

PAPER II: I produced and harvested the *Blastochloris viridis* cells, purified RC and set up the crystallization for multiple synchrotron based experiments. I tested the platform multiple times and participated in the data processing and structure refinement. I wrote the first draft and was continuously involved in improvements to the manuscript and figure creation.

PAPER III: I cultivated and harvested the *Blastochloris viridis* cells, purified RC and prepared the crystals for the experiment, participated in the experimental session at the Paul Scherrer Institute, took part in the data processing, drew figures and contributed to the manuscript methods.

PAPER IV: I was involved in image generation and data processing.

Abbreviations

Here follows a list and short explanation of the different abbreviations used in this thesis.

ATP	**A**denosine **T**ri **P**hosphate
***Bl*vir**	*Blastochloris* ***vir****idis*
Bchl	**B**acterio **Chl**orophyll
Bphe	**B**acterio **phe**ophytin
CAD	**C**omputer **A**ided **D**esign
CMC	**C**rytical **M**icelle **C**oncentration
cryoEM	**cryo**genic **E**lectron **M**icroscopy
DDM	n-**D**odecyl-β-**D**-**M**altoside
DOPC	**D**i-**O**leoyl**P**hosphatidyl**C**holine
ESRF	**E**uropean **S**ynchrotron **R**aditation **F**acility
EPR	**E**lectric **P**aramagnetic **R**esonance
FNR	**F**erredoxin **N**ADP **R**eductase
FT-SSX	**F**ixed **T**arget **S**serial **S**ynchrotron **X**-ray crystallography
GPCR	**G**-**P**rotein **C**oupled **R**eceptors
H$^+$	Proton(s)
HPLC	**H**igh **P**ressure **L**iquid **C**hromatography
G3P	**G**lyceraldehyde-**3**-**P**hosphate
GDVN	**G**as **D**ynamic **V**irtual **N**ozzle
IE	**I**on **E**xchange chromatography
LCP	**L**ipidic **C**ubic **P**hase
LDAO	**L**auryl**D**imethyl**A**mine **O**xide
LH	**L**ight **H**arvesting
LSP	**L**ipidic **S**ponge **P**hase
NADP+	**N**icotinamide **A**denine **D**inucleotide **P**hosphate
OG	n-**O**ctyl-β-D-**G**lucopyranoside
RC	**R**eaction **C**enter
MAD	**M**ultiple wavelength **A**nomalous **D**iffraction
MR	**M**olecular **R**eplacement

***RC*vir**	***R**eaction **C**enter of Blastochloris **vir**idis*
PDB	**P**rotein **D**ata **B**ank
PQ	**P**lastoquinone
PQH2	**Q** Plastoquinole
PC	**P**lastocyanin **C**arrier
PSI	**P**hoto **S**ystem **I** (number one)
PSII	**P**hoto **S**ystem **II** (number two)
P680	Special **P**air of Chlorophyll that absorb at 680 nm
RuBisCo	**R**ibulose-1,5-**Bi**sphosphate **C**arboxylase/**O**xygenase
SAD	**S**ingle wavelength **A**nomalous **D**iffraction
SAXS	**S**mall **A**ngle **X**-ray **S**cattering
SDS	**S**odium **D**odecyl **S**ulphate
SEC	**S**ize **E**xclusion **C**hromatography
SVD	**S**inlge **V**alue **D**ecomposition
SX	**S**erial **X**-ray crystallography
TR-SX	**T**ime **R**esolved **S**erial **X**-ray crystallography
TR-SFX	**T**ime **R**esolved **S**erial **F**emtosecond **X**-ray crystallography

Contents

Chapter 1

Introduction

Life on Earth as we know it is possible due to proteins and photosynthesis. Proteins are the building blocks of life, while photosynthesis harvests light to produce the energy that acts as its fuel. These indispensable molecules are involved in all biological functions and chemical processes and are thus responsible for every living organism that exists.

1.1 Proteins and their importance

All living cells are made of proteins, which modulate activities by undergoing cascades and structural changes that allow them to function. They are the machinery by which the cell is interconnected, built, and differentiated. Research on proteins and how they function is of utmost importance in many fields. Understanding how these processes work can give insight into the cause of many diseases and contribute to developing strategies to correct them. It is because of this fact that proteins have been studied for decades and continue to be one of the pillars of research. Understanding what they are and how they work is the first step towards optimization of biological processes, to cure several diseases, and even for engineering of biological machines.

1.1.1 Structure of proteins

Proteins have four levels of organization: primary, secondary, tertiary, and quaternary, which are depicted in figure 1.1. Proteins are built by an ordered sequence of amino acids, known as the primary structure. Each sequence dictates the folding and thus the three-dimensional conformation of the protein. This, in turn, determines its biological function and interactions with itself and other molecules. Even a single amino acid substitution in the primary structure can significantly alter the shape, stability, and activity of a given protein [1, 2].

The overall amino acid content is another relevant aspect because amino acids have specific properties: they can be charged, polar or hydrophobic. Thus, the overall content gives specific characteristics to each protein. A protein with the same size and the same amino acid content is not necessarily the same protein; in fact, they can be incredibly different, depending on where these amino acids are placed.

The secondary structures in proteins refers to the local three dimensional shape and arrangement of the poly peptide backbone, without considering the side chain interactions. These are mainly α helices and β sheets. These are held together by hydrogen bonds and allow further packing of the protein domains. α helices are right-handed coiled structures while β sheets can be either flat or gently twisted and curved. They are the structural pillars for the three-dimensional tertiary structure because their arrangement and interactions determine the shape, stability, and function of the protein.

Tertiary structures refers to the way these secondary structural elements and their side chains are disposed in space, namely the 3D structure. The quaternary structure appears when multiple polypeptide chains come together and are interconnected. This allows the arrangement of different chains into a machinery that allows the correct functioning by connecting the 3D structures between one another in space.

Proteins can be studied by biologists in all these levels. Many techniques have been developed throughout the years with this aim. In this work some of them will be mentioned with special emphasis on serial X-ray crystallography.

1.1.2 Amino acids - the building blocks

Amino acids play a vital role in various biological functions. There are 20 standard amino acids, all of which exhibit a central carbon atom called the α carbon [3]. This carbon is bonded to an amino group (NH2), a

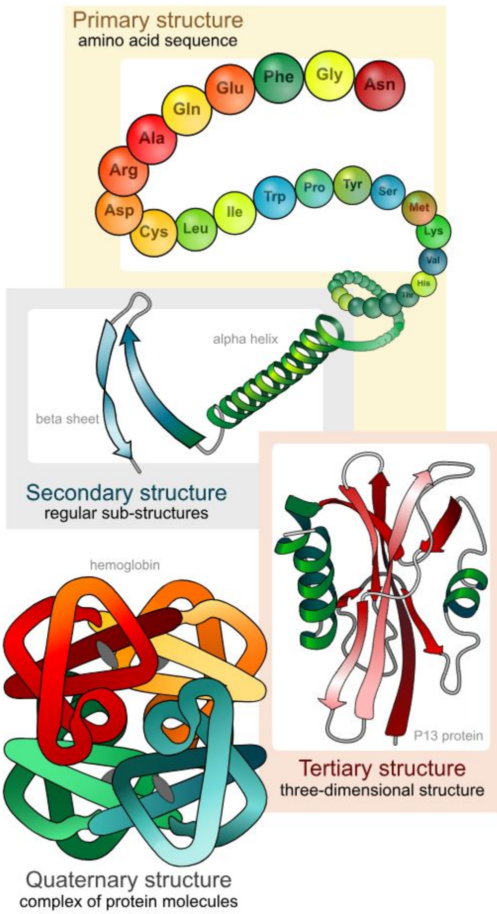


Figure 1.1: Protein structure organization. Different protein structural hierarchy are shown. In the primary structure, the amino acid sequence is shown with color coding. The secondary structure shows how the α helix and β sheets look like. The tertiary structure shows how these are organized in space and the quaternary structure shows how they interact in space to build a protein complex (credit: LadyofHats, public domain).

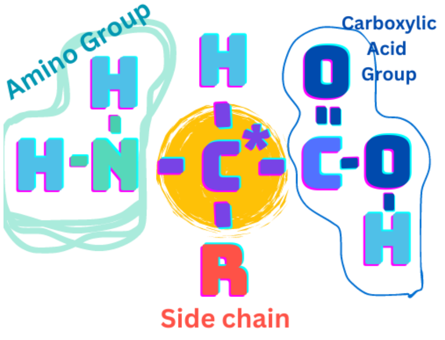


Figure 1.2: Amino acid structure. In the center and marked with yellow, the α carbon is represented, which is bound to an hydrogen and to the variable side chain (R) in red. On each side it contains the link to the amino group and the carboxylic acid group.

carboxyl group (COOH), a hydrogen atom, and a distinctive side chain, as depicted in figure 1.2. The side chain is commonly known as the R group and determines the unique properties and classification of each amino acid.

Amino acids can be categorized based on the characteristics of their side chains into Nonpolar, polar and charged groups. The nonpolar or hydrophobic amino acids have side chains that do not interact well with water. Examples include valine, leucine, and isoleucine. These are the ideal material to build membrane proteins. Polar or hydrophilic amino acids have side chains that form hydrogen bonds with water, such as serine, threonine, and asparagine. Charged amino acids can be either positively charged (basic) like lysine and arginine or negatively charged (acidic) like aspartic acid and glutamic acid.

Amino acids are linked together via peptide bonds, formed through a dehydration reaction between the amino group of one amino acid and the carboxyl group of another. Their unique side chains facilitate the formation of protein structures and influence the protein's stability, folding, and interactions with other molecules.

1.1.3 Cell membranes

Cell membranes are a component of all living cells and maintain cellular integrity, facilitating several biological processes. They act as a barrier that allows selective passage of ions and charged molecules, and separate the internal from the external environment [4]. They help maintain homeostasis by regulating the movement of ions, nutrients, and waste products in and out of the cell, making sure that the internal environment remains stable. They participate in cell communication through receptors that allow cells to communicate with each other and respond to external signals, such as hormones and neurotransmitters.

Membranes can have many functions and are located accordingly. The plasma membrane surrounds the entire cell and is present in all cells. Eukaryotic cells also contain internal membranes that form organelles, such as the endoplasmic reticulum, Golgi apparatus, mitochondria, chloroplasts and lysosomes. These membranes compartmentalize cellular functions and processes [4].

Cell membranes are described by the fluid mosaic model [5], which depicts the membrane as a dynamic and heterogeneous structure composed of a lipid bilayer with embedded proteins. This model accounts for the lateral mobility of lipids and proteins, allowing for various interactions and functions within the membrane [6]. Membranes have a specific lipid composition, including phospholipids, cholesterol, and glycolipids, which conditions membrane fluidity and the formation of specialized domains, such as lipid rafts. These domains are important for cellular signaling and protein interactions [7, 8].

1.1.4 Membrane proteins

Membrane proteins function as gatekeepers of the cell. Cellular membranes exhibit functional diversity and their role is defined and sustained by the composition and activity of their membrane proteins. About one third of proteins encoded in the genome are membrane proteins, but this number is not accurately reflected by the number of structures deposited in the Protein Data Bank (PDB) since the number is significantly lower [9].

There are mainly two types of membrane proteins. Integral membrane proteins are embedded within the lipid bilayer and can span the membrane. They often function as channels, transporters, or receptors, playing critical roles in cellular communication and transport. Alternatively, peripheral proteins are associated with the membrane at the level of their surface and can interact with integral proteins or lipid components. They are mainly involved in signaling pathways and maintaining the cell structural integrity.

More than half of all the commercially available drugs target membrane proteins because they amplify signals or participate in signaling cascades [10]. However, they are the hardest to work with and research on [11]. One of the main characteristics of membrane proteins is that they are amphiphilic, meaning that they have a hydrophobic part which allows them to be in a lipidic environment within the cell membrane while also having hydrophilic domains that allow their correct function in the aqueous environment [4]. This very convenient characteristic that allows placing and functioning in a very specific niche of the cell is rather troublesome for structural studies.

Membrane proteins need detergents in order to be extracted from the lipid bilayer, which they are part of. Since the environment provided by the detergent is not the exact same as the one the protein is in, this may lead to a loss of stability that can cause the protein to precipitate in solution. Therefore, studies of membrane proteins are difficult, but many strategies have been developed to overcome these hardships, some of which will be mentioned later in this thesis.

1.2 Photosynthesis

Photosynthesis is arguably the most important process underpinning life on earth. It directly or indirectly provides chemical energy for the biosphere by converting sunlight into chemical energy, while producing oxygen and organic compounds essential for the survival of plants and animals [12]. The term photosynthesis was coined in 1893 and was first introduced by C.R. Barnes as a way to refer to the process by which plants convert CO_2 and water to sugars and oxygen [13]. The most common organisms that come to mind when thinking of photosynthesis are plants, but these are not the only ones. In addition to plants, algae and certain bacteria are also able to convert light energy from the sun into chemical energy [14].

The basic equation of photosynthesis is often written as follows:

$$Carbon\, dioxide + water \xrightarrow{light} organic\, matter + oxygen \tag{1.1}$$

Strictly speaking, the process of photosynthesis can be divided into light dependent and light independent stages. The light dependent reaction, as the name suggests, relies on the excitation of the electrons by light causing a chain of electron transfer events through the membranes of the thylakoids of the chloroplasts. The light independent stage is most commonly known as the Calvin cycle and it produces glucose in plants by using the energy generated in the light dependent reaction.

1.2.1 Chloroplasts

A chloroplast is a specialized organelle found in the cells of plants and algae, where photosynthesis takes place. Its structure is depicted in figure 1.3. Julius Sachs and Andreas F. W. Schimper, who conceived the word "chloroplast" in 1885, observed the process of organelle division in different plant tissues and described it in their work, showing how the division of young chloroplasts in the cytoplasm of the cell takes place [15].

Konstantin Sergeevich Mereschkowski, a Russian biologist, proposed in 1905 the "symbiogenesis" theory, where he explained that these organelles are the result of endosymbiotic organisms close to cyanobacteria that have been integrated by its host in an evolutionary event [15, 16].

These organelles are enclosed by a double membrane, as shown in figure 1.3. The stroma is the space inside the chloroplast and it contains the enzymes needed for the Calvin cycle to take place. Inside the chloroplast there is an intricate system of structures bound to the membrane called thylakoids, organized into stacks called grana, where the light dependent reactions of photosynthesis take place.

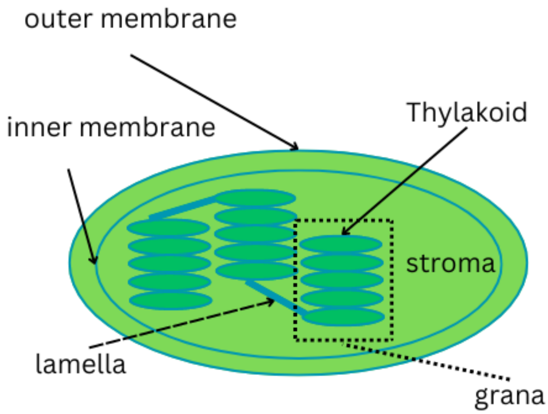


Figure 1.3: Chloroplast structure and its organization. Grana of the thylakoids are connected by lamella and located within the inner membrane, in the stroma.

Chloroplasts contain their own circular DNA and ribosomes. The DNA encodes some of the proteins needed for photosynthesis and other chloroplast functions. This fact supports Mereschkowski's theory and it was of course unknown at that time. His theory was born in the early 20th century, and the function and structure of DNA was discovered later in 1944, when the transfer of genetic information between bacteria was demonstrated by only using DNA [16].

1.2.2 Electron transfer events in Photosynthesis

Located within the membrane of the thylakoid inside a chloroplast, the electron transfer reactions in photosynthesis happens in presence of sunlight. This enables the increase in energy of an electron from a chlorophyll molecule, allowing the movement of the electrons across the thylakoid membrane. The electrons are obtained by the breakdown of water molecules, which serve in this case as weak electron donors. This in turn produces as a result the release of oxygen as a by-product. As the electrons are transported, a proton gradient is produced, which will be used by the ATP-synthase to produce ATP in the stroma. The final step is the conversion of $NADP^+$ to NADPH when high energy electrons are loaded along with H^+. This process will be described in detail.

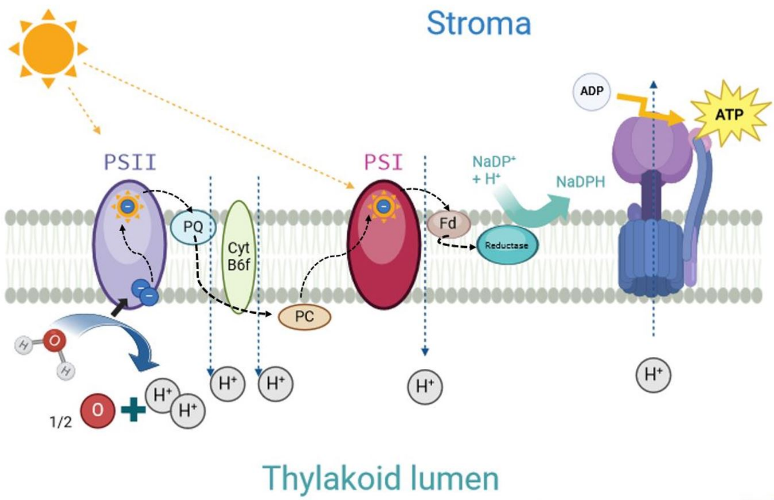


Figure 1.4: Electron transfer chain in photosynthesis The proteins involved in the electron transfer chain in photosynthesis; PSII and PSI, located in the thylacoid membrane are shown. Black dotted arrows represent the electron transfer pathway. Image created with BioRender.com.

The membrane protein involved in the first step is called photosystem II (PSII). This protein has a special pair of chlorophyll molecules that absorb light at 680 nm wavelength (P_{680}). The energy derived from sunlight is absorbed and causes an electron to be moved from one molecular orbital to another of higher energy. This causes the electron transfer to yet another chlorophyll molecule within PSII, then to a pheophytin molecule and ends up reducing a plastoquinone (PQ).

The process repeats itself, thus fully reducing the plastoquinone to plastoquinol (PQH_2) causing it to dissociate from PSII into the thylakoid membrane. The electrons from PQH_2 are then transferred to the cytochrome b_6f complex, which acts as a plastoquinol-plastocyanin oxidoreductase, thus reducing the mobile electron carrier plastocyanin (PC) [17].

Plastocyanin then migrates towards Photosystem I (PSI) and transfers electrons to the reaction center of PSI, which contains a special pair of chlorophyll A molecules called P_{700} [18]. However, these electrons have lost some energy during their journey through the electron transport chain from PSII. To regain the energy required for the journey ahead, the P_{700}^{+} chlorophyll pair absorbs light energy from the antenna pigments associated with PSI. This light energy excites the electrons within P_{700}^{+}, raising them to a higher energy level [19].

The electrons from PSI are used to reduce the final electron acceptor $NADP^{+}$ to NADPH, through ferredoxin-$NADP^{+}$ reductase (FNR) [20]. All these electron transfer events generate a proton gradient, since protons (H^{+}) are pumped from the stroma into the thylakoid lumen by the cytochrome b_6f complex and PSI. This proton gradient is used by ATP synthase to produce ATP, which, along with NADPH, provides the energy and reducing power for the light independent reactions in the Calvin cycle of photosynthesis to convert carbon dioxide into more complex molecules [17].

1.2.3 Calvin Cycle overview

The Calvin cycle comprises a series of biochemical reactions in the stroma of chloroplasts of photosynthetic organisms that allow the conversion of carbon dioxide into organic compounds (such as glucose) using the energy and reducing power generated during the light-dependent reactions [21].

It is the most widespread and well known carbon fixation pathway used by the majority of photosynthetic organisms such as cyanobacteria, algae, and plants. It utilizes the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) to fix carbon dioxide into organic compounds, primarily glyceraldehyde-3-phosphate (G3P) [22]. Other pathways, such as the reductive citric acid cycle and the 3-hydroxypropionate bicycle, use different enzymes and mechanisms for carbon fixation [21].

1.3 Purple photosynthetic bacteria

Purple bacteria are one of the most studied of the many species of bacteria. They have pigments called bacteriochlorophylls (Bchls) that give them a distinctive color, which can range from purple to red. These pigments absorb light in the infrared and visible spectrum, allowing them to thrive in low-light conditions [23].

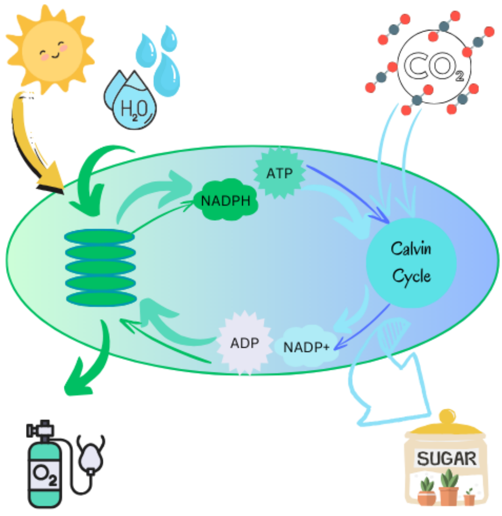


Figure 1.5: Process of photosynthesis within the chloroplast. The thylakoids contain the light dependent proteins and in the stroma, the proteins responsible for the Calvin cycle are confined.

Found in aquatic environments, such as ponds, lakes, and marine habitats, purple bacteria grow in the anoxic zones of stratified water bodies, where they form dense layers or blooms near the boundary between oxygenated and anoxic zones. These organisms require not only light but also absence of oxygen to grow phototrophically, because the synthesis of the pigments they require is repressed by oxygen [14].

Purple bacteria belong to the class Proteobacteria, which is one of the largest and most diverse groups of bacteria. They are believed to have diverged very early from other bacterial lineages and are believed to be among the earliest photosynthetic organisms on Earth [24].

Purple bacteria are primarily anoxygenic photosynthetic organisms, which means they do not produce oxygen as a byproduct of photosynthesis. Instead, they use light to convert carbon dioxide and other substrates into organic compounds, utilizing a variety of electron donors such as hydrogen sulfide or organic compounds [23, 25].

Their photosynthetic machinery, particularly the reaction centers (RCs), share remarkable similarities with the PSII complex found in oxygenic photosynthetic organisms like plants and cyanobacteria [26]. This suggests that the water oxidizing complex of PSII may have evolved from the simpler reaction centers of purple bacteria, thus playing a pivotal role in the oxygenation of the Earth's atmosphere during the Great Oxidation Event around 2.4 billion years ago [24].

The main light harvesting pigment of purple bacteria is called bacteriochlorophyll (BChl) and it absorbs light at significantly longer wavelengths in the near infrared spectra compared to chlorophyll from plants [25]. There are several kinds of BChl and the exact absorption wavelength depends on the specific type. The near infrared absorption of BChls is crucial for the photosynthetic process in anoxygenic photosynthetic bacteria, as it allows them to utilize the light energy efficiently in their natural environments, which lack visible light. This results in a clear advantage, not only does it allow purple bacteria to flourish in conditions that are unsuitable for most photosynthetic organisms, but also enables them to occupy ecological niches with minimal competition [27].

Carotenoids are also present and participate in the light harvesting process whilst providing the main pigmentation in the visible region, being responsible for absorbance spectrum between 450 and 550 nm, before the Soret bands of BChl (located in the near-UV between 300 and 400 nm) [25]. The Soret band is a strong absorption band found between 380-420 nm, in the ultraviolet-blue region, in the absorption spectrum of chlorophyll and related tetrapyrrole pigments like heme or B12 [28, 29]. It originates from a strong electronic transition in the porphyrin and determines the light-harvesting function of chlorophyll in photosynthesis [30].

There are mainly two groups of purple bacteria named sulfur and non sulfur purple bacteria, depending on their ability to metabolize sulfide compounds and supported by 16S rRNA analysis [14]. This molecular technique sequences the 16S ribosomal RNA gene, which is highly conserved among bacteria but contains variable regions that allow species identification and phylogenetic analysis.

1.3.1 *Blastochloris viridis*

Blastochloris viridis (Bl_{vir}) is one of many purple gram negative non-sulfur photosynthetic bacterium [31]. It was formerly known as *Rhodopseudomonas viridis* but was reclassified into a new genus based on 16S rRNA gene sequence analysis. The genus *Blastochloris* was made to include BChl b producing species of phototrophic bacteria [32]. It contains BChl b instead of BChl a, and has menaquinone in the Q_A quinone site, as opposed to other species of purple bacteria which have quinone [25]. It was first isolated from water bodies in Germany, near Freiburg im Breisgau [33].

Its photosynthetic system is one of the simplest known and it includes a light harvesting complex around the RC, along with soluble quinones and periplasmic cytochrome that transports the electrons between RC and bc_1 [32]. The photosynthetic membranes are tightly arranged in this organism in stacks that are apparently short in comparison to other microorganisms [34].

The main difference with photosynthesis relies in the electron donor, while PSII uses water as a weak electron donor, *Blastochloris viridis* reuses its electrons in a cyclic electron transfer process that will be explained in detail where cytochrome C_2 is the reductant or electron donor [35].

Bl_{vir} has a single circular chromosome of 3,723,225 bp, 81% of which codes for proteins. There is a photosynthesis gene cluster that contains 6 of the genes that (in addition to pufBALMC and puhA) code for the RC-LH1 core complex. Two other gene clusters code for bacteriochlorophyll and cytochrome bc_1. Cytochrome C_2, which is the electron carrier between cytochrome bc_1 and the photosynthetic reaction center, is encoded by a single gene [31].

When there is high light intensity in the growth conditions, the cytochrome C_2 and bc_1 complex ratio to the reaction centers increases, most likely to facilitate electron transport [34].

In the electron transfer process that occurs in Bl_{vir}, the RC transfers 2 electrons to a mobile ubiquinone-9. This molecule in turns needs to take up two protons from the cytoplasm to be able to transport these electrons. After that, in the form of ubiquinol (QH_2), it transfers the electrons to the cytochrome bc_1 complex and returns to its former state.

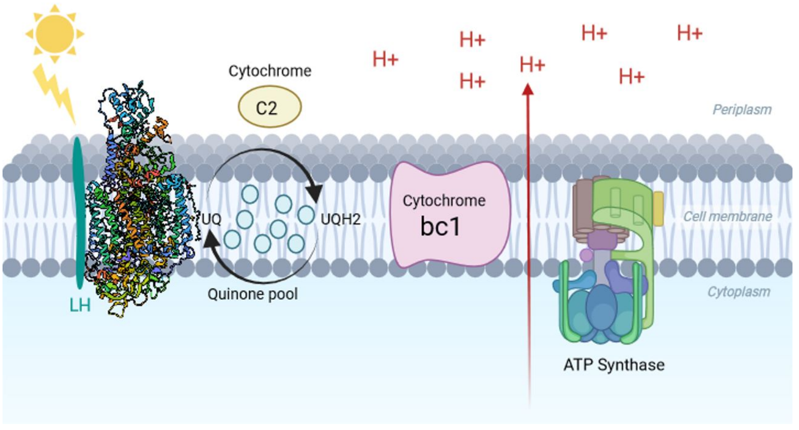


Figure 1.6: Proteins involved in the electron transfer events in *Blastochloris viridis* For the sake of clarity, only one LH complex is shown in the figure next to RC. 2 electrons are transfered from RC to UQ_9, these are passed on to Now, in the form of ubiquinol cytochrome *bc₁*. The cycle is closed by cytochrome C_2 that returns the electrons to RC. In this process, 4 Protons are transported across the thylakoids membrane towards the peri-plasmic region coupled with electron passage: two protons when the UQ molecule reduces and oxidizes, and two more when the electrons are transported through the cytochrome bc_1 protein. The result of this is the formation of a proton gradient that will then be used by the ATP synthase to generate ATP Figure created with BioRender.com

The protein responsible of closing the cycle by returning the electrons to the RC is cytochrome C_2 [36]. Protons are transported across the thylakoid membrane towards the periplasmic region coupled with electron passage: two protons when the UQ molecule reduces and oxidizes, and two more when the electrons are transported through the cytochrome bc_1 protein. The result of this is the formation of a proton gradient that will then be used by the ATP synthase to generate ATP [37]. The proteins involved in this process are shown in figure 1.6.

1.3.2 RC from Bl*vir*

Reaction centers are the main membrane protein complexes responsible for converting the energy of light to chemical energy. RC_{vir} particularly, has been studied in detail for years because its core structure of cofactors has high similarity with PSII from plants and cyanobacteria. Furthermore, two of it subunits (L and M) share some amino acid sequence homology with PSII from plants [38]. Furthermore, evolutionary analysis revealed that D1 and D2 proteins of PSII share a common ancestral origin with the L and M subunits of bacterial reaction centers [39].

Deisenhofer, Huber, and Michel determined the 3D structure of Bl_{vir}'s photosynthetic reaction center, the first integral membrane protein structure obtained through X-ray crystallography, and it was such a groundbreaking achievement that they were awarded the Nobel Prize in Chemistry in 1988 [40, 41].

RC_{vir} consists of four protein subunits named H, M, L, and C that contains 14 cofactors associated mainly with the chains L and M. The cofactors consist of 2 bacteriopheophytin (BPh-b), 4 bacteriochlorophyll (BCh-b), menaquinone-9, ubiquinone-9, 4 haem groups, a carotenoid molecule, and an iron ion [41, 42]. It is surrounded by a ring of 17 LH proteins that have the specific function of directing the photons to the RC to get started with the electron transfer chain [43, 44]. The L and M subunits contain 5 transmembrane α helices each, and the H subunit only one because it is located in the cytoplasmic surface [45].

The main difference when comparing to other RC's from other bacteria is the presence of an additional cytochrome subunit, termed C subunit [25, 46]. This is a tetraheme cytochrome subunit that functions as electron acceptor interacting with soluble electron carriers [47]. The C subunit contains the heme co-factors and is located towards the periplasm [14].

The electron transfer process takes place mainly in the L and M subunits, where the co-factors are held. Each of these subunits consist of five helices, termed A, B, C, D and E.

1.3.3 Electron transfer process in RC

The special pair of Bchl (P_{960}) is where the absorption of photons takes place in RC_{vir}, leading to an electron transfer event through the L subunit. The pathway that the electron transfer follows is referred to as the "A branch" or "active" branch [48]. From P_{960}, the electrons go to a second bacteriochlorophyll in the L branch ($BChl_L$). After that, a bacteriopheophytin located also in the L branch ($BPhe_L$) takes them, and then passes them on to a menaquinone-9 molecule that is termed Q_A because it is the primary quinone in RC_{vir}.

Usually, when an electron is transferred from the donor to the acceptor molecule, the result is a state where the donor is positively charged and the acceptor is negatively charged. When the electron returns to the donor molecule the system returns to its original state, and the charge separation is undone. Before this happens, during the charge separated state, the P_{960^+} gets an electron from the closest heme group in the C subunit [49, 50]. Immediately after, the electron in the pathway can be transferred from Q_A to Q_B (the mobile quinone) through the iron in the structure. The complete reduction of ubiquinone (Q_B) to quinol (Q_BH_2) involves the transfer of two electrons and is accompanied by the uptake of two protons and it happens in 2 steps [45, 51].

When the final acceptor of electrons is a quinone, the co-factors usually appear in an apparent two-fold rotational symmetry (C2 symmetry). This means that the structure can be rotated 180 degrees around a specific axis and appear essentially the same. Even though this is the case, the electron transfer events happen along the L subunit (the A branch) and never in the M subunit (the B branch) [52]. This happens due to structural differences because the L and M subunit might seem alike, but are truly not and thus are not considered homologues.

If assumed that both pathways could potentially be used for electron transfer events, one could argue that this has happened because evolutionary reasons, making the function of RC the most efficient if only one of these pathways is used [53]. In the hypothetical case where both quinone molecules were to simultaneously detach from their binding sites, there would be a risk that electrons could end up in an empty pocket (where the other quinone should have been). This would result in an electron being wasted, thus decreasing the efficiency of the process.

This notion is supported by the fact that electron transfer predominantly occurs through the A branch in many RCs [45, 48]. Thus, despite the near symmetrical arrangement of the co-factors, evolutionary pressures and functional optimizations have led to a preference for one branch over the other in many species.

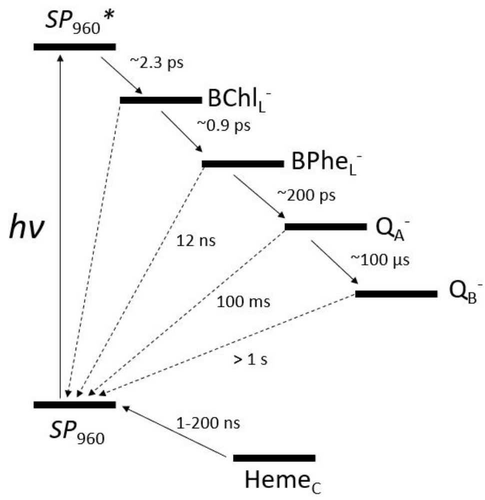


Figure 1.7: Approximate time-points of electron transfer within the RC*vir* for the first reduction of Q_B. The energy levels depicted are arbitrary, and the time-points might differ according to bibliography.

1.3.4 Kinetics of the electron transfer of RCvir

Transfer of the electron from P_{960} to $BPhe_L$ happens in 2.8 ps [54, 55]. In between, there is a short-lived intermediate of less than 1 ps when the electron passes through $BChl_L$ [56]. After the electron is in $BPhe_L$, it takes 200 ps to reach Q_A and the reduction of Q_B takes an extra 100 μs [45, 57]. Q_B gets fully reduced with a second electron transfer event. A schematic of this process is shown in figure 1.7.

1.4 Theories on protein dynamics

Chemical reactions have a wide range of time frames to happen as depicted in figure 1.8. Some reactions, for example the ones enzymes perform, happen mostly in the millisecond to second time scales [58].

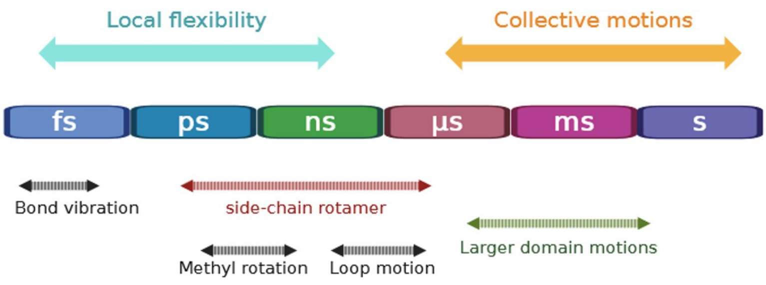


Figure 1.8: Timescales for reactions Graphic representation of dynamic processes in proteins and their time scales. Image taken and modified from [62] and created with BioRender.com.

Some others, like photosynthetic electron transfer reactions [59], photochemical reactions [60] or isomerization of molecules such as retinol [61] happen at ultra fast time scales.

1.4.1 Protein native states and their impact

There are multiple conformations that a protein can adopt at its native state [62]. This range of configurations possible is influenced by many factors such as thermal motion and specific interactions that can happen both within the protein residues, but also with the environment and neighboring molecules [63]. These are mostly systems considered ergodic, meaning it can explore its entire configurational space if given enough time. In a non-ergodic system, the time scale of the reaction is shorter than the time it takes to explore all possible states. The reaction rate depends on which configurations are accessible during the reaction, not on the possible range of states [64].

There is still a matter of whether or not these sub states and the transitions that happen affect or not biochemical reactions, whether these changes contribute only to the equilibrium or if they have an impact during the course of a reaction. Namely, if only the equilibrium free energy barrier is altered or the reactions energetics are also affected. In order to fully understand chemical kinetics it is needed to know if protein dynamics themselves can influence reaction rates beyond just altering static thermodynamic properties. Most established theories do not allow efficient control of reaction rates through the inherent movements of atoms within proteins. This possibility raises the question of whether a dynamic mechanism needs to be considered to explain how proteins modulate reaction rates. Given that protein dynamics contribute to reaction energetics, the inclusion of this concept will aid in the development of new models that tackle enzyme function, protein folding, and the design of molecules to modulate biochemical reactions. By integrating static and dynamic properties together in a model, a more comprehensive prediction of chemical kinetics in biological systems will be reached. This question was targeted by several studies performed in RCs given that their timescale falls in the range of picoseconds [59]. In order to understand how protein dynamics influence electron transfer reactions in the bacterial photosynthetic RC, molecular simulations were performed between Bphe and the primary quinone (Q_A) [65]. It was seen that protein and water dynamics restrict the available configurational space during electron transfer reactions, influencing the activation barrier in a time-dependent (non-equilibrium) manner. Thus, a shift from equilibrium-based to kinetic models is required to consider the configurations the system can take in a certain time frames, that might not be the total configurations possible due to time constraints. In this study, certain mutations were also shown to affect flexibility near the Bchl site, altering the dynamics and energetics of the reaction. Therefore, showing interactions between protein motions and water molecules modifies the energy barrier and ultimately influences the reaction outcome.

Marcus theory of electron transfer assumes equilibrium conditions for the reaction, which means that the system has enough time to explore its configurational space. The activation energy and reaction rates are based on thermodynamic parameters that are constant over time. It remains a relevant model for reactions that happen in conditions where nuclear relaxation is fast relative to the reaction time scale, but in the case of RCs where electron transfer processes happen on time scales as short as few picoseconds, non-ergodic kinetics need to be considered. In this case, the reorganization energy and activation energy are not static thermodynamic properties but are dependent on the nuclear rearrangement that is possible within the time of the reaction.

1.4.2 IR spectroscopic studies in RC

Since a central question in the field is whether sub-picosecond conformational changes occur in RCs during electron transfer, and how such dynamics align with Marcus theory, many spectroscopic studies have been performed. There are a few lines of reasoning and indirect evidence that point toward sub-picosecond conformational changes in RC proteins. Spectral and electronic state dynamic studies provide evidence of the temporal progression of the primary electron transfer reaction of RC [66]. The ultra-fast rise and decay of absorption features, as well as the appearance of long-lived states, support the hypothesis that the protein environment undergoes structural changes to facilitate efficient electron transfer and stabilization of charge-separated states.

Research combining quantum dynamic theories with quantum chemical calculations suggests that intramolecular vibrations contribute to sub-picosecond charge separation in the PSII reaction center [67]. These vibrations enhance the robustness of charge separation, indicating rapid structural adjustments within the protein. While several studies have provided evidence supporting sub-picosecond conformational changes RC proteins, some research suggests that large-scale conformational changes may not occur within such rapid timescales. A simulation study on the PSII reaction center indicated that the D1 and D2 proteins do not undergo significant large-scale conformational changes within sub-picosecond timescales [68]. This suggests that while ultra-fast structural motions may occur, substantial conformational rearrangements might require longer timescales.

Another study using steady-state absorption and fluorescence spectroscopy in RC from *Rhodobacter sphaeroides* to investigate the initial steps of charge separation in bacterial reaction centers within picoseconds [69] suggest that ultra-fast electronic dynamics can occur without the need for large-scale conformational rearrangements in the protein.

Electron paramagnetic resonance (EPR) studies on RC of *Rhodobacter sphaeroides* indicated that efficient electron transfer occurs without the necessity for ultra-fast conformational changes [70].

Ongoing research continues to explore these dynamics, aiming to provide a more comprehensive understanding of the fundamental processes that drives photosynthesis.

1.5 Scope of this Thesis

The focus of this thesis is the production of the RC protein from *Blastochloris viridis* and its subsequent crystallization for diverse studies related to the structural changes associated with its function. The main methods used were fixed target serial synchrotron crystallography (FT-SSX) and time resolved serial femtosecond X-ray crystallography (TR-SFX).

PAPER I presents new conditions used to obtain micro-crystals of *Blastochloris viridis* RC in lipidic cubic phase (LCP). These conditions use seeds from detergent grown crystals and LCP micro-crystals diffracted to 2.25 Å resolution. Ubiquinone is included within the crystallization environment and the electron density observed in the Q_B pocket is improved.

In **PAPER II**, RC*vir* was crystallized using detergent based protocols to validate a fixed target method to perform structural studies at synchrotron radiation sources. This system is conveniently designed for ease of use and compatible with most standard X-ray crystallography beam-lines.

In **PAPER III**, crystallization conditions from paper I were used in TR-SFX studies of RC*vir*. Specifically, to address the question of whether or not there are structural changes within the protein prior to the primary electron transfer reaction in the RC*vir*.

PAPER IV, utilizes both TR-SFX and TR-XSS to study RC*vir*'s response to multi-photon absorption processes. In particular, **PAPER IV** shows how a global ultrafast conformational change increases in amplitude as more photons are absorbed in RC*vir*.

Chapter 2

Methodology

2.1 Protein purification

Protein purification is essential for studying the structure, function, and interactions of proteins, as well as for various applications in biotechnology, pharmaceutical development, and research. It involves separating a specific protein from a complex mixture of other proteins and molecules, such as lipids, nucleic acids, and RNAs that are present in biological samples.

There are several techniques available for protein purification, each with its own advantages and limitations [9]. The choice of purification method depends on factors such as the properties of the target protein, the source of the sample, and the desired purity level. Additionally, the purification process can impact the stability and activity of the target protein, making it essential to optimize the purification conditions and assess the quality of the protein obtained [71]. In this chapter, an overview of the most commonly used techniques is provided.

2.1.1 Cell growth

The first step towards protein purification is cell growth because it determines the amount and quality of the protein to be produced. There are several aspects to consider when producing a certain protein. The very first one is the choice of expression system since it influences the protein folding, post translational modifications and expression levels [72]. The growth conditions can also impact the cell growth and protein yield, thus determining the best temperature, pH, aeration, and nutrient availability is a fundamental aspect [73].

Another element to consider is that the lipids in the host membranes of different expression systems might vary making it an important factor when the protein that is going to be expressed is indeed a membrane protein [74].

Another component to take into account is the induction strategy. The expression system chosen is the one that will determine what induction strategies we can use. The timing and concentration of the inducer can also condition what result we will end up having. The most common inducer when working with bacterial expression is Isopropyl β-D-1-thiogalactopyranoside more commonly known as IPTG, but in our case is the light exposure. This factor plays a vital role in maximizing protein production and keeping cellular stress and proteolytic degradation at a minimum [75]. Furthermore, when inducing a membrane protein, the fact that proper insertion into the membrane also needs to be considered [74]. A widely used tool to improve the protein production yield is genetic engineering, where modified organisms are used. These organisms have certain modifications that increase the chances of success for protein expression. Codon optimization is one of the strategies used as well as the implementation of a stronger promoter for protein expression [75], [73]. The addition of specific chaperone proteins needed for the correct folding of the protein of interest is another improvement that can be performed to the host cell [76].

In a small number of cases, the production from the native source yields large amounts of protein that enables purification without any modifications needed. After the cell with the protein content is obtained and harvested, it is stored at -80 °C for subsequent purification.

2.1.2 Solubilization and detergent

This step is required when the protein that we need to purify is a membrane protein. These proteins are integral part of the membrane and in order to extract them from their native environment, one must mimic it with detergents as closely as possible [77]. This is no easy task since there are a great number of detergents available and each one has specific properties. Furthermore, we might need to mix them in certain proportions to achieve the best result possible [9]. A screen is often needed to find what the ideal detergent is and what ratio to use [78]. There has been important developments in screening methods for membrane protein purification over the years [9]. There has also been research that contributed to improving the purification yield of membrane proteins and enhancing the specific recovery of membrane protein-lipid interactions by designing a range of detergent mixtures [79]. Rational design of detergents is an active field of research since there is a lack of understanding of what the role of the detergent structure is regarding protein purification yield [78].

The general structure of a detergent consists of a water soluble head group and a water insoluble tail group. The head group provides the detergent with its water solubility and charged characteristics and usually is an ionic or polar group. The tail group is usually a long, linear or branched hydrocarbon chain, which can be saturated or unsaturated and is responsible for the detergent's ability to interact with and solubilize non-polar substances, such as lipids and the hydrophobic residues of membrane proteins [79]. The balance between the hydrophilic head group and the hydrophobic tail group determines the overall properties and behavior of the detergent [80].

The concept of protein coverage by detergent micelles is crucial for maintaining membrane protein stability and functionality in solution. It has been proposed that decrease in lateral pressure may be one key component in protein destabilization due to partial removal of lipid molecules [81]. Molecular simulation studies proposed that the increase in hydrophobic mismatch as well as the membrane curvature and protein structure are important factors that influence the behavior and stability of membrane proteins [82, 83]. Hydrophobic mismatch happens when there is a difference between the hydrophobic regions the membrane proteins and the thickness of the lipid bilayer which they are part of [84]. Novel detergents have been developed to improve the solubilization of membrane proteins while minimizing hydrophobic mismatch effects [79].

Detergents are referred to as harsh or mild, depending on its chemical structure, charge, and ability to disrupt or preserve the native structure and function of proteins. A detergent is considered harsh if it has a charged or ionic head group such as sodium dodecyl sulfate (SDS). Ionic detergents can strip away essential lipids and co-factors, disrupt electrostatic interactions, and denature proteins. It can also disrupt hydrophobic interactions that stabilize the protein's tertiary structure causing denaturalization. They can even dissociate oligomeric protein complexes by interfering with protein-protein interactions, destabilizing and inactivating multi subunit membrane protein complexes. These kind of detergents are extremely useful in some protein separation techniques but are not ideal for applications such as protein purification. On the other hand, a detergent is considered mild if its head group is non-ionic or zwitterionic. Zwitterions are molecules that possess both positive and negative charges but remain electrically neutral. Examples of these include n-dodecyl-β-D-maltoside (DDM) and lauryl dimethylamine N-oxide (LDAO). These are less denaturing and can form tight but non disruptive interactions with proteins [79].

A strategy is to use one detergent for purification and then exchange it during the purification process for another one, depending on the downstream needs for the sample.

Other properties that need to be taken into account when choosing a detergent is the aggregation number and critical micelle concentration (CMC) [9]. The aggregation number is the number of molecules of detergent contained in a micelle. The CMC refers to the minimum concentration needed for a micelle to form [79]. Detergents need to be in the micellar state in order to be able to disrupt the membrane structure and form the hydrophobic interactions that enable them to successfully extract the protein by forming complexes with the protein and the membrane [79, 85]. DDM is a widely used detergent for membrane protein purification, since it forms large micelles, which tend to be more effective for shielding the hydrophobic regions. Smaller micelles, like the ones n-Octyl-β-D-Glucopyranoside (OG) forms, might not cover fully all hydrophobic regions causing the protein to aggregate and precipitate [79]. Excessively large micelles can interfere with protein-protein interactions and may hinder certain experimental techniques, such as crystallization [77]. Mild detergents also tend to preserve protein-protein interactions in oligomeric complexes which is important for maintaining the structure and function of multi subunit membrane protein assemblies, such as RC.

Detergent hydrophobicity is also a factor to consider because hydrophobic detergents (like OG and NG) are better at solubilizing highly hydrophobic proteins, but can also destabilize them. These properties are important to be taken into account as well as what the downstream technique is going to be. For crystallization, smaller detergents are preferred because they tend to expose more regions of the protein which can be paramount in crystal lattice formation [85].

Last but not least, the conditions such as pH, temperature, and ionic strength, can also influence the behavior of the detergent and impact on the protein stability. Choosing the right detergent for successful membrane protein for purification along with the ideal conditions to use is of great importance, as it helps solubility and stability of the purified protein while aiding maintaining its native structure and function.

2.1.3 Chromatography

Once the protein has been solubilized, it needs to be separated it from other cell components. Chromatography is an analytical technique for separating molecules in a mixture and it is widely used for several applications such as drug formulation and quality control, pollutants analysis in the air, water, or soil and of course for protein purification and analysis.

The fundamentals of chromatography rely on the differential partitioning of compounds between two phases, a stationary phase and a mobile phase [86]. The stationary phase is fixed in place. The mobile phase, on the other hand, moves through or over the stationary phase, carrying the components of the mixture and facilitating the movement of the analytes. The molecules in the mixture move according to the properties they have thus exhibiting a specific retention factor. This is defined as the ratio of the distance traveled by the compound to the distance traveled by the solvent front and helps identifying compounds. It basically quantifies how far a substance travels relative to the solvent front during a chromatography [87].

There are different types of chromatography methods that are classified according to the mobile and stationary phases they use. Gas chromatography uses a gas as the mobile phase and is often used for volatile compounds. It employs a liquid stationary phase coated on a solid support [86]. High Performance Liquid Chromatography (HPLC) and Ultra-High-Performance Liquid Chromatography (UHPLC) are variants of liquid chromatography. They use a liquid mobile phase, like column chromatography but with increased pressure [88]. In the case of RC*vir* purification, a liquid chromatography system developed by Cytiva called ÄKTA system was used, that works in a lower-pressure range, making them more suitable for the separation and purification of delicate biological molecules [89].

2.1.4 Purification of RC*vir*

Once the protein is in solution, two different chromatographic techniques are used to isolate RC.

In the first place, an ion exchange (IE) chromatography is performed. This method relies on the charge of the protein to be isolated and the choice of stationary phase [90]. The stationary phase used for RC*vir* is POROS anion exchange resin 50 HQ. This resin has a size of 50 μm and is rigid. It is made of quaternary polyethyleneimine, a strong anion exchanger.

Anion exchange chromatography is commonly used to purify or separate biomolecules such as proteins, nucleic acids, and polysaccharides that carry a net negative charge at the working pH [88]. Anions in the sample bind to the positively charged groups on the stationary phase. The strength of binding depends on factors like charge density and size of the anions. Weakly bound anions elute first, while strongly bound anions require a higher ionic strength to be displaced. They can be eluted using a gradient of increasing ionic strength, pH, or polarity [87]. For RC*vir*, increasing concentration of NaCl was used.

After concentrating the fraction with RC, a second chromatographic separation is performed. In this case, a method called size exclusion chromatography (SEC) or gel filtration is used. It is a separation technique that separates molecules based on their size. The stationary phase consists of porous beads with a pore size designed to allow entry of smaller molecules while excluding larger ones [90]. As the sample passes through the column, the larger molecules that cannot enter the pores of the beads move faster through the column and elute first. Smaller molecules that can penetrate the pores have a longer path to travel and thus, elute later [87].

2.2 Crystallization

A pure sample is needed for crystallization to be performed. When X-ray diffraction studies are carried out on crystals of proteins, researchers can determine the structure of the protein forming the crystal at atomic resolution. This is a very useful tool for drug design, as understanding protein structures aids in the design of drugs that target specific proteins, particularly in the field of medicinal chemistry [91].

Although extremely useful, this technique is quite challenging and is the bottleneck for many structural studies. Proteins can be complex to crystallize, certain factors such as post-translational modifications and conformational flexibility can affect crystallization trials. Thus, finding the right conditions for crystallization often requires extensive screening and optimization, involving various precipitating agents, pH levels, and temperatures.

2.2.1 Theory - crystallization process

The principles behind protein crystallization rely on the fact that proteins when in a supersaturated solution overstep the maximum solubility they can reach. Thus, they are pulled out of the solution by being tightly packed and ordered together in a process known as nucleation [92]. When in optimal conditions, proteins begin to spontaneously order themselves in a crystal that will keep growing and getting ordered until they reach an equilibrium [93].

For this process to occur, the protein needs to be both pure and concentrated. In addition, the protein needs to be in the presence of a substance called precipitant that is used to reduce the solubility limit of the protein [92].

There is a diagram that depicts the crystallization process as a function of the protein and the precipitant concentration, which is shown in figure 2.1. This diagram needs to be determined experimentally for

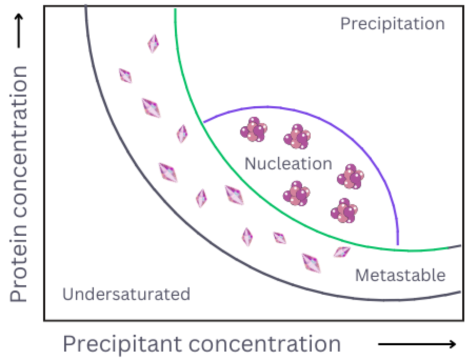


Figure 2.1: Phase diagram of crystallization process showing the effect of protein concentration vs. precipitant concentration. Water evaporates leading to increased precipitant and protein concentration, hopefully reaching the nucleation zone where molecules spontaneously start ordering themselves. Crystals will grow until an equilibrium is reached.

each protein because both the protein concentration and the precipitant concentration varies according to the properties of the sample [94].

The nucleation process involves the formation of small clusters called nuclei of protein molecules and can be spontaneous or induced by factors such as the presence of seed crystals [95].

As the molecules begin to get ordered together, the concentration of protein begins to decrease until an equilibrium is reached. When the metastable zone is reached, the crystal stops growing [94]. Conditions, such as temperature, pH, and ionic strength, can be optimized to control the size and quality of the crystals obtained [96]. There are other possibilities of events that can happen if the ideal concentrations are not reached. The most common outcome is that the protein precipitates owing to high salt and protein concentrations.

There are different approaches to obtain protein crystals and different strategies. One of the most straightforward methods is batch crystallization, where the protein solution is mixed with a precipitant and allowed to equilibrate [94]. The most commonly used method is vapor diffusion, and it is presented in detail later in this work.

The final goal of protein crystallization is to get the best data quality possible, and this is directly dependent on the quality of the protein crystals obtained. It is not uncommon to test different temperatures for crystal growth since the rate at which the molecules deposit themselves has a direct impact on crystal quality. Usually, the slower the process happens, the more time they have to get better ordered, tightly packed and the crystal diffraction has better data quality [96]. Seeding is another approach that might increase the crystal quality, where very small previously grown crystals are introduced to provide a nucleation site instead of waiting for the process to happen spontaneously [97].

Some proteins can be successfully crystallized, but in some cases some modifications are needed in order to make the protein prone to crystallization such as affinity tag removal, protein engineering or addition of certain molecules during the process of crystallization. Some affinity tags used for purification purposes can block crystal contacts making the process harder. Flexible regions such as terminal tails are also not favorable if present for crystallization making the process more likely if removed [98]. There are certain molecules such as specific antibodies or ligands that might be added during the crystallization process that aid in the process working as chaperons for the correct molecule disposition to form crystal structures [99–101]. This process is usually called co-crystallization.

Molecules in protein crystals are held together by a network of non-covalent interactions. These interactions occur not only between the protein molecules, but might also involve the surrounding solvent molecules [102]. The nature of these non-covalent bonds, which include hydrogen bonds, van der Waals forces, and electrostatic interactions, contribute to the distinctive properties of protein crystals resulting in a more delicate and less rigid crystal structure. As a consequence, protein crystals tend to be susceptible to physical stress and environmental changes. This makes protein crystals sometimes challenging to handle and manipulate, requiring careful consideration in their preparation and analysis.

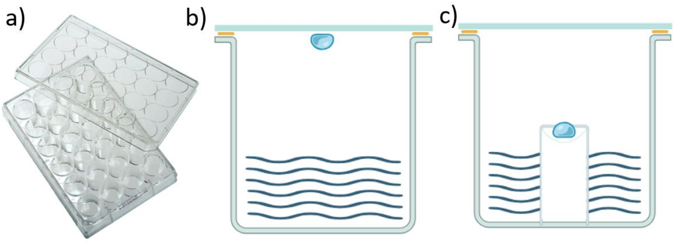


Figure 2.2: Vapor diffusion setup for crystallization Picture of a 24 well plate a) for hanging drop vapor diffusion crystallography setup. b) This schematic shows how the experimental schematic setup for the hanging drop method. The protein is hanging from the cover lid and sealed with grease, allowing vapor diffusion from the protein solution towards the precipitant. c) Schematics for the sitting drop method, where the protein solution is sitting on a well, sealed with a cover lid and grease. The protein solution is surrounded by the precipitant allowing the vapor diffusion to take place. Image created with BioRender.com

2.2.2 Vapor diffusion

This crystallization method involves placing a drop of protein solution in a sealed environment with a higher concentration of precipitant. The water molecules will then migrate from the protein drop to the reservoir with the precipitant leading to supersaturation of the protein solution in a very slow and controlled way. Variations of vapor diffusion are micro batch, sitting drop and hanging drop. They use very small volumes and can be automated for high throughput screening.

Hanging drop technique

The setup for the hanging drop method is a small drop of a protein solution suspended from a glass cover slip, which is then inverted over a reservoir containing the precipitant solution, and it is shown in figure 2.2b. This creates a vapor diffusion environment, where water evaporates from the drop into the reservoir, leading to an increase in the concentration of the protein solution and promoting crystallization.

It is important to control the conditions such as temperature and concentration gradients, which can significantly influence the crystal growth.

This method can be adapted for various types of proteins and conditions, making it a flexible choice for crystallization experiments and one of the most widely used setups. There are some considerations that need to be taken into account when performing this method. First, the size of the drop should be a volume that allows sufficient protein content, but small enough to minimize evaporation, mantain a stable environment and not fall by its own weight [96]. The initial concentration of the protein solution should be optimized to achieve supersaturation, which is essential for crystal formation as it can be seen in the phase diagram shown in figure 2.1. The choice of precipitant and its concentration must be carefully considered, as it directly impact the crystallization process. Last but not least, maintaining a consistent temperature is vital for reproducible assays and it also influences the stability during the crystallization process.

Sitting drop method

The hanging drop and sitting drop methods differ in their setup. In the sitting drop method, a drop of protein solution is placed directly on the surface of a well in a crystallization plate, with the precipitant solution in the well below it. The drop sits on the bottom of the well, with the precipitant solution below it, as shown in figure 2.2c. In this method vapor diffusion occurs, but instead, the drop is in direct contact with the well surface. Thus, it can be easier to set up and might reduce the risk of contamination and can facilitate removal of crystals.

Improvements from the 24 to the 96-well plate design have expanded its capacity, making the process more efficient and cost-effective, which has led to its widespread application [103].

2.2.3 Lipidic cubic phase - LCP

Lipidic cubic phase (LCP) crystallization is a technique used for crystallization of membrane proteins [85]. The LCP environment, formed by specific lipids such as monoolein, which is shown in figure 2.3, mimics the natural membrane by forming a bicontinuous cubic phase and providing a favorable matrix for the solubility, stabilization, and crystallization of these proteins [85, 104].

This approach was reported by Landau et al, where they mixed lipids, water and membrane proteins to form a complex 3D array with a system of aqueous channels that intercommunicate and provide a matrix for crystallization [105]. The hydrophobic proteins are able to diffuse within the bi-layer created by the lipids in the mixture while the hydrophilic components can move throughout the aqueous channel system. This method has evolved over decades and has continuously been improved and researched. It has been shown that even though the size and shape of the protein crystals are directly modulated by the cubic phase transport properties, it does not affect their space group [104]. Innovations in crystallization methods, including the use of sponge phases as liquid analogs of cubic phases, have also emerged, enabling the successful crystallization of challenging membrane proteins like bacteriorhodopsin and G protein coupled receptors (GPCRs) [106].

A new molecule of monoacylglycerol called 7.10 MAG for LCP crystallization has been produced, which is shorter than monoolein by one methylene in chain length, and with a cis carbon-carbon bond that is located two carbons closer to the glycerol group. This molecule is shown in in figure 2.3 under a monoolein molecule. These structural changes alter the phase behavior and the microstructure of the mesophases while exhibiting compatibility with other lipids, such as cholesterol and dioleoylphosphatidylcholine (DOPC), over a wide range of concentrations and temperatures, allowing to optimize crystallization conditions and potentially improving the quality and size of protein crystals [107].

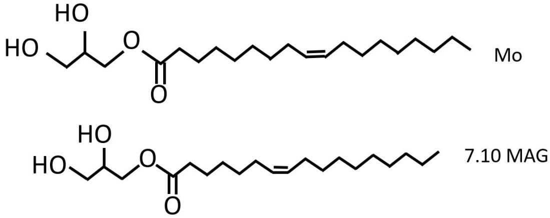


Figure 2.3: Comparison of Monoolein (at the top) and 7.10 MAG (at the bottom) are shown.

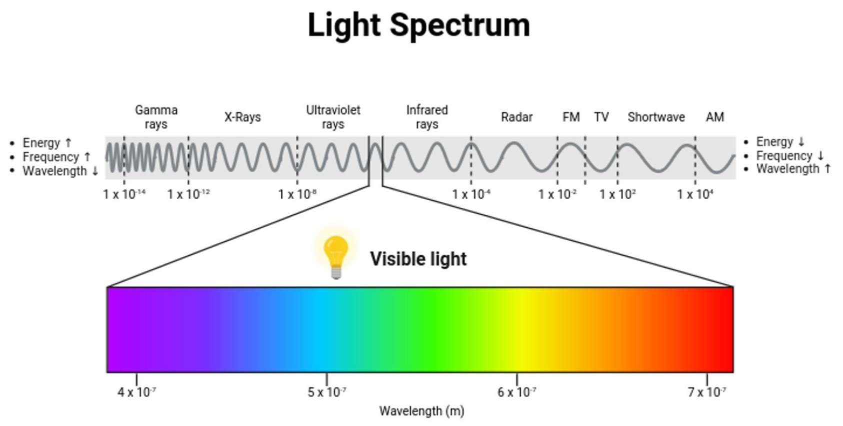


Figure 2.4: Light Spectrum Light spectrum from Gamma rays, X-rays, ultraviolet rays, visible light, infrared and FM, TV, shortwaves and AM waves. As can be seen in the figure, from left to right the energy decreases, the frequency decreases and the wavelength increases. Image generated with Biorender.com.

Recent studies have explored the influence of lipidic molecular structure on the mesophase properties and crystallogenesis, revealing that modifications, such as incorporating anionic lipids, can enhance water pore sizes, facilitating protein inclusion and diffusion [104]. This research allows high-resolution structural insights that can aid in drug discovery and membrane protein functionality studies.

2.3 X-ray diffraction

X-rays are a type of electromagnetic radiation with a wavelength shorter than ultraviolet light, but longer than gamma rays, represented in figure 2.4 in comparison with other radiation types. They were discovered accidentally by Wilhelm Conrad Röntgen, a German physicist in 1895, and the name was given because the radiation was (at that moment) of an unknown nature [108].

Nowadays X-rays are routinely used for many purposes and society as a whole has relied on them for decades. One of the most important applications for everyday life is medical imaging and diagnosis. X-ray imaging for radiography is used to visualize bones, organs, and other internal structures for diagnostic purposes such as detecting fractures, lung diseases, and tumors [109]. Computed Tomography (CT) scans use X-rays to create detailed cross-sectional images of the body, providing information and enabling medical conditions to be diagnosed [110]. Mammography uses low-energy X-rays to examine breast tissues for the early detection of breast cancer [111]. Another routine use for X-rays is non-destructive testing and inspection in industries to inspect materials and components for defects, cracks, or flaws without causing damage, ensuring the safety and integrity of structures, machinery, and equipment [112]. Security screening such as the ones used at airports, borders, and other security checkpoints also relies on X-rays to inspect the content and detect dangerous or unauthorized goods [113].

X-rays have wavelengths in the range of 0.1 to 100 Å, which matches the interatomic distances in solids, particularly in crystals, where these distances typically range from 1 to 2 Å (0.1 to 0.2 nm) [114]. This correspondence makes X-rays ideal for probing atomic structures through techniques like X-ray diffraction, because for resolving atomic scale structures, the wavelength of the radiation used must be on the same scale as the structure that is going to be studied [102].

In our particular case, we are interested in the diffraction caused by X-rays when a crystal is exposed. Due to the periodic arrangement of the atoms in the crystal, X-ray beams are scattered by the electrons surrounding the atoms in the crystal lattice. This scattering phenomenon results in constructive and destructive interference patterns, which are recorded as diffraction spots on a detector.

2.3.1 Theory

The key principle governing X-ray diffraction is Bragg's law, which describes the conditions for constructive interference of the scattered X-rays. It was observed by father and son in 1913 that, by exposing crystalline materials to X-rays, a pattern was produced that was a result of the reflection of the beam in different planes according to the crystal structure [115].

William Lawrence Bragg, 23 years old at the time, formulated the law that now bears their name. This work earned them the Nobel Prize in Physics in 1915, which made William the youngest Nobel laureate in physics up to date [116].

Bragg's law reads as follows:

$$n\lambda = 2dsin\theta \tag{2.1}$$

Where

- n: is an integer that corresponds to the order of diffraction.

- λ: is the wavelength of the X-rays.

- d: is the interplanar spacing between the planes of atoms in the crystal.

- θ: is the angle between the incident X-ray beam and the atomic planes, most commonly known as scattering angle.

For constructive interference to occur, the path difference between X-rays scattered from adjacent planes must be an integer multiple of the X-ray wavelength. This condition is satisfied when the angle of incidence (θ) and the interplanar spacing (d) follow the relationship described by Bragg's law [117] and it is shown in figure 2.5.

The molecules within a crystal are disposed in space with a certain order. The unit cell is the smallest repeating unit of the crystal structure, and it contains the full symmetry of the crystal lattice. When repeated in three dimensions, it forms the entire crystal structure. The corners of the unit cell corresponds to the lattice points in the crystal lattice.

The unit cell contains all the symmetry elements of the crystal lattice. The arrangement of atoms within the unit cell determines the symmetry of the crystal. The dimensions of the unit cell, such as the length (*a, b, c*) and the angles (α, β, γ) define the lattice parameters of the crystal structure.

The asymmetric unit is the smallest unique part of a crystal structure that, when subjected to the symmetry operations of the space group, generates the complete unit cell. It represents the fundamental building

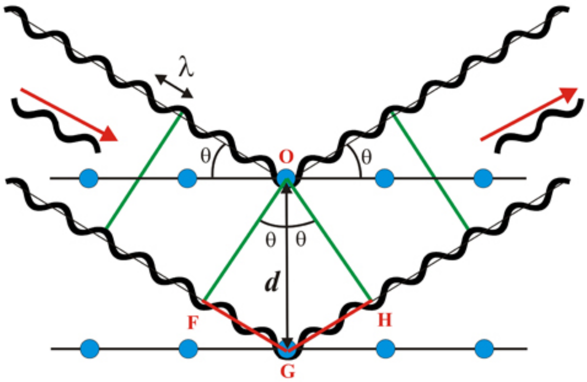


Figure 2.5: Bragg's Law diagram A representation of the atoms of a crystal (depicted in blue) and the atomic planes (d) is shown. The incident X-ray beam is depicted in black. All the X-ray waves are in phase with each other. θ is the scattering angle and it is the same as the incident angle θ. λ is the wavelength of the X-rays. Image from wikipedia commonsby *xmartin*, licensed under CC BY-SA 4.0. Source: Wikimedia Commons.

block of the crystal's repeating pattern. By applying the symmetry operators specific to the crystal's space group, the asymmetric unit can be transformed to produce the entire unit cell, which is the basic repeating unit of the crystal lattice. So, within the unit cell the asymmetric unit is the portion that contains all unique structural information. This simplifies the description and analysis of crystal structures by focusing on the essential structural elements that are not repeated.

There are 14 distinct types of crystal lattices in crystallography, known as Bravais lattices, which are categorized into seven crystal systems based on their symmetry. These can be cubic, tetragonal, orthorhombic hexagonal, rhombohedral/trigonal, monoclinic and triclinic and they are shown in figure 2.6 along with their characteristic features.

The 14 Bravais lattices are defined by four types of unit cells which can be primitive (P), body centered (I), face centered (F), and end centered (C) and are shown in figure 2.7. As it can be observed in figure 2.6, not all crystal systems can support all types of unit cells. In the primitive unit cell, the lattice points are located at the corners. It can be found in all

Figure 2.6: Table of crystal lattices This table shows the seven crystal systems and the 14 types of Bravais lattices.

Crystal class	Characteristic	Type of unit cell (Bravais Lattices)
CUBIC	a=b=c A=β=γ=90	P, I, F
TETRAGONAL	a=b≠c A=β=γ=90	P, I
ORTHORHOMBIC	a≠b≠c A=β=γ=90	P, I, F, C
HEXAGONAL	a=b≠c A=β=90 γ=120	P
MONOCLINIC	a≠b≠c A=β=90 γ≠120	P, C
TRICLINIC	a≠b≠c A≠β≠γ≠90	P
TRIGONAL	A=b=c A=β=γ≠90	P

seven crystal classes. The body centered has an additional lattice point at the center of the unit cell and can be found in cubic, tetragonal, and orthorhombic classes. The face centered has lattice points at the center of each face and is found only in the cubic and orthorhombic crystal classes. The end centered has lattice points at the center of two opposite faces in addition, and it is mostly found in monoclinic and orthorhombic systems.

Different centering types can lead to more efficient packing of atoms in the crystal structure. The centering affects the symmetry operations possible within the unit cell. Crystal lattices exhibit various symmetry operations, including rotation, reflection, and inversion. Different centering types result in characteristic systematic absences in X-ray diffraction patterns, which help in narrowing down possible space groups for a given crystal structure.

A system was developed to describe planes and directions in crystal lattices called Miller indices. They provide a standardized notation for identifying specific crystal planes and directions within a crystal structure. Miller indices are a set of three integers, typically denoted as (h, k, l),

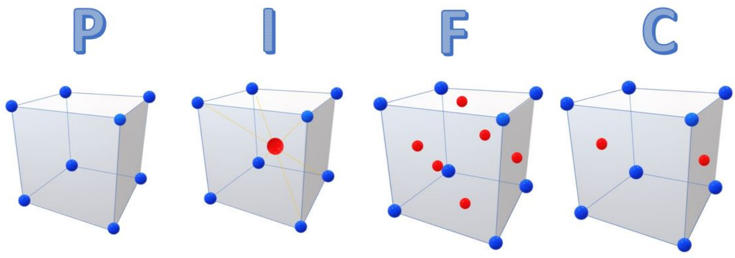


Figure 2.7: Unit cells There are four types of unit cells, and these are represented as follows: P (primitive),which allows atoms only at the corners. Unit cell type I (body centered)displays atoms at the corners and also at the center of the unit cell. Type F (face centered) has atoms placed at the corners and as well as at the centers of the faces. Last but not least, C (end centered) exhibits atoms at the corners and at the centers of two opposite faces.

that describe the orientation of a plane in a crystal lattice relative to the unit cell axes. The diffraction from the Miller index (*h, k, l*) forms a Bragg peak, and its intensity correlates with the electron density in the crystal planes associated with that particular Miller index.

Reflections (Bragg peaks) have a specific position that depends on the crystal lattice, namely the cell dimension and the space group that the crystal exhibits. This information, along with the intensity allows the calculation of the unit cell [118]. The diffraction pattern is a representation of how the arrangement of atoms in the crystal interacts with the incoming X-rays. The pattern obtained after diffraction is a representation of the reciprocal lattice. The reciprocal lattice has the same Laue symmetry as the real crystal's lattice, thus by knowing the reciprocal lattice and applying Fourier transformation, the real lattice can be calculated.

The relation between real and reciprocal lattices is explained by the Ewald's sphere, which is a geometrical representation that helps to understand the conditions under which diffraction occurs and it is shown in figure 2.8. Diffraction happens when the reciprocal lattice points intersect the Ewald sphere, leading to the constructive interference of scattered waves that forms the observed diffraction pattern. In order to solve a structure, it is ideal to obtain the maximum amount of reflection data

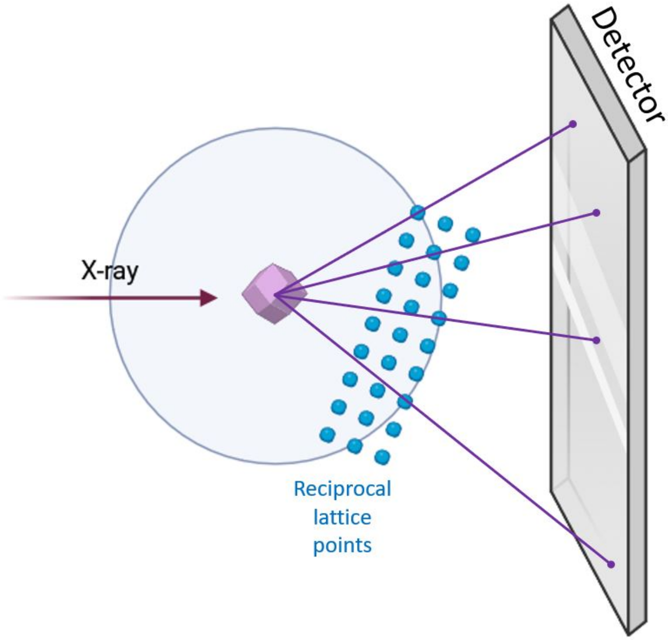


Figure 2.8: Ewald sphere representation For diffraction to happen, a reciprocal lattice point must lie on the surface of the Ewald sphere. This means that, for a given wavelength, the diffraction condition is met when the angle between the incident and diffracted waves satisfies the diffraction condition given by Bragg's Law.

from a crystal. For this to happen, the reciprocal space must be thoroughly sampled. This is done by rotating the Ewald sphere around its origin, thereby positioning as many reciprocal lattice points as possible in a condition where Bragg's law happens. Practically, this is the reason why the crystal is rotated for data collection.

There is an equation called the structure factor equation (F_{hkl}) that accounts for both the amplitude and phase of scattered X-rays from the atoms in the crystal. It holds information about the arrangement of atoms within the unit cell, their positions, and the number of electrons and it reads as follows:

$$F_{(hkl)} = \sum_{j=1}^{N} f_j \, e^{2\pi i(hx_j+ky_j+lz_j)}$$

The terms contained within the equation are:

- $F_{(hkl)}$: is the structure factor for the set of Miller indices hkl.

- N: corresponds to the number of atoms present in the unit cell.

- f_j: is called atomic scattering factor of the j-th atom, which describes how strongly the atom scatters the X-rays and is related to the number of electrons.

- (x_j, y_j, z_j): these are the fractional coordinates of the j-th atom within the unit cell, namely their positions.

- hkl: the Miller indices that define the planes in the crystal lattice.

- $e^{2\pi i(hx_j+ky_j+lz_j)}$: this is the phase factor, and it accounts for the relative positions of the atoms with respect to the crystal lattice planes. This term describes the phase shift of the scattered waves based on the atomic positions.

The structure factor is a sum of the contribution of each one of the atoms in the unit cell of the crystal. The atomic scattering factor f_j is summed up for each atom, and the exponential term for the atoms position accounts for the phase of the scattered wave depending on the position of the atom in question.

The amplitude of the structure factor is determined by the intensity of the diffraction spot. The phase of the structure factor that is encoded in the complex exponential is not directly measured in diffraction experiments. Since it is an essential piece of data to be able to construct the electron density of the crystal, this is commonly known as the "phase problem" in crystallography. There are several methods available to tackle this problem, which are explained in the data processing section.

2.3.2 Classic X-ray diffraction crystallography

Crystallography consists of determining the precise arrangement of atoms by studying the diffraction patterns produced when a beam of particles interacts with a crystal sample. Crystals are solid materials with a highly ordered microscopic structure, where atoms, molecules, or ions are arranged in a repeating 3D pattern [114]. The ordered nature of the protein crystal makes the scattered electrons from the atoms in the molecule interfere constructively and yield a diffraction pattern according to Bragg's law 2.1.

Crystals exhibit symmetry operations, such as rotation, reflection, inversion or translation and the combination results in 230 unique space groups, which describe all possible crystal symmetries [114]. In the case of protein molecules, and because of their chiral nature, there is only 65 possible space groups out of these 230 existing groups [118]. Only rotation, translation and rototranslation can be applied.

The diffraction pattern recorded on the detector contains information about the crystal structure, including the unit cell dimensions, symmetry, and the positions of atoms within the unit cell. By analyzing the intensities and positions of the diffraction spots or rings the 3D arrangement of atoms in the crystal lattice can be determined. In classic crystallography setups, also called cryo-crystallography, one single crystal is contained inside a loop and is rotated while exposed to the beam to collect a complete dataset. A complete dataset includes all unique reflections as determined by the symmetry of the space group of the crystal, which are mathematically equivalent or related to each other. Thus, only one reflection from each equivalence class is needed, rather than measuring every single reflection that appears in the diffraction pattern. In crystallography, the triclinic crystal system represents the lowest symmetry. In this system, symmetry is minimal, so nearly every reflection in the diffraction pattern is unique. Consequently, the total number of reflections that must be collected to solve the structure is significantly higher. In contrast, the cubic crystal system has the highest symmetry. In cubic crystals, symmetry relationships often make reflections in one part of the diffraction pattern equivalent to those in other parts. This equivalence reduces the number of unique reflections that need to be collected, simplifying the data collection process.

Even though there are some constraints associated with the use of crystallography, it has remained one of the most commonly used techniques in structural biology for many years. The first one is that the crystals should get to a certain size. This is because larger crystals produce stronger diffraction, a weak signal from a small crystal can result in poor-quality diffraction data and lower resolution. Furthermore, bigger crystals tolerate radiation damage better, while smaller crystals usually cannot tolerate a complete data collection. The second requirement is that the crystal needs to be flash frozen in liquid nitrogen for transport and data collection. This is to reduce the rate of radiation damage, though unfortunately it is not possible to do this completely [119]. During data collection in synchrotrons, a stream of cold nitrogen gas is constantly applied to keep radiation damage to a minimum. Last but not least, it is critical to include a cryo-protectant to the crystallization setup before the flash freezing process because the water molecules within the crystal might destroy it.

2.3.3 Serial X-ray crystallography

Serial X-ray crystallography (SX) involves collecting X-ray diffraction data from a large number of micro crystals in a serial fashion, rather than from a single large crystal as in traditional crystallography [120].

This technique offers certain advantages, such as the possibility to study proteins at room temperature, which helps avoiding potential artifacts derived from the process of cryo-cooling [121]. Furthermore, this approach opens the door to study protein movements through crystallography at room temperature avoiding the potential bias caused by the cooling process [122]. Since it requires much smaller crystals, it broadens the range of bio-molecules that can be studied. This allows structural studies of challenging proteins, for instance membrane proteins and protein complexes. These proteins are harder usually to crystallize or have crystal size constraints.

In SX, X-ray diffraction data is collected from each crystal at a given time, thus one image is obtained per crystal. The diffraction data from thousands of microcrystals randomly oriented are combined to form a full

3D model of the molecule being studied. In order for this to be possible, crystals are delivered into the X-ray beam in a continuous stream, such as using a liquid jet, or another specialized delivery method. X-ray free electron lasers (XFELs) allow to use extremely short, intense X-ray pulses that makes possible data collection is recorded before significant radiation damage occurs [123], which laid the theoretical foundation for their use in structural biology. Although initially it was an exclusive technique for XFELs, serial synchrotron crystallography (SSX) is a standard practice nowadays opening the bottleneck to millisecond time resolved experiments to many more users [124]. This is owed to technical upgrades where Synchrotrons can nowadays produce brighter and more focused X-ray beams, with beamlines being designed exclusively for this purpose and the introduction of fast improved detectors capable of capturing diffraction patterns at high speeds [125].

2.3.4 Time resolved serial X-ray crystallography

Time-resolved serial X-ray crystallography (TR-SX) is a powerful technique that allows researchers to study the structural dynamics of macromolecules at room temperature. It was introduced with the theoretical concept of "diffraction before destruction" [123], and was supported by the first time-resolved study published on nano-crystals of PSI at LCLS in the year 2011 where a 8.7 Å resolution was achieved [126].

TR-SX utilizes short X-ray pulses to capture molecular movies of proteins undergoing structural changes during a reaction. The experiment involves triggering a reaction in protein micro-crystals and measuring the structural changes at specific time delays after reaction initiation. These experiments are often referred to as pump-probe techniques, and they involve two steps. The first step is the pump pulse, where the sample is excited using a stimulus such as a laser pulse. The pump gives rise to a conformational change or a reaction in the protein molecules of the crystals. The second step is the probe pulse, where a second pulse, an X-ray in this case, generates the data collection. Intermediate states and transition dynamics can be captured by varying the time delays between the pump and the probe pulses.

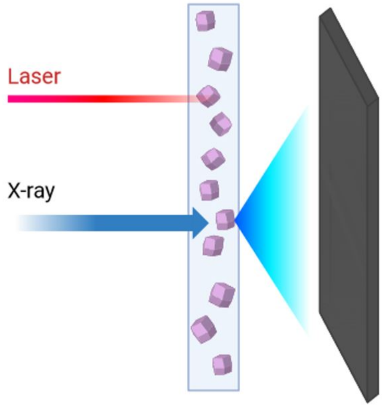


Figure 2.9: Experiment setup for TR-SX Graphic representation of the stream of crystals and the laser (pump, depicted in red) with a time delay from the X-ray pulse (probe, depicted in blue).These two appear apart to illustrate the time delay needed between them.

TR-SX resulted a great tool to visualize the structural basis of biological processes in real time, providing detailed insights into the mechanisms of protein conformational changes and enzyme catalysis mechanisms [120].

Figure 2.9 shows how an experimental setup looks like for this technique. As it can be seen in the image, the protein being studied needs to be triggered with a specific measurable time delay. Thus, a precise alignment of the pump and probe pulses is of utmost importance to get accurate time-resolved data. Ensuring the crystal in the X-ray and pump region overlap at the exact moment of excitation is challenging. Another important consideration, is that the crystals must be uniform within a size range to ensure consistent excitation and diffraction.

The trigger depends directly on the properties of the system being studied. If the protein needs a specific substrate to function or redox agent or cofactor, that can be exposed to the crystal and will diffuse and trigger the reaction [127]. In this work, light was used to trigger the protein, since it is a photoactive protein. There are other less common ways to use this technique and trigger the structural changes needed. Short bursts of strong electric fields applied to protein crystals, combined with time-lapse X-ray imaging, can be used to see how proteins change shape in space and time to explore how proteins work through their movements [128]. Work using solvent-based temperature jumps with TR-SX has been performed that showed widespread atomic vibrations in enzymes that evolved into localized structural changes that, compared to the inhibited protein, gave important information on the inhibition mechanism of the enzyme [129]. While the time resolution is lower in synchrotrons compared to XFELS, typically nanoseconds to milliseconds, TR-SSX allows studying slower processes like ligand binding, allosteric transitions, and enzimatic reactions.

2.4 Data collection and experiments

2.4.1 Synchrotrons

A synchrotron is a type of particle accelerator that makes electrons obtain extremely high energies through the application of electromagnetic fields. Synchrotrons are used as a powerful source of radiation, which is highly intense and has several properties. It can be fine-tuned from infrared to hard X-rays. Hard X-rays have shorter wavelengths and higher penetrating power (less than 1 nm). Soft X-rays have longer wavelengths (typically 1-10 nanometers) and are better suited for studying electronic structures, chemical environments, and magnetic properties.

This light source has a pulsed nature that allows measurements with pulses that can be as short as few picoseconds in the case of ESRF. It also allows polarization dependent studies.

These properties make synchrotrons tools for a range of applications, such as protein crystallography, small angle X-ray scattering, materials science, condensed matter physics, chemistry, medical imaging, and more [120].

The most important components of a synchrotron include an electron gun to shoot electrons into a linear accelerator, where they are accelerated to relatively low energies. Then there is usually a booster ring, where the electrons are injected into to be further accelerated to higher energies. After this, there is the storage ring, into which the high energy electrons are injected. It consists of a large circular vacuum chamber with bending magnets and other magnetic devices that keep the electrons circulating at nearly the speed of light. As the high energy electrons circulate in the storage ring and are forced to change direction by bending magnets, they emit extremely bright synchrotron radiation tangentially to the ring. This radiation is then channeled and optimized for various experimental techniques and applications using specialized beamlines and instrumentation [130].

These facilities are spread world wide. There are more than 50 synchrotron sources up to date, which are rather unevenly distributed. The number of synchrotrons continues to grow as new facilities are built and existing ones are upgraded, making this technology available to more people.

2.4.2 XFELS

The development of XFELs began with the Linear Coherent Light Source (LCLS) in 2009 in USA [131], followed by advanced facilities such as SPring-8 Angstrom Compact Free Electron Laser (SACLA) in 2011 in Japan [132], the European XFEL in 2017 in Germany [133] and SwissFEL in 2018 in Switzerland [134]. An X-ray Free Electron Laser (XFEL) is a type of linear particle accelerator that generates extremely intense and ultra-short pulses of X-ray radiation. XFEL radiation is extremely bright, billions of times brighter than synchrotron X-ray sources. It can reach ultra short pulses, shorter than 10 femtoseconds with approximately 10^{12} photons per pulse, at rates of several MHz, providing unmatched temporal resolution. The X-ray wavelengths can be tuned to be hard or soft. Unlike synchrotrons, XFELs use linear accelerators and require long facilities to achieve the necessary electron energy, an example being the European XFEL with 3.4 km length [131]. Another of the properties is high spatial coherence, which means that the X-rays emitted from different points within the source can interfere constructively or destructively over long propagation distances enabling advanced imaging techniques.

These properties make XFELs useful for structural biology and time resolved studies, especially for time resolved studies of ultrafast processes like photochemical reactions, protein conformational changes,chemical bond formation, and enzyme catalysis mechanisms on the femtosecond to picosecond timescales using pump-probe techniques.

XFELs are composed in the first place by an electron gun, where electrons are generated and accelerated using a linear accelerator. Then there is an undulator, where the high-energy electron beam is passed through. Here, relativistic electron beams experience microbunching due to interactions with emitted X-rays, aligning electrons into small packets that emit coherent radiation and spaced by approximately the wavelength of the emitted X-rays. This process enhances the beam intensity exponentially. The undulator is a periodic magnetic structure where the oscillating magnetic fields cause the electrons to undergo a "wiggling" motion, emitting X-ray radiation in the process. The emitted X-rays are amplified through self-amplified spontaneous emission (SASE), producing intense, coherent beams directly used for experiments. The interaction between the electron beam and the emitted X-rays causes the X-ray intensity to grow exponentially along the length of the undulator, creating an intense beam.

2.4.3 Sample delivery systems

Both TR-SFX and TR-SSX require specialized sample delivery techniques. Here we will briefly go through the main approaches available. There are some other approaches that will not be covered thoroughly, such as tape-drive systems (crystals are deposited onto a thin, moving tape) or droplet-based microfluidics.

Sample injectors

When using Liquid Jet Injection, microcrystals are suspended in a liquid, which is typically a buffer, and injected into the X-ray beam using a fine capillary, often aided by gas pressure. The continuous flow allows for high-speed data collection with minimal interruption, and it is ideal for TR-SFX, where rapid time-resolved data is needed. The drawback is that it requires a massive amount of crystals, and only works for systems where the crystal environment can remain stable in a liquid suspension. It is very common in XFELs to use Gas Dynamic Virtual Nozzle (GDVN), where crystals flow within a very thin liquid stream into the beam. GDVN was the approach used in 2011 for the proof of concept publication for TR-SFX in PSI [126].

Fixed target

Fixed target methods involve placing microcrystals onto solid support, such as chips or meshes, which are then scanned by the X-ray beam. These substrates are mounted on specialized sample holders and positioned in the X-ray path. This approach minimizes sample consumption, since the crystals are stationary rather than flowing continuously [135]. It also offers flexibility in using different carrier media, including viscous materials that may not be suitable for injection methods [136]. Additionally, it simplifies the experimental setup [137].

However, drawbacks include the potential for increased background scattering and the need for precise scanning mechanisms [137]. Physical damage to the samples during handling, such as crystal breakage or accidental loss during chip preparation, is a possibility [138], making it preferable to avoid this step when possible.

To address these challenges, various fixed target techniques have been developed, which will be discussed in chapter 4.

2.4.4 Data processing overview

When we have successfully purified and crystallized our protein, chose the method and X-ray facility along with the ideal sample delivery method and successfully collected our data, it is time for data processing.

In this section, an overview of the process is presented. The raw data consists of a series of images that contain the diffraction of our crystal samples and are stored as packs of images in .h5 files. These images are in fact the images from the detector, and crystallographic softwares such as MOSFLM [139] and XDS [140] are used. These programs aim to sort the data by using the Miller indices (*hkl*) according to the parameters of the unit cell of the crystal and taking into account the intensities. This process can be divided into stages, first of which is called autoindexing and consists of determining what the unit cell is and the crystal lattice orientation, which might indicate the Laue group. This is performed by identifying diffraction spots within diffraction patterns and then a mathematical model is created to describe the lattice parameters and relate them to the observed diffraction spots. A procedure called post-refinement [141] is used to further refine the initial estimate of the unit cell parameters obtained during autoindexing and to evaluate the degree of imperfections or misalignment within the crystal, known as mosaicity [139]. The last step integrates the images. Here, the positions of the Bragg reflections in each image are predicted, and an estimate is obtained of the intensity and the uncertainty of each reflection. The algorithms used for these purposes might vary between available software, such as XDS (X-ray Detector Software) [140], MOSFLM [139], DIALS (Diffraction Integration for Advanced Light Sources) [142] and CrystFEL [141, 143]. Most beamlines have automated software which gives quick information on the sample and automatically goes through the data processing described.

Data processing step by step

The first step is to reduce the data size by extracting the useful information from the raw data files. First, the unit cell dimensions need to be determined as long as the crystal symmetry, in a process called indexing. Afterwards, the intensity of each diffraction spot needs to be integrated in the images. This step is called integration. Scaling is then applied, by where corrections for variations in X-ray intensity, crystal absorption, and other factors are done. The last step is called merging, and it consists of combining multiple observations of the same reflection. It Averages symmetry-related reflections and combines partial reflections from different images.

During merging, data redundancy and multiplicity become significant. Multiplicity refers to the number of times a reflection is measured, and higher multiplicity improves the signal-to-noise ratio (S/N). This enhances the statistical reliability of intensity measurements, facilitates accurate atomic position fitting, strengthens the refinement process, and increases confidence in the final structural model.

Data quality assessment

Data quality indicators are calculated, such as R-merge and I/σ. These are statistical measures used to assess the quality and reliability of X-ray diffraction data obtained. R-merge is a metric that evaluates the consistency of intensity measurements for the same reflection across multiple observations. Roughly, it measures how reproducible the intensity measurements are. A lower R-merge indicates better data quality, indicating more consistency in the intensities observed. For a dataset to be considered good, the R-merge value should be lower than 10%. Sigma (σ) refers to the precision of the measurement, and its the estimated standard deviation or uncertainty in the measurement of diffraction intensities. A smaller sigma indicates intensity measurements with higher precision, which generally translates to higher-quality data. This is because the signal to noise ratio for a reflection is often expressed as:

$$\mathsf{S/N} = \frac{I}{\sigma(I)}$$

- I is the measured intensity.
- $\sigma(I)$ is the standard deviation of the intensity.

In data processing, sigma values are used to filter out weak or unreliable reflections. Typically, data with $I/\sigma(I) > 2$ are considered reliable for structural analysis.

The processed data is evaluated by checking specific parameters. From the data analysis obtained, the resolution limits are determined as well as the completeness of the dataset. Signal-to-noise ratios are calculated and evaluated, along with R-factors (measures of data consistency). In cases of weak reflections or high-resolution data, where intensities may be lower, the high multiplicity achieved in serial crystallography provides robust statistical support, enabling better S/N and more precise structural models. The serial crystallography metrics that need to be considered are the half-Set Correlation Coefficient most commonly known as $CC_{1/2}$ and R_{split}. $CC_{1/2}$ is a measure of the Pearson correlation coefficient calculated between two randomly divided halves of a given dataset and it is a statistical measure to quantify how one variable changes with respect to another. It provides a measure of the consistency and quality of the measured diffraction intensities, particularly at high resolutions, and is used as a reliability indicator for merging data. A high $CC_{1/2}$ value is close to 1 and indicates a strong agreement between the two half datasets, implying reliable data. It can be used to determine the resolution cutoff, as it reflects the strength of the signal relative to noise in the data. R_{split} is a measure of how reproducble is a measurement in two independent halves of a dataset. It whather the two halves agree with each other, and is useful for assessing the quality of integrated intensities in data with high multiplicity or varying quality. It is expressed as a percentage, with lower values indicating better agreement.

Phase determination

X-ray detectors only record intensities while the phases remain unknown. This is widely known in crystallography as the "phase problem". There are different methods that can be used to get the phase information. The most common one is called molecular replacement, and it is based on the fact that proteins with similar sequences often have similar 3D structures. It uses the known structure of a related protein to estimate the initial phases for the unknown structure. First, a structurally similar protein is chosen, typically based on sequence homology or functional similarity. The orientation of the search model in the new crystal is determined by comparison of the model with the unknown structure. Once the orientation is known, the position of the model in the unit cell is determined. Initial phases are calculated based on the positioned search model. Then, the initial model is refined against the experimental data to improve the fit and accuracy. Nowadays this is the most straightforward method to use and most proteins have similar proteins that can be used for this purpose. Moreover, in this scenario a structure obtained by AlphaFold [144] can be used [145].

Nonetheless, there are other methods that can be used if the chosen protein has no fitting MR model to use or the AlphaFold model obtained does not show high trust values. Heavy atoms scatter X-rays more strongly than lighter atoms, creating measurable differences in the diffraction pattern. They can be introduced into the crystal, either by soaking or co-crystallization. One method that exploits this is called multiple wavelength anomalous diffraction (MAD) and involves collecting data at two or more wavelengths, typically three that are called peak, inflection, and remote wavelength. The peak wavelength maximizes the anomalous signal so that it can obtain the strongest possible anomalous scattering for accurately locating the positions of anomalous scatterers in the unit cell. The inflection wavelength minimizes the real part of the anomalous scattering. It provides the largest difference in scattering factor compared to other wavelengths. The remote wavelength provides additional phase information to offer a reference point away from the absorption edge. It is usually chosen at an energy significantly different compared to the absorption edge. Often at a higher energy (shorter wavelength) than the peak and inflection points. It provides a dataset with minimal anomalous scattering contribution. It helps in resolving phase ambiguities by providing a control dataset for comparison with the peak and inflection data. In this way it can improve the overall accuracy of the phase determination. These differences are used to calculate phase angles through mathematical algorithms. Common choices for anomalous scatterers include selenium (often incorporated as selenomethionine in proteins), mercury, and various other heavy metals [146].

Another method is single wavelength anomalous diffraction (SAD) and it requires data collection at only one X-ray wavelength, typically chosen to maximize the anomalous signal of the heavy atoms present [147, 148]. The success of SAD phasing depends on the strength of the anomalous signal relative to noise in the data. MAD generally provides more robust phase information compared to SAD, especially for larger or more complex structures.

Combination of SAD with Molecular Replacement (MR-SAD) has also been incorporated into an automated platform for structure determination, where improved performance was shown in challenging cases [149]. Phase errors greater than 70 degrees during poor MR or SAD phases can be improved by applying MR-SAD allowing a significant portion of the model to be determined automatically.

Electron density map - model building and refinement

Using the measured intensities and estimated phases, an initial electron density map is calculated through the use of Fourier transform methods. The map obtained represents the distribution of electrons in the unit cell.

An atomic model is built into the electron density map obtained, this process is automated using a molecular graphic software such as COOT [150], Moorhen, Mol* [151] (also an online tool) or similar. Manual corrections are needed to make sure that the model fits the data obtained, such as making sure the residues align with the electron density obtained, modifying side-chain conformations to matching the density more accurately, correcting the backbone or adding waters. After these adjustments are performed, the model needs to be refined to improve the fit between the model and the experimental data. These two steps are iterative until the model cannot be improved any further.

Model validation

The final model is checked to ensure the accuracy and reliability of the results. There are several considerations to take for model validation. Firstly, geometric parameters are assessed, including bond lengths, bond angles, and torsion angles. The backbone dihedral angles (ϕ,ψ, ω) should also fall within allowed regions that define the 3D structure and influences the folding [150]. Secondly, the model should agree with the experimental electron density map obtained and the structure should make sense from a chemical perspective, especially for ligands and metal interactions. It should account for all expected atoms and molecules.

There is some parameters that can be checked such as B-factors and Ramachandran outliers. B-factors represent the degree of thermal motion or static disorder of atoms in a crystal structure. They typically range from about 10 to 100 $Å^2$. A low value is located between 10-20, a moderate value between 20-40 and above 50 is considered a high value [152]. Higher B-factors indicate greater atomic mobility or positional uncertainty. They tend to increase as resolution decreases, reflecting increased uncertainty in atomic positions. The distribution of B-factors should be reasonable and consistent.

There are several tools available for model validation purposes in the crystallographer toolbox. One of the most commonly used is a program named Coot [150], which enables model building, refinement, and validation. It offers various validation features, including Ramachandran plot analysis and density fit checks [153]. It is continuously been improved and also offers a web based version called Moorhen. Another tool is called MolProbity [154] and it checks for geometric criteria, atom clashes, and rotamer outliers. There are also tools that have many options available. PHENIX [155,156] is a conglomerate of programs that includes validation tools for both crystallography and cryo-EM structures. CCP4 [157] is another such tool which has a cloud version and a deskstop option. It offers various validation tools, including PROCHECK for geometric validation. There is a PDB Validation Server that provides a report on model quality before deposition to the Protein Data Bank and a three-dimensional space developed for describing protein backbone conformation called DipSpace which can be used for validation. In cases involving metal occupancy refinements, especially in the case of metal-protein complexes there is a program called SHELX Suite [158].

Structure deposition

After the structure is considered ready, it is validated and deposited in a public database called Protein Data Bank (PDB) for biological macromolecules and Cambridge Structural Database (CSD) for small molecules. Each structure gets a unique code assigned, which can be used to open the structure in software such as Coot, for not only visualization but direct comparison through superposition. In this way, researchers all over the world have access to the structure.

2.4.5 Resampling

To enhance experimental outcomes, a common approach is to control experimental conditions and replicate the experiment. However, simply gathering more data may not always be sufficient or feasible, particularly in the case when experiments are performed in XFELs. This is mostly due to time and sample constraints. To enhance the reliability of the results and obtain a more accurate estimate of uncertainties, statistical techniques are used to maximize the information extracted from the data. One such approach, which makes use of the high multiplicity of serial crystallography data and the fact that each diffraction pattern is derived from different crystals, is the resampling method. Resampling involves randomly picking images or intensities from the dataset, generating new datasets from the original measurements, allowing the evaluation of how data variations affect the final structural model. This method offers a clearer understanding of potential errors in the experimental process. In this case, the Bootstrap method of resampling [159] generates new data samples from the original dataset without the need for additional experiments or external sources. It relies on random sampling, where each member of the population has equal chance of selection and provides a robust way to estimate uncertainties in coordinates and other structural parameters. By generating multiple datasets and analyzing variations in the resulting models, it offers insights into the confidence levels of atomic positions and structural features. It helps to mitigate phase bias, which can arise from the influence of the initial or resting state model. This is because varying phases across resampled datasets provides a more unbiased representation of structural changes. In cases where data is limited or noisy, the resampling approach enables the generation of almost independent datasets, effectively enhancing the robustness of structural analysis. Last but not least, the use of methods such as singular value decomposition (SVD) in conjunction with resampled data improves the signal-to-noise ratio of difference Fourier electron density maps. As a result, visualization of subtle electron density changes becomes clearer, making it easier to interpret conformational transitions.

2.4.6 Electron density maps

An electron density map is a three-dimensional representation of the distribution of electrons within a crystal, which helps to visualize the atomic structure of the crystallized material. It is derived from the X-ray diffraction data and provides information about the positions of atoms in a crystal lattice. There are two ways of comparing the data measured with a model. In the first one, the electron density is depicted in a map called

$2F_o$ - F_c, where F_o represents the observed structure factors derived from experimental data, and F_c represents the structure factors calculated from the model. The second approach is through F_o - F_c maps, which not only highlight areas that align with the model but also indicate regions where modifications might be needed, such as adding or removing atoms or adjusting side chain orientations. Most commonly, neutral areas where there is good agreement between the observed and calculated data is colored blue in Coot. When Green appears, it indicates areas where the observed data (F_o) is stronger than the calculated data (F_c), suggesting that the model underestimates electron density in that region. Practically, it means that the model might be missing electron density, such as an unmodeled atom or an incorrect positioning of an atom or side chain. It implies that additional atoms or modifications to the model are needed in those areas. Red regions usually represent areas where the calculated data (F_c) is stronger than the observed data (F_o), meaning that the model overestimates electron density in that region. This could indicate that the model includes atoms that don't fit the experimental data, or that the positioning of atoms might be incorrect. Red regions are often indicative of parts of the model that need to be adjusted or removed, such as in the case of incorrect side-chain conformations or misplaced atoms.

Isomorphous Difference Density Maps

These maps are also called Difference Fourier Electron Density Maps, and are used to visualize changes in electron density between two different sets of measurements. This map is obtained by subtracting the electron density of the native map from the electron density of the derivative map as follows:

$$\text{Isomorphous Difference Map} = \rho_{\text{derivative}} - \rho_{\text{native}}$$

where $\rho_{\text{derivative}}$ and ρ_{native} are the electron densities for the derivative and native crystals, respectively.

Typically, these maps compare a dark or resting state versus a light-activated state and highlight differences in electron density that occur during transitions, such as between a ligand-bound and unbound state, or before and after a chemical reaction.

This method is sensitive to small changes in electron density, more than any other type of electron density map [160]. This sensitivity makes it an ideal tool for detecting localized and subtle conformational changes, such as the movement of side chains or small structural rearrangements. However, this technique assumes that both measurements: dark and light states, share the same unit cell and space group. This means it is most useful for cases where the changes are confined to specific regions of the molecule. It is less effective in situations where large, widespread conformational changes occur throughout the entire protein.

Positive differences in electron density are indicated by areas of increased density in the difference Fourier map. It typically reflects new electron density that was absent in the resting state but has arisen due to the reaction or transition of interest. Conversely, negative differences or regions of decreased density, suggest that electron density has been lost as the reaction progresses.

It is generally assumed that any bias in phase estimation, favoring the resting state , does not greatly distort the quality of the difference Fourier maps. This assumption holds in many cases, as the maps are primarily used to detect small, localized changes in electron density. However, the impact of this phase bias can vary depending on the specifics of the experiment, especially when the phase estimation is not accurate or the differences between the two states (resting and activated) are large or widespread. Despite this, the technique remains the leading approach in TR-SX, and is regarded as one of the most reliable measures of experimental success in structural studies [160, 161].

Omit maps

It is a type of electron density map used to evaluate the quality and correctness of a structural model, particularly during model refinement. It helps identify regions where the model may need improvement, such as areas with incorrect atom placements or missing atoms. In the process of refining a structural model, an initial model is built based on the experimental data. To evaluate whether a particular part of the model is fitting the experimental data correctly, the omit map is generated by removing a specific region. The omitted region can be a residue, a side chain, or even a larger section of the protein. The remaining part of the model is then used to calculate a new electron density map without the omitted region. The omit map is then calculated from the remaining parts of the model, and the electron density in the omitted region is observed. If a well-defined peak of electron density appears in this omitted region, it suggests that the region was missing or incorrectly modeled in the initial model, and the new data can be used to refine it. On the other hand, if no significant density appears in the omitted region, it indicates that the omission was appropriate, and the model does not need adjustments in that area.

Omit maps are particularly useful for refining models where parts of the structure are uncertain, such as disordered regions or areas of the molecule that are hard to model. They help ensure that any modifications made to the structure are supported by the experimental data and prevent over-fitting, where a model becomes too closely tailored to a particular dataset.

In summary, an omit map helps validate and refine structural models by evaluating regions that may require adjustments, making sure that the final structure reflects the experimental data accurately.

SVD

Singular Value Decomposition (SVD) is a powerful mathematical technique often used in data analysis to decompose and interpret complex datasets. In the context of crystallography, SVD helps to improve signal-to-noise ratios and uncover patterns in the data. The process of doing an SVD analysis if as follows: A set of difference Fourier maps is arranged into a matrix, where each column represents a flattened map or specific features of the map. SVD decomposes this matrix to identify dominant patterns and separates them from the background noise. The principal components represent the most significant features in the data, such as real structural changes in the protein or other macromolecule. By focusing on the principal components, SVD filters out random noise or inconsistencies in the data, resulting in improved clarity and signal-to-noise ratio in the maps. In time-resolved studies, SVD can also track how these dominant features evolve over time, providing insights into the dynamics of conformational changes.

2.5 X-ray solution scattering - XSS

This technique consists of studying the structure of macromolecules in solution, as opposed to crystallography, which requires a protein to crystallize. This results in a clear advantage in the case of proteins with flexible regions that do not crystallize or in the case of large protein complexes, which might also be reluctant to form a crystal structure [162]. The structural information provided is low-resolution and is useful to complement high-resolution methods such as X-ray crystallography or cryo-EM as well as computational approaches such as molecular modeling to study dynamics [163].

Proteins are disordered in solution and scattered randomly in space, and in this technique the intensity is measured as a function of the scattering angle, which can be small or wide. The scattering pattern reflects the overall shape, size, and structural arrangement of the molecules, and mathematical models are used to interpret the data and reconstruct 3D models of the macromolecule. As mentioned before, this technique can be performed using small or wide angles, and the name of the technique varies with each approach, varying also the information each one provides. Small angle X-ray scattering or SAXS typically uses 1-5 $^{\circ}\mathrm{C}$ and gives information on global shape and size of macromolecules [164]. Wide angle X-ray scattering or WAXS collects at larger scattering angles, above 5 $^{\circ}\mathrm{C}$, and shortens the distance to the detector yielding details about secondary structure or local order [165, 166]. This technique can also be performed in a time resolved manner and is referred to as TR-XSS, where snapshots of structural changes at different time points are captured by using the pump-probe principle previously described for Crystallography in chapter 2 [167]. At wide angle, it is possible to observe how the X-ray scattering properties of the solvent change as a sample is heated [168]. This property we utilize in paper IV.

Chapter 3

Paper I

Lipidic cubic phase serial femtosecond crystallography structure of a photosynthetic reaction centre

3.1 Overview

Serial crystallography is a technique that enables structural analysis of micro-crystals, which are too small for conventional X-ray crystallography. This creates the need for research on new strategies to produce small well-diffracting showers of crystals.

This study describes the conditions for the growth of micro-crystals from the photosynthetic reaction center of *Blastochloris viridis*. Using a lipidic cubic phase (LCP) crystallization matrix, a seeding protocol was developed with detergent-grown crystals that have a different crystal packing.

Former sponge phase approaches (LSP) yielded the second SFX structure from an XFEL [169,170]. This was extended to the first atomic resolution structure of a membrane protein using SFX data-collection, but the hit-rate was low and the resolution obtained was 3.5 Å. Detergent protocols were screened, which enabled to go from RC*vir* macro-crystals to micro-crystals. Then LCP was tested to improve the resolution. When exposed to XFEL radiation, a diffraction resolution of 2.25 Å was achieved. This resolution constitutes a 0.15 Å improvement compared to earlier micro-crystal forms.

Ubiquinone was included in the crystallization media and as a result the electron density observed within the mobile Q_B pocket was comparable to that of other co-factors in the structure.

In light of this, LCP micro-crystallization conditions could potentially support time resolved diffraction studies of electron transfer reactions to the mobile quinone, potentially allowing the observation of structural changes related to the complete reduction of this ligand.

3.2 Results and Discussion

3.2.1 Crystallization screening results

Considering the reduced sample requirements of an LCP injector, we investigated the feasibility of obtaining an LCP crystal form of RC for serial crystallography. The micro-crystallization protocol developed consists of using crushed vapor diffusion grown crystals as seeds for the growth of LCP crystals. In addition, ubiquinone 2 (UQ_2) is added to the crystallization conditions in a ratio of 100:1 UQ_2:RC. The reason behind this that the final electron acceptor (Q_B) is lost during the purification process.

The results of all the crystallization trials performed can be seen in figure 3.1. Initial LCP conditions were chosen based on previous studies [171] and also taking into account commercial screening protocols specifically developed for LCP crystallization. RC crystallization screening results showed that small molecule diols or triols that were used as additives yielded crystals that are shown in figure 3.1a. This suggests that they might behave in a similar way to heptanetriol, promoting contacts for crystal formation [172, 173]. When these conditions were tested in a gas tight syringe setup, small needle shaped crystals were obtained that are shown in figure 3.1b. These crystals that yielded diffraction with 2.4 Å resolution. However, because of their characteristic shape, the hit rate obtained was low. Diffraction was attributed to the largest crystals exclusively. These conditions, having proven effective in crystal formation, were used for screening in glass crystallization plates [174]. The results obtained with this approach are shown in figure 3.1c, where a different crystal shape can be appreciated. Diffraction obtained for these sturdier looking crystals reached 2.2 Å resolution, but given the large size of the unit cell, the dataset could not be indexed.

Seeding is a common tactic often used to improve the quality of protein crystals. Since LCP crystals are hard to crush, an alternative approach was tried whereby the seeds used were obtained from vapor-diffusion RC crystals [175]. The seed stock was mixed with the protein in solution and then joined in a 40:60 ratio with monoolein to form the LCP mix. Each glass well would then contain 100 μL precipitant solution with 10-15 μL LCP-RC-seed mix. When using this approach, crystals shown in figure 3.1d could be obtained that exhibited a high hit rate in SFX and were processed to obtain a structure with 2.25 Å resolution.

During the crystallization screenings, it was noted that the size of the crystals obtained in LCP conditions varied owing to the monoolein ratio used. The more monoolein, the smaller the crystals with higher density. When the precipitant concentration was increased, the crystals were larger, but an increase in the heptanetriol concentration also induces the transition to the sponge phase (LSP). Different seed concentrations were tested but did not appear to affect the end result in LCP setups. While the approach is unconventional, since it mixes two different kinds of crystallization setups, the seeds aided in the LCP crystal formation although the space group obtained with these two approaches is different. LCP crystals have the $P2_12_12$ while detergent crystals have $P4_32_12$ [175].

3.2.2 Diffraction obtained and data collection

The samples were transported at room temperature to the SACLA XFEL facility in Japan and data collection took place at the BL3 beam-line in November 2018. A structure was produced with the diffraction data that reached 2.4 Å resolution with 93,000 indexed images which is shown in figure 3.2.

Even though the seeds were produced from vapor-diffusion crystals in space group $P4_32_12$, crystals obtained grown in LCP were from the space group $P2_12_12$. These results agree with those obtained earlier for LSP crystals [176]. Internal distance-matrix analysis to compare the internal spatial relationships of atoms within a molecule [175, 177] revealed a structure with an average C-α distance difference of 0.35 Å compared to

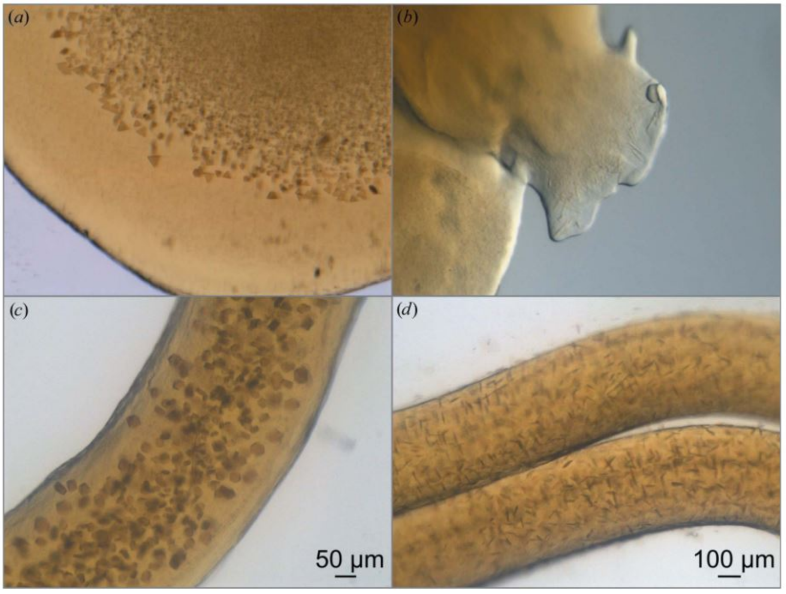


Figure 3.1: Micro-crystallization trials a) Results of the first stages of screening with different additives. b) These crystals were grown in air-tight syringes and exhibit needle like shape. They reached 2.4 Å resolution diffraction at SACLA with a low hit rate. c) First results obtained after switching to glass wells [174] which got 2.2 Å resolution but their long crystal axis prevented them from being indexed. d) Resulting crystals from the seeding approach that yielded 2.24 Å resolution.

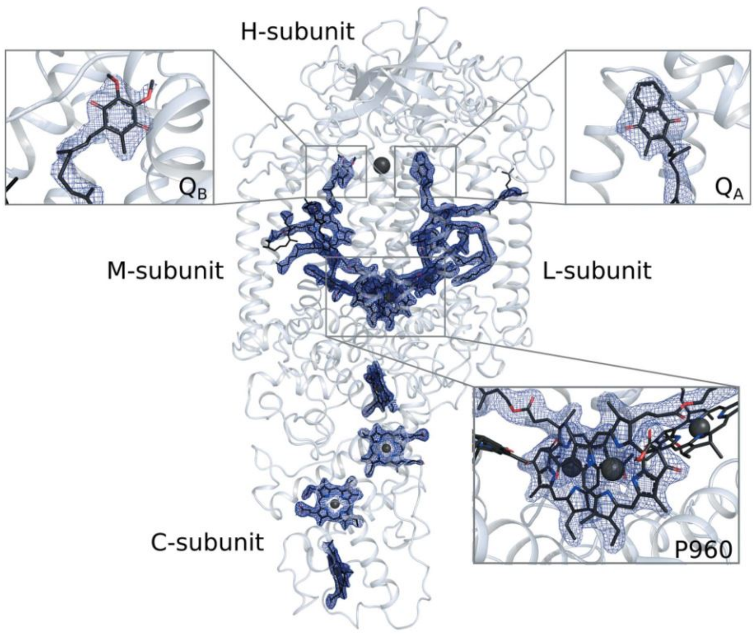
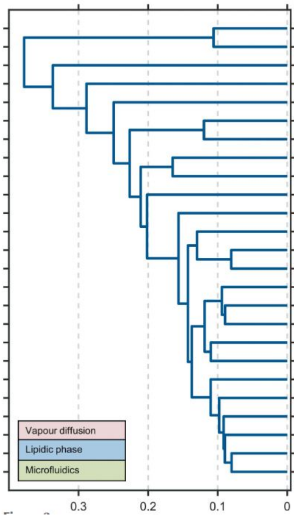


Figure 3.2: Electron density from SFX LCP crystals With a 2.25Å resolution, these LCP crystals exhibited similar quality to other cofactors in the mobile Q_B for the $2F_{obs}$-F_{calc} (shown in blue). Contour was set as 1σ, where σ corresponds to the standard deviation of the electron density of the map



PDB ID	Year	Res (Å)	Collaboration	Xray Source	Temp (K)	Space Group
7Q7Q	2019	2.25	Neutze	XFEL	293	P 21 21 2
7Q7P	2018	2.40	Neutze	XFEL	293	P 21 21 2
1R2C	2003	2.86	Norris	Synchrotron	293	P 43 21 2
4CAS	2013	3.50	Neutze	XFEL	293	P 21 21 21
4AC5	2011	8.20	Neutze	XFEL	263	P 21 21 21
2X5V	2010	3.00	Neutze	Synchrotron	293	P 21 21 2
2X5U	2010	3.00	Neutze	Synchrotron	293	P 21 21 2
2WJN	2009	1.86	Neutze	Synchrotron	100	P 21 21 2
2WJM	2009	1.95	Neutze	Synchrotron	100	P 21 21 2
3D38	2008	3.21	Ismagilov	Synchrotron	100	P 31 2 1
1PRC	1988	2.30	Michel	Rot. Anode	271	P 43 21 2
5O64	2014	3.40	Neutze	XFEL	293	P 43 21 2
5O4C	2015	2.80	Neutze	XFEL	293	P 43 21 2
5NJ4	2016	2.40	Neutze	XFEL	293	P 43 21 2
3T6D	2011	1.95	Cogdell	Synchrotron	100	P 43 21 2
3T6E	2011	1.92	Cogdell	Synchrotron	100	P 43 21 2
2I5N	2006	1.96	Ismagilov	Synchrotron	100	P 43 21 2
3G7F	2009	2.50	Norris	Synchrotron	100	P 43 21 2
1VRN	2005	2.20	Norris	Synchrotron	100	P 43 21 2
2JBL	2006	2.40	Ditchfield	Synchrotron	263	P 43 21 2
2PRC	1997	2.45	Michel	Synchrotron	263	P 43 21 2
3PRC	1997	2.40	Michel	Synchrotron	263	P 43 21 2
7PRC	1997	2.65	Michel	Synchrotron	263	P 43 21 2
6PRC	1997	2.30	Michel	Synchrotron	263	P 43 21 2
5PRC	1997	2.35	Michel	Synchrotron	263	P 43 21 2

Figure 3.3: Tree analysis of RC_{vir} structures taken from PDB This table contains the PDB identity number of each structure considered, the year it was uploaded, the resolution obtained, research group, X-ray source used, temperature and space group. The colors code for the crystallization method used, pink being vapor diffusion, blue LCP and green micro-fluidics. The sorting of the structures obtained are related to the mean difference of their internal Cα distances as previously described [177]. The entries from this work are 7q7q and 7q7p.

previous structures. This is shown in figure 3.3, where it can be seen that not even the LSP structures solved with the same space group appear to be related. The LCP micro-crystals exhibited unique unit cell axes of a = 84.9, b = 125.3, c = 182.7 Å. When compared to former crystal packing obtained, a and c axes were to some extent similar to some earlier structures in LSP (PDB id 5x5u and 2wjm) while the b-axis resulted shorter by 20 Å for earlier room-temperature and 13 Å in cryogenic structures.

A second dataset was collected at SwissFEL in October 2019 with 160,000 indexed images which resulted in a resolution improvement to 2.25 Å. The unit cell dimensions obtained were a = 84.7, b = 125.1, c = 182.4 Å. These values closely matched the ones obtained from the SACLA structure dataset. These findings suggest that crystal packing differences have a greater impact on structural refinement than the data-collection method. Both LSP and LCP $P2_12_12$ crystals exhibit 2D stacking, with the LCP form showing tighter packing due to its shorter b-axis.

3.3 Structural observations

3.3.1 Mobile quinone occupancy and position

The Q_B pocket exhibited electron density corresponding to the ubiquinone that was incorporated to the crystallization mix (figure 3.4c). The addition of UQ_1 or UQ_2 to the LCP crystallization conditions was possible owing to the fact that it dissolves effectively in monoolein. This was not the case for previous LSP structures [178] [176], where monoolein took the place of native ubiquinone (figure 3.4d). The electron density obtained for Q_B (figure 3.4c) was comparable to the electron densities obtained for other co-factors in the molecule (figure 3.2). B-factor values quantifies from 10-100 Å^2 how much an atom deviates from its average position. This movement can be due to thermal motion or static disorder, but high B-factors can also indicate poor crystal packing or unresolved electron density. Usually, a higher B-factor indicates greater atomic mobility or uncertainty over the position assigned, while lower values suggest that the atom is more rigidly fixed in place. If 100% occupancy is assigned, the B-factor value obtained for Q_B is 60.9 Å^2 and 39.0 Å^2 for Q_A. The difference in value observed when comparing these values may be attributed to the increased mobility attributed to Q_B, though it can also be due to the occupancy value assigned being lower than 100% in reality.

In figure 3.4a, two alternative positions of Q_B are shown according to data from previous studies [179], where only the proximal position makes its reduction possible (depicted in red). SFX LCP structure obtained in this study (figure 3.4a, black color) agrees with the proximal active form previously described, the only difference being mainly a partially repositioned Phe216$_L$ head group which was proposed to facilitate proximal binding of ubiquinol. When compared with structures showing native UQ_9 [180](shown in figure 3.4b), both the Phe216$_L$ head group and proximal position concur. Its also worth mentioning that only the first isoprene unit can be confidently modeled owing to the flexibility of the UQ tail. Initially, the distal positioning of UQ was accepted as predominant

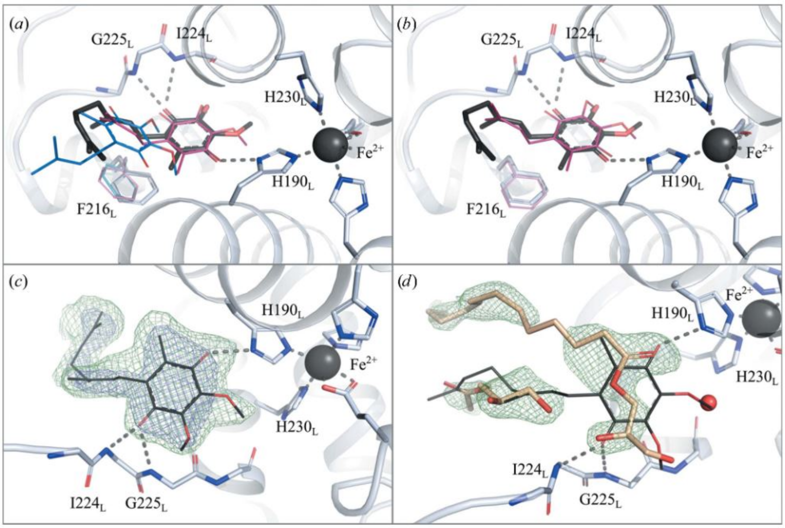


Figure 3.4: Q_B binding site overview a) Ubiquinone (black) is positioned superposed with the proximal (red) and distal (blue) binding sites for Q_B in RC_{Sph}, taken from PDB identities 1aig and 1aij. b) A comparable result is depicted for RC_{vir} taken from PDB id 2i5n. c) Density obtained for the LCP SFX structure with added quinone to the crystallization environment. d) LCP structure 2wjn from PDB showing a monoolein molecule placed where the ubiquinone is placed. The omit electron density maps F_{obs}-F_{calc} is shown as a green mesh with 3σ. Electron density maps colored blue ($2F_{obs}$-F_{calc}) have σ set as 1.

due to electron density observed [181]. This notion was then changed by Fourier Transform Infrared Spectroscopic studies [182] that showed that Q_B and Q_B^- occupy the same position. It was later shown [183] that the pH has a direct influence on the quinone position, being that at higher pH (above pH 9), a decrease in the proximal Q_B population was observed. This indicates a pH-dependent interaction influencing the proximal position at higher pH. Flash freezing studies have shown a preference for Q_B binding predominantly in the proximal position in both neutral and charge-separated states [184].

3.3.2 Second plausible binding site

Electron density features supporting a second binding site was spotted in the SFX LCP RC_{vir} structure for ubiquinone and it is shown in figure 3.5. Even though the B-factor value obtained for this alternative quinone binding site is 91 Å^2 when 100% occupancy is assumed, this finding is supported by other crystallographic structures obtained of RC_{vir} [180] [52] [185] [186] and also by a cryo-electron microscopy (cryoEM) [187]. It has not been yet proposed a function for this second site, even though the two amino acids in the L and M chain that might stabilize this interaction are in fact conserved between species [185]. The most relevant and worth noting is that it is stable enough to have appeared in several crystallization experiments, thus may be one of many places where the quinone gets stabilized by the protein when it moves to the binding site.

3.4 Conclusions

This new crystallization method using LCP and applying seeding with detergent based obtained crystals yielded high quality crystals and made possible a resolution improvement. Moreover, it makes possible time-resolved studies, enabling us to see the mobile quinone in the process.

A second quinone binding site was observed that is supported by other studies, stable enough to be repeatedly observed in several crystallization setups.

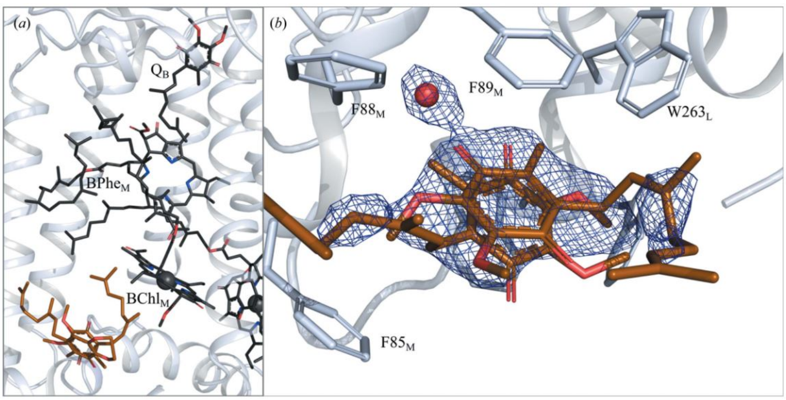


Figure 3.5: Structure proposed for the ubiquinone molecule a) The ubiquinone molecule is shown in orange occupying its proposed space. b) Electron density map $2F_{obs}$-F_{calc} contoured at 1σ obtained with the new protocol containing the added ubiquinone. As it can be seen, there are two possible conformations that this molecule can occupy.

Chapter 4

Paper II

A user-friendly goniometer-compatible fixed-target platform for macromolecular crystallography at synchrotrons

4.1 Overview

Crystallization using fixed-target platforms is an efficient method for macromolecular structure determination and it offers advantages such as reduced sample consumption and simpler handling given that allows direct crystallization. A reliable and cost-effective system is presented, featuring X-ray transparent membranes integrated into ‘chips’ tailored for high-throughput hanging-drop protein crystallization setups. Crystals grown on these chips can be sealed in a ’sandwich’ and packed into customized 3D-printed cassettes to be transported at room temperature to synchrotrons. The chips can be directly mounted onto the goniometer without alterations to beam-line optics thanks to a custom made 3D-printed base holder. This platform was validated using diverse beamlines, crystal samples, and data collection strategies. High-resolution data were obtained at room temperature from both single-crystal and serial crystallography experiments on fixed targets. Lysozyme was resolved to 1.9 Å at room temperature and 1.6 Å under cryogenic conditions using single crystal crystallography. Meanwhile, more complex proteins such as the reaction center from *Blastochloris viridis* (RC*vir*) and cytochrome c oxidase (CcO) were resolved at 3.3 Å and 2.3 Å, respectively, using fixed-target serial crystallography. When compared to existing platforms,

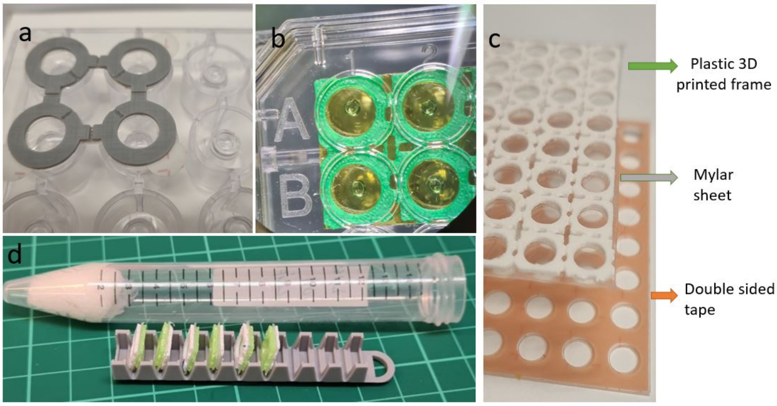


Figure 4.1: Array of devices The chips can be used to hanging method techniques where they fit the wells on 24 well plates a), or 96 well plates b). c) Shows how the setup looks for assembly, showing the printed frame with the mylar sheet on top of the double sided tape. d) The 3D printed cassette to transport assembled chips for data collection next to the 15 mL Falcon tube.

these chips proved to be useful for crystallization, crystal screening, transport and data collection, making them a useful tool for structural biology research.

4.2 Design and Description

4.2.1 Design of crystallization platform components

Computer-aided design (CAD) technology was used in AutoCAD (Autodesk 2020) to develop 3D models around a framework compatible with standard 24-well (shown in figure 4.1a) and 96-well (Greiner and TPP, in figure 4.1b) plates for hanging-drop crystallization assays. The 'backbone' design of the frame was organized as an array of flat, circular 'chips'(shown in figure 4.1a, b and c) that were assembled with a layer of X-ray transparent membrane (either Mylar or Kapton) of variable thickness (3 to 12 microns) using double-sided adhesive tape as shown in figure 4.1c.

Figure 4.2: Data collection The chips can be placed onto the holder and into the goniometer for data collection in most beam-lines, as shown in these two pictures.

The 0.75 mm thickness of the 3D-printed support allows a lightweight nature and flexibility, ensuring that the platform can be easily divided with a scalpel. The chip design consists of two devices:

- Crystallization Chip:

A 'frame-adhesive-membrane-adhesive' disposition, suitable for crystallization. It allows to directly stick the device and seal the wells for vapor diffusion crystallization setups.

- Crystal encapsulation chip:

This design includes a single layer of double-sided tape, making it 'frame-adhesive-membrane'. When joint with the 'crystallization chip', it is used to encapsulate crystals in a sandwich structure. This prevents the crystals from drying out.

The adhesive layer provides a 70 μm spacer for encapsulated crystals, which can be chosen to be doubled in case of using two 'crystallization chip' devices to pack the sample. This approach is shown in figures 4.3d and e.

4.2.2 Crystal storage and transportation

For storage and transport of crystal samples at room temperature, a compact 3D-printed device was designed that consists of a stack of chip-holders or "cassettes". These cassettes are designed to fit within a 15 or 50 mL Falcon tube as shown in figure 4.1d and 4.4. Each Falcon tube is filled with moist cotton soaked in a crystallization solution or buffer, and the tube is tightly sealed with para-film to maintain humidity around the crystals during transport to synchrotron sources.

4.2.3 Goniometer mounting design

To facilitate crystal mounting on the goniometer, the previously reported goniometer-compatible flow-cell device was adapted [130] from holder of a capillary to hold the sandwiched chips for alignment and data collection. The original design was altered to incorporate grooves for holding or clipping various chip sizes and arrays (figure 4.2). The height of the holder was kept low to prevent any perturbation in the beam-line optics.

During data collection, the sandwiched chips are removed from the cassette, placed in the groove of the holder, and mounted on the goniometer. A magnetic disk needs to be inserted into the base of the plug making sure the polarity fits the goniometer at the beam-line. The final version of the device is shown in figure 4.3a, b and c.

4.3 Results and Discussion

4.3.1 Structures obtained and data

The chip setup for data collection was tested in several beam-lines, with different proteins and different sample types. In table 4.1, the results obtained are shown. Lysozyme dataset was obtained at cryogenic temperature with rotation collection, and also at room temperature. Last but not least, an LCP sample was used with Cco crystals. As can be seen in the table 4.1, the resolution ranged from 1.6 to 3.3 Å, and all the structures could be obtained by using the fixed target setup.

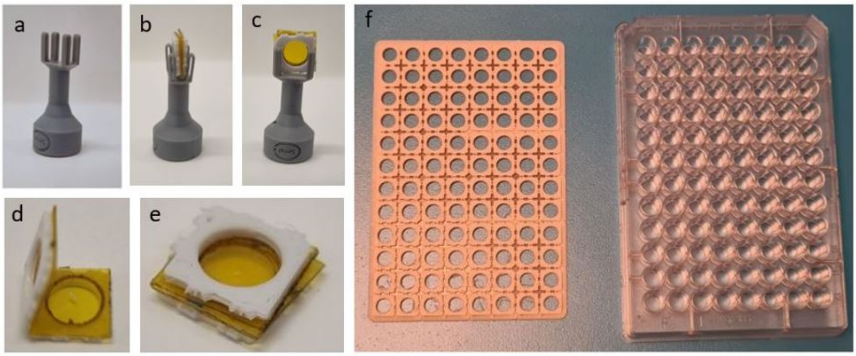


Figure 4.3: 96 well plate chip design The pictures shown illustrate how the holder looks a) without the chip and b) and c) with the chip attached and from different angles. d) illustrates the possibility of closing two crystallization chips together and e) shows the closed chip. f) Picture of the 96 well plate chip side by side with the 96 well crystallization plate.

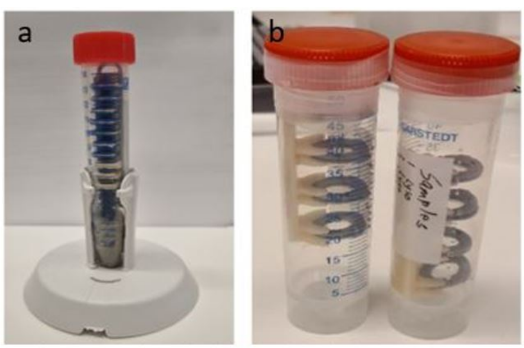


Figure 4.4: Chip holder design for transport Both 24 well plate and 96 well plate chips can be successfully transported using a) 15 mL Falcon tubes and b) 50 mL Falcon tubes.

4.3.2 Data Collection Strategies

Table 4.1: Statistics of all data collections Data collection statistics are shown for the lysozyme single crystal dataset, both room temperature and cryo-cooled. For RC*vir* both 24 well plate chips and 96 well plate chips were used, one of which was crystallized directly as a hanging drop from the chip mylar. CcO is the only LCP sample used.

	Lysozyme	**Lysozyme**	**RC*vir***	**RC*vir***	**RC*vir***	**ba3-type CcO**
Chip design	96-well	96-well	24-well (loaded)	24-well	96-well	96-well
On-chip crystallization (Yes/No)	Yes	Yes	No	Yes	No	No
Crystal type	Single crystal	Single crystal	Micro-crystals	Micro-crystals	Micro-crystals	LCP micro-crystals
Method	Rotational	Rotational	SSX	SSX	SSX	SSX
Data Collection						
Beamline	PF BL5 A	AR-PF NW12A	MAXIV- BioMAX	MAXIV- BioMAX	ESRF-ID29	MAXIV- BioMAX
Collection temperature (K)	100	293	293	293	293	293
Space Group	P43212	P43212	P43212	P43212	P43212	C121
Cell dimensions						
a, b, c (Å)	78.0 78.0 37.2	79.1 79.1 37.9	224.2 224.2 113.8	223.8 223.8 113.5	224.2 224.2 113.8	145.8, 100.3, 96.62
α, β, γ (°)	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90, 126.8, 90
Resolution (Å)‡	39.0 - 1.9	39.6 - 1.6	36.0 – 3.3	36.0 - 2.8	36.0-3.3	37.2 – 2.3
Rsplit (%)†‡			32.5 (107.3)	21.6 (82.9)	13.8 (142.6)	23.2 (54.7)
I/σ(I)‡	16.1 (2.4)	16.0 (2.1)	3.1 (0.99)	4.1 (1.0)	6.6 (0.9)	3.07 (1.5)
CC($^1/_2$)‡	99.1 (18.2)	99.9 (62.4)	87.9 (36.7)	93.6 (52.9)	99.9 (57.1)	85.3 (57.0)
Completeness (%)	95.4	95.7	100	100	100	99.9
Multiplicity‡	6.2	6.0	691.2	412.0	2258.7	2.0 (2.0)
Number of collected images	900	800	70133	150855	201910	32613
Number of indexed patterns			18128	36048	72801	5852
Number of total reflections	56118	95667	58248555	62309072	114467587	97482
Number of unique reflections	9071	16028	84271	151216	50678	49574
Refinement						
Resolution	38.98 - 1.9	31.4 - 1.6	36 - 3.3	36.8 - 2.8	49.6-3.3	32.17 - 2.30
R_{work} / R_{free} (%)	22.2/26.6	18.8/21.5	24.0/29.9	48.9/52.6	22.7/28.3	22.5/26.1
Number of atoms	1040	1073	10313	10402	10256	6420
Average B factor (Å^2)	33.8	20.50	80.02	59.69	98.4	50.59
R.m.s deviations						
Bond lengths (Å)	0.008	0.005	0.018	0.014	0.015	0.014
Bond angles (°)	1.00	0.76	2.19	1.61	1.92	1.94
‡Values in parenthesis is those of the highest resolution shell						

The platform demonstrated its adaptability by supporting different data collection conditions and samples in multiple synchrotron facilities and protein systems, all of which are shown in table 4.1. Diffraction studies were performed both at cryogenic and room temperature for lysozyme, which proofed compatibility with many experimental setups. For cryogenic data collection, lysozyme crystals were flash-frozen and mounted onto the chip, yielding high-quality data at 100 K. At room temperature, the setup facilitated direct alignment and data collection, by minimizing sample handling and ensuring efficient workflow integration.

In serial synchrotron crystallography (SSX), fixed-target strategies at BioMAX (MAXIV, Sweden) and ID29 (ESRF, France) highlighted the platform's robustness and ease of use. Data from CcO and RC*vir* crystals

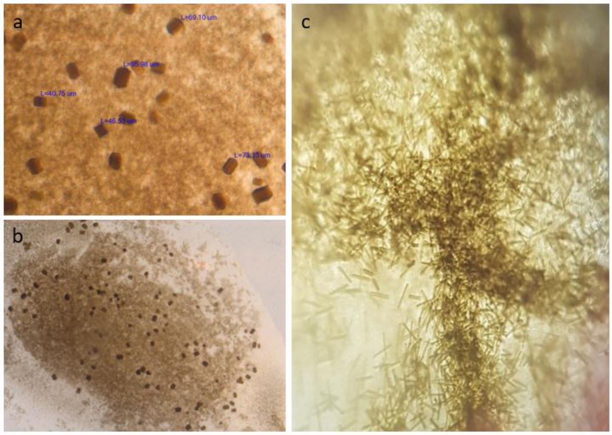


Figure 4.5: RC*vir* crystals obtained Here, the different morphologies exhibited by the change in crystallization setup are shown. In figures a) and b) the crystals grown directly on the chip are shown and c) shows sample from sitting drop method.

were collected from both LCP and solution crystal samples respectively. Despite minor challenges such as movement of liquid-containing samples during data collection, the chip-based approach maintained data quality, reinforcing its utility for high-throughput experiments and enabling rapid acquisition of datasets suitable for structure determination.

4.3.3 Crystallization and performance

RC*vir* crystals were grown directly on the 24 well chip by using the hanging drop method. As can be seen in figure 4.5, the morphologies and sizes obtained for different setups were not alike. While these chip-grown crystals exhibited good diffraction properties, partly due to their size, their utility for complete dataset collection was limited due to lower completeness. This highlights the need for optimizing crystallization conditions on-chip to achieve the desired crystal size and quality. However, crystals grown in sitting drops provided robust datasets, underscoring the chip's flexibility in accommodating various crystal preparation techniques.

In addition to crystallization, the chips served as an alternative to traditional loops for mounting larger crystals. Although this resulted in approximately 40 degrees of data loss due to the chip frame obstructing the X-ray beam, sufficient data for structure elucidation were successfully collected, demonstrating the platform's practicality and versatility.

4.3.4 Comparison with alternative chip options

Many platforms are continuously being developed to simplify the process of obtaining serial crystallography datasets. Figure 4.6 shows a table that compares some methods to the chip developed in this study. Focus was placed on practicality during sample placement, manipulation, and data collection.

One of the convenient aspects of the available devices that was assessed is the need to manipulate the base holder as to successfully collect data in any beam-line. Base manipulations are needed for eight of the devices considered, while four do not require extra manipulation on the device for placement on the goniometer during beam-time. The straightforward setup without the need for specialized holders arises from the use of the spine system commonly available at beam-lines. This significantly streamlines the data collection process, making it faster and more efficient.

Another factor examined was the potential for direct crystallization on the device, referred to as "in situ" crystallization in the table. Only four out of the twelve devices support in situ crystallization, enabling samples to be obtained directly on the device avoiding the need for additional handling prior to measurement. This not only saves time but also reduces the risk of crystal damage during transfer. Based on these two advantages alone, only two devices from the twelve evaluated combine the ease of using the spine holder with the ability to perform in situ crystallization [188].

Additionally, disposable devices offer advantages, such as minimizing cross-contamination risks and eliminating the need for cleaning, which enhances efficiency. They are particularly useful for high-throughput sample processing, as there is no concern about wear and tear. However, their use generates waste and may lead to supply chain vulnerabilities. Compared to the devices reviewed, four are fully disposable, two are partially disposable, and two do not specify.

Most of the devices employ a raster scan strategy for data collection, though three also allow directed data collection. The choice between these strategies depends on factors such as sample type, size, and consistency. For the samples tested in this study, liquid-contained crystals were observed to move across the chip, particularly in the 24-well plate chip and the 96-well plate device. However, this movement did not affect the quality of the data during the collection. The impact of this movement on time-resolved data collection remains a topic for further discussion. Notably, no movement was observed when the sample was embedded in LCP (lipidic cubic phase). Most devices are compatible with both room and cryogenic temperature data collection and do not require specialized sample-loading systems.

Figure 4.6: Table comparing available chip options. Various characteristics of diverse available options for crystallographic data collection have been considered. The first row named "base" evaluates whether the device can be directly used with no need for modifications for data collection at the beamline or if it requires a special add-on. The second row compares whether the platform allows crystallization directly using the device, which enables skipping sample handling. Disposability, reusability, and temperature conditions were also compared.

Base	in situ	disposable	reusable	collection strategy	R-temp	Cryo-temp	sample load system	Reference
spine	no	no	Yes	raster	Yes	Yes	yes	Illava et al, 2020
customized	yes	no	yes	raster	Yes	no	no	Ren et al, 2020
customized	yes	unknown	unknown	directed	yes	no	yes	Mueller et al, 2015
spine	yes	unknown	unknown	raster	Yes	Yes	no*	Lieske et al, 2019
customized	no	yes	no	raster	Yes	unknown	yes	Lee et al, 2020
customized	no	partially	yes	directed	yes	unknown	yes	Carrillo, Mason et al. 2023
customized	no	partially	unknown	directed, oscilation	Yes	unknown	yes	Wierman, Paré-Labrosse et al. 2019
customized	no	yes	no	raster	Yes	Yes	no	Doak, Nass Kovacs et al. 2018
customized	no	yes, mylar	yes	raster	Yes	Yes	no	Rabe et al, 2020
customized	no	no	Yes	raster	Yes	Yes	no	Bjelčić, Sigfridsson Clauss et al. 2023
spine	no	no	Yes	raster	Yes	Yes	no	Park, Choi et al. 2020
spine	yes	yes	no	raster	Yes	Yes	no	This work

4.4 Conclusion

This study highlights the value of the 3D-printed chip platform by addressing the challenges of efficient crystallization, storage, transportation, and mounting processes. The modular and versatile design accommodates various plate sizes and crystal types, enabling broad applicability in both single and serial crystallography. The storage system effectively maintains humidity for crystal samples, safeguarding integrity during transport and data collection. Furthermore, the chip-holder placement on top of the goniometer ensures quick alignment and efficient data collection, making the experimental workflow straightforward. The potential the design has to be packed with extra space might result in an advantage in certain setups, such as when working with LCP. The platform has the potential to become an indispensable tool for diverse crystallographic applications, fulfilling the initial motivation of creating a more accessible and efficient solution for structural biology research.

Chapter 5

Paper III

No observable Sub-picosecond structural impulse within a photosynthetic reaction centre

5.1 Overview

Photosynthesis is arguably the most important process on Earth, as it enables the life of most organisms directly or indirectly, both on its surface and in the oceans. Photosynthetic reaction centers are the molecules that enables the conversion of sun light to chemical energy. Light-induced conformational changes on millisecond [189] and picosecond [190] timescales have previously been by revealed Time-resolved X-ray crystallography studies of the *Blastochloris viridis* reaction center. Investigations were here extended into the femtosecond domain to determine whether photon absorption triggers sub-picosecond structural responses. Neither global nor local structural motions were detected associated with the photoexcitation of the special pair of bacteriochlorophylls. The continuous development of the field along with machine learning algorithms for merging and scaling serial X-ray crystallography data [191] might in the future be able to recover a faint structural signal. However, the findings presented in this work indicate that no consistent structural impulse is observable on sub-picosecond timescales.

5.2 Background

There is an ongoing debate about whether protein conformational changes assist charge separation or stabilize the charge-separated state. Prior studies on X-ray crystallography and time-resolved spectroscopy show mixed evidence for light-induced protein motion and its role in charge transfer. Flash-freeze X-ray crystallography suggested that conformational changes occur in the binding pocket of the mobile quinone (Q_B) upon receiving an electron, indicating that protein dynamics may contribute to stabilizing the charge-separated state [179]. Similarly, time-resolved Laue diffraction studies observed localized movements, such as Tyr162L shift, on millisecond timescales after photo-activation [176].

Femtosecond infrared spectroscopy revealed low-frequency vibrations in the reaction center, which were attributed to coherent nuclear motion of the co-factors or their coupling with the surrounding protein environment [192–194]. These conflicting results underscore the complexity of protein-co-factor interactions and the need for further investigation into the dynamics underlying charge transfer events in photosynthesis.

Here, the aim is to investigate protein conformational dynamics during sub-picosecond charge separation using time-resolved serial femtosecond crystallography (TR-SFX).

5.3 Results and Discussion

Studies of time-resolved infrared spectroscopy on the RC special pair have shown low-frequency oscillations of 300 fs period that diminishes over picoseconds and are influenced by mutations in the special pair (SP) surroundings [192,193]. If the SP vibration pattern is affected by changes in the nearby protein structure, these vibrations must be linked to the vibrations of the larger protein [194]. However, some studies suggest that the observed vibrations are not due to the SP or the surrounding protein, but rather originate within the bacteriochlorophyll (BCh) molecules themselves [195]. Bacteriochlorophyll molecules contain a porphyrin-like structure made up of pyrrole rings. The specific assignment of these vibrations would be localized on deformations in pyrrole ring I and could result

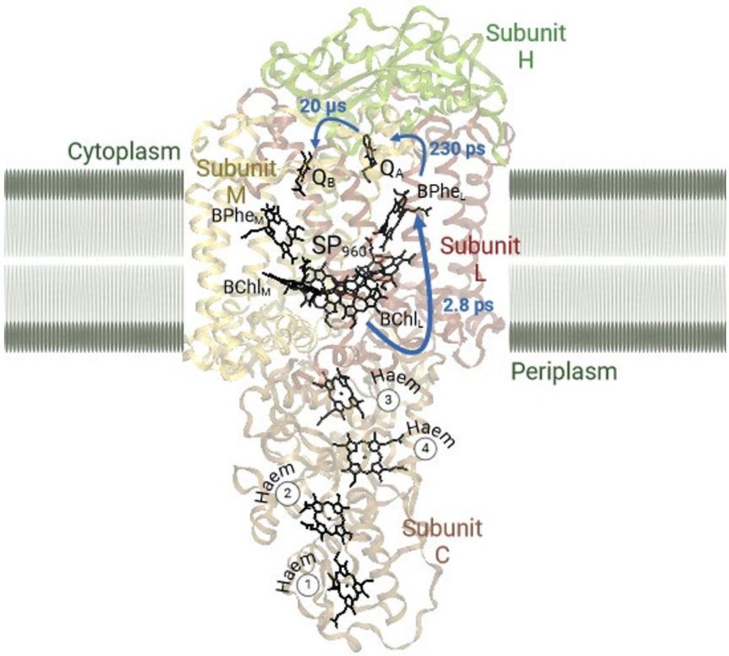


Figure 5.1: Reaction center from *Blastochloris viridis* A representation of the RC*vir* is shown with the three first electron transfer events are marked by blue arrows and the time-points of each one shown. The co-factors are colored black, the H subunit in green, the M subunit yellow, the L subunit dark red and the C subunit is brown.

in precise changes on molecular shape or energy levels. Alternatively, the vibrations might involve a collective deformation across all pyrrole rings in the BCh molecule, which could be communicated to the surrounding molecule. This would suggest that the entire core of the molecule participates in the motion.

Several other spectroscopic time-resolved studies strongly support the idea that the initial charge separation reactions are controlled by protein conformational changes [65, 196, 197]. These studies tracked how tryptophan absorbance changed over time during charge separation in the wild-type and mutant protein. The data revealed that the protein undergoes conformational changes during this stage, which could control or influence the charge-separation reaction. The interplay between protein dynamics and water interactions governs reaction rates, demonstrating that the traditional equilibrium approach to reaction energetics is insufficient to explain rapid processes such as electron transfer, which depend heavily on time-sensitive dynamics [65].

TR-SFX study suggests that, following excitation, the electron density associated with the special pair of RC*vir* increases between 1 ps and 20 ps, indicating some initial structural or electronic change. After 300 ps, this change begins to decrease, potentially as the system relaxes back to its initial state [190]. While the data shows electron density changes over time, they are not definitive proof of a particular structural change or mechanism. The changes in electron density near the SP are linked to structural or dynamic changes within the protein, but more research is needed to confirm what these changes mean and how they impact the overall function of RC*vir*. Global conformational movements in the protein regions surrounding the L-branch co-factors were observed which closely matched the time scale of electron transfer reactions occurring between these co-factors [190].

The aim of this work is to expand on previous findings and explore whether structural observations could help validate and explain the signals detected by sub-picosecond vibrational spectroscopic studies. As it is shown in figure 5.2f, former studies revealed strong electron density difference in the vicinity of the SP. Notably, as it is shown in figure 5.2a-e, when the same analysis is performed on the sub-picosecond time-scale these features appear much weaker. These weaker features might be due to the fact that the occupancy exhibited by the photo-excited SP* state appeared considerably reduced. Another possibility is that these motions might need a short period of picoseconds to grow their amplitude. It must be noted that in this case, LCP grown micro-crystals were used as opposed to the earlier TR-SFX studies [175], which had detergent based RC*vir* crystals. This was due to the fact that a 0.15 Å resolution improvement was obtained when the LCP sample was used, thus it would potentially allow the detection of fainter movements.

There are two consequences derived from the practical change in sample preparation. The first one is that the morphology of the crystals differ, not only their size and shape but also the space groups are in fact different. The second consequence is that the injector used for the sample in this case was a high-viscosity micro-jet injector [198], as opposed to the GDVN micro-jet injector [199] previously used. The practical difference is that the diameter of the sample for the GDVN injector is 10 μm while the one used for this study results in a 75 μm diameter. Thus, the type of injector affects the optical properties of the micro-jet, with the high-viscosity injector producing a thicker jet that scatters light more. Additionally, LCP preparations often turn cloudy, complicating the estimation of how much light is scattered. The laser wavelength used to excite the samples was 960 nm and was the same for both experiments. The fluence (energy of the laser per unit area) was slightly reduced from 21 mJ/cm^2 [190] to 18 mJ/cm^2 in the current study. Even with this small reduction, the fluence remained theoretically far above the threshold for single-photon excitation under ideal conditions. Nonetheless it is a possibility that due to this difference, the photo-activation of the RC*vir* LCP crystal sample might have been lower. Regrettably, no method to quantify micro-jet based TR-SFX losses in experiments has been accepted so far [161, 200, 201]. Regardless of the quality of the data, the results suggest that the movements of the BCh co-factors in the SP region were extremely small (low amplitude) during the sub-picosecond timescale. These motions were observed to increase in amplitude over the course of a few picoseconds [190].

TR-SFX studies have revealed anisotropic global conformational changes in some rhodopsins [200, 202, 203], myoglobin [204], and photoactive yellow protein [205, 206]. RC*vir* exhibited ultra-fast conformational changes

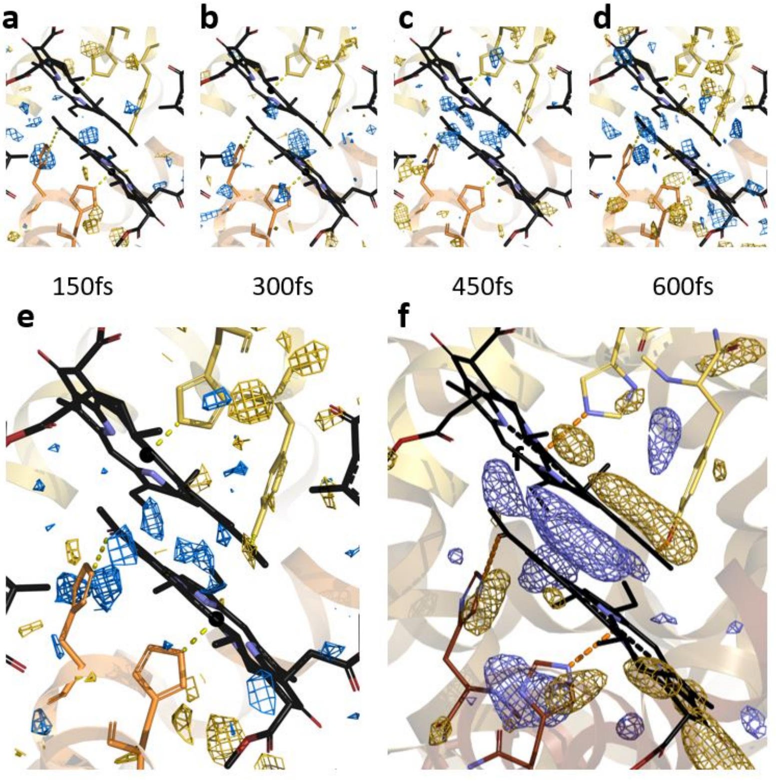


Figure 5.2: Electron density changes in RC*vir*'s special pair Isomorphous difference Fourier electron density maps F_{OBS}(light)-F_{OBS}(dark) are shown. For the individual time-points a) 150 fs, b) 300 fs, c) 450 fs and d) 600 fs, all contoured at 2.8 σ. Principal SVD component analysis is shown for e) 50, 150, 300, 450, 600, 750 fs and 1 ps contoured at 2.8 σ along side f) 1, 5, 20, 300 ps and 8 μs (from previous work) contoured at 3.2 σ. Positive electron density is shown in blue, and negative electron density is shown in gold.

drawn by extreme multi-photon excitation conditions using TR-WAXS at an XFEL, supporting the 'protein quake' hypothesis of rapid energy dissipation through quake-like motions [207]. Conformational changes in protein crystals have been observed when an electric field of approximately 108 V/m is applied to a single crystal, as the field influences the charge distribution and prompts structural adjustments. Similarly, the redistribution of charges around the SP during excitation in RC*vir* is estimated to reach the same magnitude. Even after applying resampling techniques and conducting numerous structural refinements [159], no signs of global protein movements could be extracted beyond the minimal noise inherent to the data obtained (figure 5.3). Thus, no characteristic conformational changes can be seen on sub-picosecond time-scales for RC*vir*. While some molecular events, such as chemical bond photodisossiation [204] or the photoisomerisation of a double bond [202, 203],can create strong shock waves that propagate through a protein, the energy transition from SP to its excited state SP* does not seem to generate such an effect. It cannot be ruled out that there are effects derived from this process, there is a possibility that the change of sample delivery [208] or significant increase in data collection might show evidence of conformational changes. Another possibility is that the TR-SFX data may be complicated by a mixture of single-photon and multi-photon excited states, indicating that using lower pump-laser fluence to limit excitation to the single-photon domain [204] could enhance the observed difference electron density.

Because the scaling of the data is an important part of the data analysis of TR-crystallography, machine learning algorithms [191] might aid in the improvement on the signal to noise ratios shown in the difference map obtained (figure 5.2).

Although there is room for further refinement in data analysis, this TR-SFX study currently presents a weak difference signal that aligns with previous findings [190] but lacks the statistical significance required to make a definitive observation.

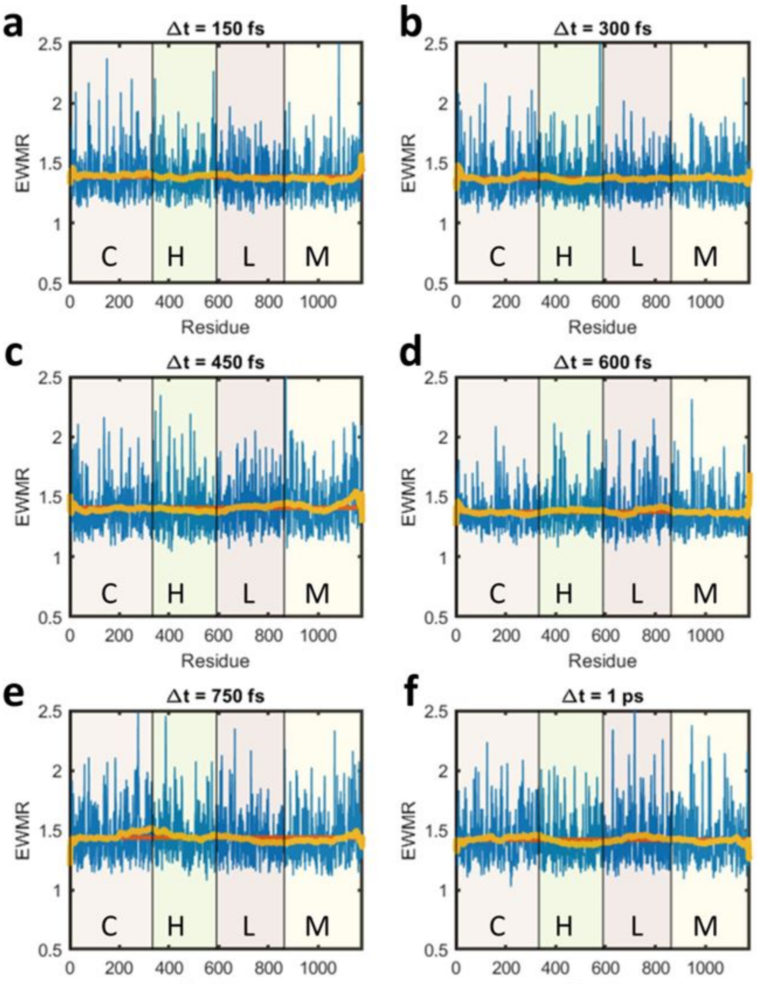


Figure 5.3: Error analysis of Cα displacements The analysis was performed for a) 150 fs, b) 300 fs, c) 450 fs, d) 600 fs, e) 750 fs and f) 1 ps after photo-activation. C, H, L and M describe the different subunits.

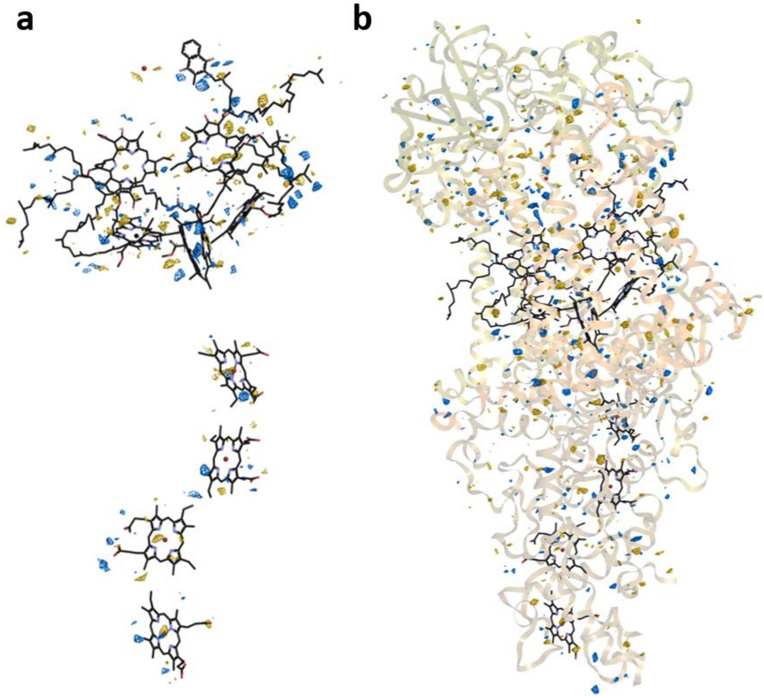


Figure 5.4: Electron difference maps F_{obs}(light)-F_{obs}(dark) obtained with SVD analysis for a) the co-factors with a 3.5 σ and b) in the overall molecule with a 2.5 σ. The positive difference is depicted in blue while the negative one is gold colored.

5.4 Conclusions

In this work a positive electron density difference consistent with prior observations was obtained. Unfortunately, the quality of the data obtained was greatly decreased due most likely to the change in the sample delivery system or the sample itself. Thus, no evidence of large-scale protein motion during sub-picosecond charge separation could be seen. The data suggests that in the case there are nuclear motions, these are localized and minimal.

Chapter 6

Paper IV

Multi-photon induced conformational changes in a photosynthetic reaction centre

6.1 Overview

Time-resolved X-ray solution scattering (TR-XSS) combined with time-resolved serial femtosecond crystallography (TR-SFX) at a free-electron laser were applied to study structural changes in RC isolated from *Blastochloris viridis*. Different pump-laser fluence conditions were applied after 5ps photo-excitation using an 800 nm pump. The observed difference in X-ray scattering, indicative of protein conformational changes, varies with the square root of the energy deposited by multi-photon absorption. TR-SFX results support the theory that multi-photon-induced conformational changes create thermal driven expansion of the protein structure surrounding the co-factors involved in light absorption. These global motions do not occur in single-photon absorption, thus are not considered essential for the initial charge separation events within RC.

6.2 Background

Conserved RCs catalyze the primary charge separation steps in photosynthesis in both plants and photosynthetic bacteria. In the case of *Blastochloris viridis*, its RC (RC_{vir}) is the integral membrane protein that makes this process possible. This complex supports key co-factors, including bacteriochlorophylls (P_{960} and BCh), bacteriopheophytins (BPh), menaquinones (Q_A and Q_B), and a non-heme iron. The special pair of Bchls (P_{960}) absorb at 960 nm wavelength, while the two auxiliary bacteriochlorophylls (BCh) do so near 800 nm.

Upon photo-oxidation of P_{960}, charge separation occurs as an electron is transferred to BPh and further to Q_A and Q_B. This process, coupled with electron transfer to and from the tetra-haem subunit, leads to the release of quinol into the membrane. For each photon absorbed, two protons are transported across the membrane.

It takes 3.2 ps for the electron to be transferred from the P_{960} to the BPh. The surrounding co-factors can also absorb and transfer energy towards the P_{960} on a timescale of 290 fs. It has been debated for years whether RCs undergo rapid conformational changes upon light activation or not. This debate has been going on for so long because these changes occur on extremely short time scales, making them challenging to be observed. There are also limitations associated to the experimental techniques to allow to see such quick changes and also, the intricate structure of RCs hinders the isolation and study of such specific conformational changes. Furthermore, it is highly debated whether the observed changes are functionally relevant or side effects of light activation. Understanding this issue is key for elucidating the mechanism used to absorb and transfer energy into the biosphere, since rapid conformational shifts can potentially be driving the high efficiency of the process.

Arnlund et al used TR-XSS and captured ultra-fast structural changes in RC*vir* within a few picoseconds after multi-photon excitation using an 800 nm pump [207]. Although ultra-fast global motions have been detected in solution, the use of high laser fluence at 800 nm led to around 800 photons being absorbed per RC_{vir} molecule, making it difficult to assess the biological relevance of these structural changes. The half-rise time for these changes was approximately 1.4 ps, with maximum amplitude observed at 7 ps, similar to the charge separation rate which is 3.2 ps. However, TR-SFX using a 960 nm pump revealed only minor initial changes, with larger structural shifts occurring later at 300 ps [190].

To investigate further the persistence of global conformational changes induced at 800 nm, data was examined when the laser fluence is reduced to the single-photon level. The amplitude of these changes varied with the square root of the energy absorbed, seen as sample heating in TR-XSS data. Power-titration experiments with time-resolved infrared spectroscopy confirmed coherent multi-photon excitation at 800 nm, but no similar pattern was observed with a 960 nm pump [190]. TR-SFX data showed global structural changes at 5 ps across two pump-laser fluence values, supporting previous interpretations based on molecular dynamics simulations [207]. These multi-photon-induced changes at 800 nm differ from those seen with 960 nm excitation, contributing to the debate on their relevance under multi-photon conditions [200, 201, 204].

6.3 Results and Discussion

6.3.1 Heating induced by multi-photon absorption measured by time-resolved X-ray solution scattering

A pump-laser spot size of 375 μm diameter was used to record TR-SFX data with an energy distribution strongest at the center and gradually decreasing to 13.5 $\%$ of the maximum intensity at the radius of the spot (I$/0=$e^2). The pulse energies used were 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 μJ per pulse to excite the sample. In figure 6.1a, TR-SFX data obtained at high laser fluence can be seen. All the most important features of the previously reported data for a time-point of 5 ps were observed under strong multi photon activation conditions.

Data was recorded by TR-WAXS two minutes for each of the laser fluences, yet instead of collecting data for all two minutes at one energy level and moving to the next, the eleven energy levels were cycled repeatedly 3 times. Thus, different pulse energies were switched multiple times during the experiment. More data was collected for the laser fluence values lower than 40 μJ/pulse since the signal is weaker. Figure 6.1a shows the XSS data time delay ΔS(q,Δt) for all the runs with different time delays. It can be seen in the plot that big fluctuations happen in the different runs, some of which exhibit X-ray scatter amplitudes that are larger, even though all of them share the same amount of energy.

The first step in the analysis was to extract the signal that is characteristic of a sample being heated. In order to do this, a singular value

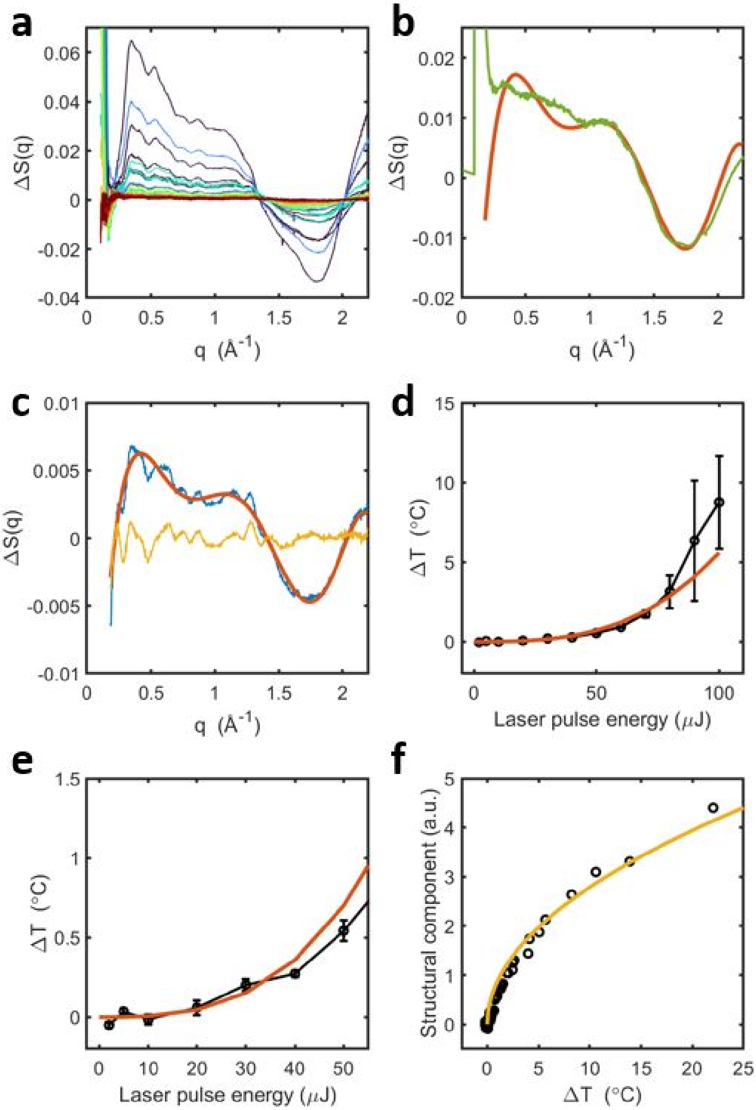


Figure 6.1: TR-XSS of RC$_{vir}$ a) Difference XSS data for all time-delays with different pulse energies. The thermal heating (red) was extracted from an SVD analysis from a sub set of spectra and fit to an eighth order polynomial expression in b). The (orange) XSS is a linear decomposition of TR-XSS data from a 800 nm pump laser RC*vir* (previous work) [207].

Figure 6.1: (continued)c) Linear decomposition of a TR-XXS difference curve (blue) for 5 ps, along with the heating (orange) and the component expressing oscillations characteristic of protein structural changes (yellow). In d) Temperature jump at 800 nm as a function of the laser pulse energy for RC*vir*. e) data from d) zoomed in. f) an amplitude of protein motions versus the absorbed energy (circles) plot, a linear decomposition of measurements from c).

decomposition (SVD) was used to calculate the principal component using a data subset that showed similar amount of heating. After this, an order eight polynomial fit was applied to the principal component. For the time delay Δt=5 ps (shown in figure 6.1b, depicted in red) the heated component extracted proved similar to the same component extracted in a similar previous analysis [207] (showed in figure 6.1b, green line for comparison). Subsequently, the polynomial fit for heating (figure 6.1c) was subtracted from a measurement that is representative (shown in figure 6.1c in blue). As it can be seen, the result is (shown in figure 6.1c colored in yellow mustard) showing oscillations that are characteristic of a protein conformational changes [207,209] The pattern obtained is similar to such patterns obtained in previous work [207] where these oscillations were related to ultra-fast protein conformational changes when TR-XSS was used.

In previous work, a heating calibration was established to be able to quantify the energy that was conveyed into the sample by the 800 nm pump laser [207]. In order to do this, the heating XSS spectrum obtained in figure 6.1b was matched to the heating effect observed in the basis spectrum at later time delays. This latter time spectrum was then scaled using controlled heating data obtained from samples exposed to a regulated air stream. Multi-photon induced temperature jumps obtained for all the energy values of the pump laser pulse can be seen in figure 6.1d. To emphasize the power dependence when there is lower laser fluence, the laser pulse energy increase vs time is shown in figure 6.1e.

The absorption cross section at 800 nm is a measure of the probability that a photon will be absorbed by a molecule at this wavelength. Given that this measurement is constant, as we increase the fluence the laser-induced temperature jump would be expected to maintain a linear relationship. In other words, doubling the fluence would result in roughly double the temperature increase, and tripling the fluence would triple the temperature increase, and so on. It can be seen that the dependence shown in 6.1e is not linear. In fact, the best fit to this data was provided by a cubic power law expressed as:

$$\Delta T = \sigma_1 F + \sigma_2 F^2 + \alpha_3 F^3 \tag{6.1}$$

where ΔT is the increase in temperature, σ_1 and σ_3 are constants (cross sections in this case) and $\times F$ is the laser fluence (energy). This cubic dependence suggests that higher fluence leads to disproportionate temperature increase, implying non-linear heating behavior in the system.

The values of the cross sections σ_1, σ_2 and σ_3 were adjusted to find the best match between the experimental data and the theoretical model by minimizing the difference between them. This process is called "least-squares fit" and it involves calculating the sum of the squared errors between the values predicted by the model and the observed experimental values, and then finding the values of σ_1 and σ_3 that make this sum as small as possible.

Based on the data, the absorption cross section at 800 nm depends on the pump laser fluence. This is typical of multi-photon absorption, where several photons are absorbed simultaneously to excite the system. The data follows a third-power dependence on the laser fluence, which would mean that three photons are involved simultaneously in the excitation process. If three photons were absorbed at 800 nm, the energy absorbed by the system for excitation would be 267 nm (because 800 nm divided by 3 = 267 nm).

6.3.2 Micro-jet scattering loss quantification

If a 30 mg/mL of RC*vir* is subjected to a temperature jump of 7 ± 1 °C then the number of photons absorbed per molecule will be $800{\pm}110$ [207]. Taking into account the ratio from this scale, by measuring the resulting temperature jumps we can estimate how many photons are absorbed per molecule at different laser pulse energies:

- $\Delta_T = 8.7 \pm 2.9$ at 100 μJ/pulse would be 1000 ± 330 photons absorbed per molecule.

- When $\Delta_T = 0.54 \pm 0.06$ using 50 μJ per pulse, the absorbed photons per molecule would be 61 ± 7.

- With a measured $\Delta_T = 0.06 \pm 0.01$ if 20 μJ/pulse is used the absorbed energy per molecule would be 7 ± 1.

In the low pump laser fluence domain, if a linear extrapolation to a power of 1.7 μJ is performed, each RC*vir* molecule would absorb 0.6 photons. This corresponds to the lowest pump laser fluence used, at low enough laser energy that absorption scales linearly with intensity so that it can be assumed that each photon interacts independently with the molecule.

To estimate the number of photons absorbed per RC*vir* molecule assuming single photon absorption, the following equation can be used:

$$N_{\text{photons}} = F \cdot \frac{\sigma_{800}}{h\nu} = \frac{F}{h\nu} \cdot \frac{\varepsilon_{800}\,\text{nm}}{N_A \cdot 1000} \tag{6.2}$$

where each variable is:

- F: The fluence of the laser pulse. This quantifies the energy delivered per unit area by the laser, averaged over the full width at half maximum (FWHM) of the laser spot.

- σ_{800}: The absorption cross-section of RC*vir* in its resting state at 800 nm (the wavelength of the laser).

- $h\nu$: The energy per photon, calculated as the product of Planck's constant (h) and the frequency of light (ν).

- N_A is Avogadro's number.

- $\varepsilon_{800\,\text{nm}}$ the molar extinction coefficient at 800 nm.

RC*vir* exhibits a high molar extinction coefficient [210], $\varepsilon_{800\,\text{nm}} \approx 180,000 \pm 20,000\,\text{M}^{-1} \cdot \text{cm}^{-1}$, which reflects the molecule's strong light absorption at 800 nm. This coefficient is related to the absorption cross-section, since both quantify the ability of a molecule to absorb light but in different units.

For a Gaussian beam with energy per pulse $E = 1.7\,\mu$J and a $1/e^2$ spot diameter of $375\,\mu$m, we find that the fluence averaged over the full width spot diameter:

$$F_{\mathsf{FWHM}} = 2.2\,\mathsf{mJ/cm}^2$$

Given that no energy loss, perfect overlap between the beam with focal spot and perfect alignment of the XFEL beam with the maximum of the laser spot, then 4.1 photons would be absorbed per RC*vir* molecule using equation 6.2. According to the data measured and previous calculations performed, only 0.6 photons are absorbed per RC*vir* molecule. This observed value is about 15% of the predicted absorption under ideal conditions. If the prediction is adjusted by using the value of fluence averaged across the FWHM of the focal spot, which would be a more realistic approximation of the energy distribution, the predicted number of absorbed photons decreases to 3.0 photons per molecule with the experimental absorption of 0.6 photons per molecule representing about 20% of the calculated absorption. The data suggests then that roughly 80-85% of the pump laser energy is lost due to scattering from the micro-jet. Even though the micro-jet appears to be smooth in snapshots, it is in reality in constant fluctuations due to turbulent flow, surface tension effects, scattering effects, and energy loss.

A debate has emerged within the scientific community related to whether TR-SFX studies have been performed under conditions of multi-photon excitation. Most studies, with one exception [204], consider that the experiments have indeed operated in a multi-photon regime [200,203, 211]. This is because the dimensionless product for the TR-SFX experiments is $F \cdot \sigma_{800}/h\nu > 1$, and this implies that the energy conveyed by the incoming pump laser pulse exceeds the single-photon energy threshold [161]. If this condition is assumed, two or more photons are being absorbed simultaneously by the molecules, leading to multi-photon excitation. Such behavior can significantly alter the dynamics of a system, affecting the timescales of the reactions and the interpretation of the data.

The main issue regarding this debate is that the experiments with micro-jets do not happen under ideal conditions which often assumed in theoretical models. Factors such as energy loss during the interaction between the pump laser and the micro-jet complicate the understanding of the absorption process.

6.3.3 Power dependence changes in protein structure

In figure 6.1c (yellow/mustard line), oscillatory features can be seen in Δ_S(q)Δ_t=5 ps after the effect of heat was removed. This fingerprint is characteristic of structural changes when measured with TR-XSS. In figure 6.1c, a decomposition is shown of the heating component and the conformational changes. This decomposition enabled to follow the amplitude of the structural changes in the protein as a function of the energy that was conveyed by the multi-photon process. In figure 6.1f, it can be seen that the amplitude of the perturbation on the protein spectra varies with the square root of the energy absorbed. This absorption matches the behavior of an harmonic oscillator, a system that oscillates back and forth in a predictable way where the total energy increases with the square of the oscillation's amplitude. If TR-XSS data is taken into account, the structural changes of RC*vir* do not continue when the energy is low (less than one photon absorbed per molecule). It is known for a fact that photosynthesis works at a low energy level, where each molecule absorbs only one photon at a time. Thus, the TR-XSS technique isn't sensitive enough to pick any changes in the protein structure under these conditions.

6.3.4 TR-SFX study using 800nm pump

A 800 nm pump laser was used with the aim to excite the co-factors near $\mathrm{P}960$, using a single time delay of 5 ps and under multi-photon excitation conditions. The setup used was identical to the setup used for TR-SSX data collection. Two pump lasers were used with intensities of 50μJ/pulse and 20 μJ/pulse. These intensities correspond to 61 ± 7 and 7 ± 1 photons absorbed per molecule, respectively. This reasoning assumes that energy losses in the micro-jet containing micro-crystals are comparable to those in a micro-jet with solubilized RC*vir* samples, though this assumption may not necessarily be the case due to the nature of the sample. Data extracted from TR-SFX experiments using a 960 nm laser pump was processed and a 3.4 Å resolution structure was obtained [190]. The isomorphous F_{obs}(light)-F_{obs}(dark) difference Fourier electron density maps obtained can be seen in figure 6.2a-c. Even though the quality is not up to the best SFX standard, at lower laser pump fluence some difference density peaks can be seen over 3.4 σ, but it is hard to interpret it as protein structural changes. To judge whether the changes are significant, a look is taken to the σ value, which represents the root mean square (RMS) electron density of the unit cell, which is the smallest repeating unit in the crystal structure. A peak above 3.5 σ means it stands out from the background noise by a certain statistical threshold, indicating that it's likely a real signal. Thus, while the low energy laser pulse creates some noticeable electron density changes, these are subtle and no connection can be made to any particular shifts or alterations in the structure of RC*vir*.

When the laser energy was 50 μJ/pulse, in figure 6.2a and b, there are some positive and negative electron density shown that are associated with the co-factors. Since we are in the high energy pump laser energy this means multi-photon excited sample. It was previously shown by MD simulations using XSS curves [207] that the data is consistent with laser induced heating of the trans-membrane αhelices of RC*vir* owing to the heating of the co-factors. TR-SFX data agrees with this statement given that both bacteriochlorophylls co-factors show features in the difference density map that hints an displacement outwards from the two fold pseudo symmetry axis of RC*vir*'s structure 6.2.

A means to estimate coordinate errors through a resampling protocol of SFX data was applied to asses the significance of the movements observed relative to these errors [159], This protocol was applied to evaluate the movements observed in figure 6.2a to c with the time delay of 5 ps. In figure 6.2d, where 20 μJ/pulse was applied, no marked difference was observed (figure 6.2d). This is not the case when 50 μJ/pulse was applied to the sample (figure 6.2e), where structural perturbations are observed around the bacteriochlorophyll co-factors. If the same is done for the 960 nm laser pump TR-SFX data, few perturbations can be seen (figure 6.2f) for the same time delay. When the time delay is 300 ps (image 6.2g), some perturbations in the L subunit can be observed, which is the active branch.

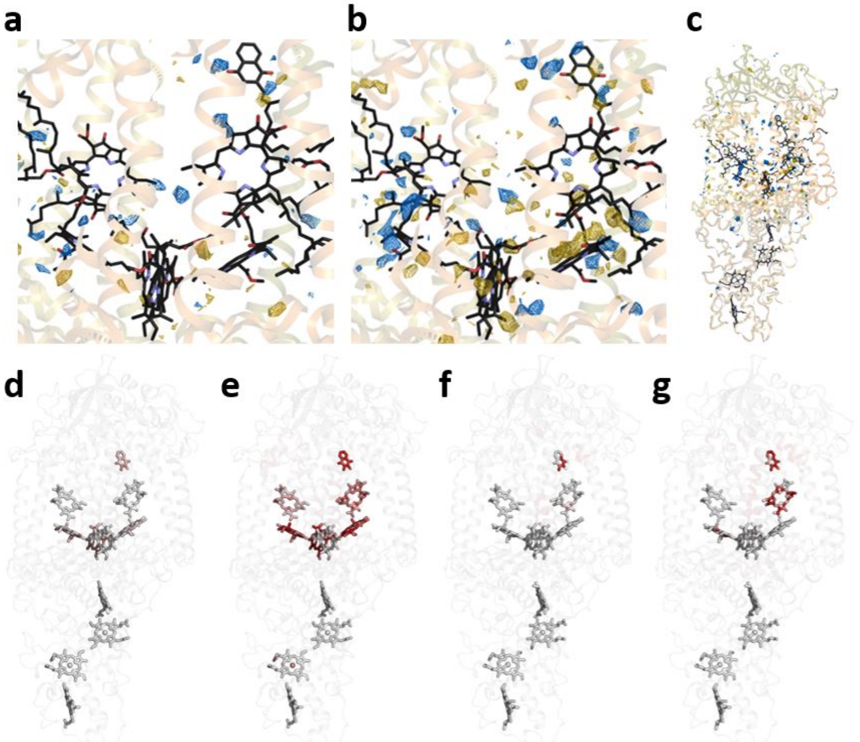


Figure 6.2: TR-SFX measurements of RCvir at 800 nm A 3.5 Å resolution difference Fourier electron density map for pump laser fluence of a) 20 μJ/pulse and b) 50 μJ/pulse. Both were focused at $\frac{1}{e^2}$ of 375 μm and a time delay of 5 ps. An overview of the whole molecule difference electron density map is shown in c). For all images, the contour is set to $\pm 3.5\sigma$ (root mean square electron density within the unit cell), where positive difference electron density is colored blue and negative difference electron density is shown in gold. These were estimated by calculating a set of 100 structures from resampled serial crystallography data [159,190]. The color of the co-factors reflect relative significance of movements induced by the laser from white (not significant) to red (more significant). (d) With a laser energy of 20 μJ/pulse at 800 and a time delay of 5 ps very few perturbations appear. With 50 μJ/pulse (e), the co-factors that are embedded in the membrane show thermally induced conformational changes. When the wavelength of the laser is set to 960 nm, and focused at $\frac{1}{e^2}$ of 323 μm and an energy of 11.8 μJ per pulse, (f) and (g) are obtained for 5 and 300 ps time delays, respectively.

RC*vir* is a big protein, thus a long distance overview of an isomorphous electron density map and establishing clear motions is a daunting task. As it can be seen in figure 6.2c, difference electron density peaks do not arrange to show specific motions. Regardless, resampling resulted a useful tool to allow identification of movements in the protein [159]. This was the case even where these movements are so small that they approach the errors in the data coordinates. When the analysis was performed with focus on coordinate changes of Cα atoms, for 20 μJ/pulse at 800 nm some regions stand out. This can be seen in figure 6.3a and c, where the Cα coordinate errors are highlighted. At 50 μJ/pulse, a backbone motion associated with the M subunit appears, which can be observed in figure 6.3b and d.

If a comparison is made between these structural perturbations, at 50 μJ/pulse it appears to be less symmetric when the fitting is performed against TR-XSS data [207]. A possible explanation is that Bch from the M subunit might for some reason be better suited to absorb at 800 nm compared to the one located in the L subunit, when the energy is high and non linear absorption happens. When a comparison is made of the structural response at 960 nm [190], at 5 ps time delay there is a slight perturbation in the H and L subunit of RC*vir* structure which can be seen in figure 6.3e and g. When the time delay is of 300 ps, figure6.3f and g, the structural response increases. It must be noted that the structural perturbations obtained for the Cα analysis varies for each wavelength used. When 960 nm laser pulse was used, the changes are associated with the L branch, where the movement of electrons in the charge separation process happens. In contrast, when 800 nm pulse is used, the changes observed are associated to the M branch and are caused by multi-photon excitation of the Bch co-factors.

6.4 Conclusions

TR-XSS data was obtained at two laser fluences, expanding on previous work [207] where multi-photon excitation was used to generate RC*vir* structural changes. At 800 nm, multi-photon absorption happens as a

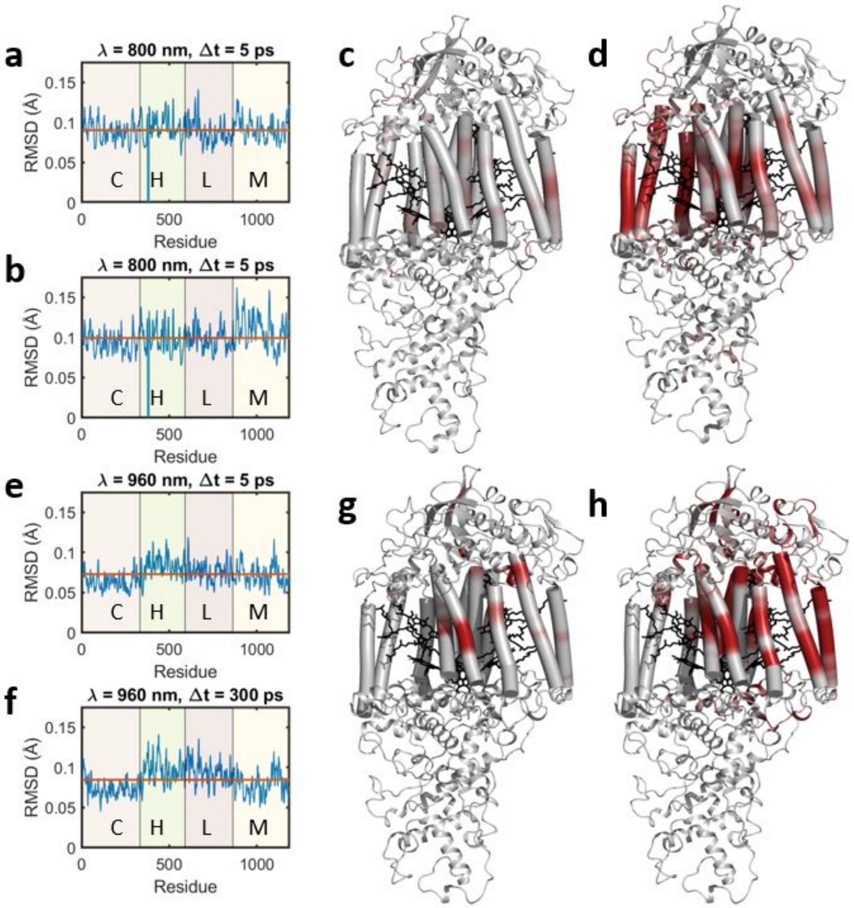


Figure 6.3: RCvir structural change response to different laser fluence RMSD (root mean square deviation) of Cα atoms after 5 ps delay 800 nm wavelength pump photo-excitation with (a) 20 μJ pulse and (b) 50 μJ. Cα atom perturbations relative to the mean of RMSD values colored in red, a different representation of data presented in (a and b). Structural changes induced by different laser power RMSD plots when 960 nm after (e) 5 ps and 11.8 μJ/ pulse and (f) 300 ps. (g and h) Cα atom perturbations relative to RMSD values for (e) and (f) respectively with higher perturbations colored in red.

cubic function of the laser power used (figure 6.1). Time resolved spectroscopic data shows that there is a dependence upon the power used at 800 nm wavelength, with a marked effect upon the Bchl co-factors. The clear advantage in the setup used is the possibility to measure the heating induced by the laser. This heat was compared to the linear absorption cross-section, to measure of how much energy is absorbed by the sample per unit area. The expression given in equation 6.2, which has been used to predict how much energy is absorbed by a sample, theoretically does not reflect reality in an accurate way. In fact, the theoretical prediction was found to be much higher than the measured absorption being that it overestimated the absorbed energy by a factor greater than five. This suggests that previous theoretical assumptions did not account for losses and practical imperfections in experiments that affect energy absorption when multiple photons are absorbed simultaneously.

When the laser is set to 800 nm, TR-SFX data agrees with previous TR-XSS data showing co-factor related structural changes [207]. There is an interesting asymmetry effect in the Cα perturbation observed when using 800 or 960 nm wavelength laser. An 800 nm laser pump activates changes in the M subunit due to heating of the co-factors triggered by multi-photon absorption, while 960 nm activation causes slower changes related to the L subunit and the electron transfer pathway [190]. It must be noted that no specific conformational change could be characterized for 800 nm wavelength at 20 μJ/ pulse (7 ± 1 photons absorbed per molecule) for RC*vir* in solution. Thus, TR-XSS might lack the necessary resolution for multi-photon induced conformational changes unless more photons are absorbed in the reaction. By contrast, previous results with 960 nm TR-SFX studies show that the time-scale of the motions observed along with the pathway through the L branch suggests an ultra fast response to primary charge separation reaction [190].

Chapter 7

Conclusions and future perspectives

This thesis has focused on structural biology using X-ray crystallography, and in particular on the photosynthetic reaction center of *Blastochloris viridis*. In **PAPER 1** a novel crystallization method utilizing LCP and detergent-based crystals for seeding has successfully produced high-quality crystals, resulting in improved resolution serial femtosecond X-ray crystallography structure of RC_{vir}. This advancement will enable time-resolved X-ray diffraction studies to be pursued that study movements of the mobile quinone after one and two light pulses. This SFX structure also identified a stable second quinone binding site, corroborated by multiple crystallization publications. In **PAPER 2** challenges related to crystallization, storage, transportation, and data collection when using fixed-target serial crystallography at synchrotron radiation sources were addressed by a 3D-printed chip platform. This innovation holds the potential to become a widely-used tool in structural biology research. This tool has not yet been applied to time resolved X-ray diffraction studies, but it should result in valuable tool for reducing the sample volumes when aiming to recover molecular movies of protein structural changes upon light activation. In **PAPER 3**, a time-resolved X-ray diffraction study of the *Bl. viridis* reaction center was performed on a sub-picosecond time-scale, which is prior to the first light-driven electron-transfer event of photosynthesis. Weak positive difference electron density was observed between the bacteriochlorophylls of the special pair, and this was consistent with prior TR-SFX studies on slower time-scales. However, changes in the sample delivery system or sample composition may have lowered the crystallographic occupancy of the photo-excited species within the LCP micro-crystals developed in this thesis and used in this study. If true, the lower occupancy would have limited the detection of large-scale protein motions on a sub-picosecond time-scale and prior to charge separation. Alternatively, if nuclear motions are present, they must be of very low-amplitude and highly localized. An extension of this work is to implement data processing tools based upon machine learning algorithms, and this may allow the detection of slight protein motions with higher confidence and potentially distinguish interpretable features from the noise inherent in these maps. In particular, these new tools focus on the steps of scaling and merging serial X-ray crystallography data, and this is a critical step when the difference electron density signal is weak. Another possibility may be to collect an order of magnitude more TR-SFX data so as to improve the sensitivity of the overall experiment, and to utilize even smaller crystals. The use of the fixed-target platform may also hold some advantages over a LCP micro-jet, but this should be proven using synchrotron radiation sources prior to additional studies using XFEL radiation. In **PAPER 4**, TR-XSS data collected at varying pump-laser fluences expanded on previous research. At 800 nm, multi-photon absorption significantly affected the bacteriochlorophyll co-factors, with distinct responses in the M and L subunits that were very distinct from the motions visualized when using a 960 nm pump laser. Theoretical models of the energy absorbed were found to overestimate the absorbed energy measured from the heating observed in TR-XSS data, suggesting a need for more sophisticated models when discussing the influence of multi-photon absorption in time-resolved X-ray diffraction experiments. Future efforts should focus on developing more accurate theoretical frameworks that incorporate practical losses and experimental imperfections. These studies collectively advance the understanding of the structural and functional dynamics of the *Blastochloris viridis* reaction center, offering contributions to photosynthetic research and protein X-ray crystallography.

Bibliography

[1] Y. Bromberg and B. Rost, "Correlating protein function and stability through the analysis of single amino acid substitutions," *BMC bioinformatics*, vol. 10, pp. 1–9, 2009.

[2] P. C. Ng and S. Henikoff, "Predicting the effects of amino acid substitutions on protein function," *Annu. Rev. Genomics Hum. Genet.*, vol. 7, no. 1, pp. 61–80, 2006.

[3] J. F. Atkins and R. Gesteland, "The 22nd amino acid," *Science*, vol. 296, no. 5572, pp. 1409–1410, 2002.

[4] P. J. Quinn, D. Chapman, and A. D. Keith, "The dynamics of membrane structur," *Critical Reviews in Biochemistry*, vol. 8, no. 1, pp. 1–117, 1980.

[5] G. Vereb, J. Szöllősi, J. Matko, P. Nagy, T. Farkas, L. Vigh, L. Matyus, T. Waldmann, and S. Damjanovich, "Dynamic, yet structured: the cell membrane three decades after the singer–nicolson model," *Proceedings of the national academy of sciences*, vol. 100, no. 14, pp. 8053–8058, 2003.

[6] G. L. Nicolson and G. Ferreira de Mattos, "A brief introduction to some aspects of the fluid–mosaic model of cell membrane structure and its importance in membrane lipid replacement," *Membranes*, vol. 11, no. 12, p. 947, 2021.

[7] P. Khakbaz and J. B. Klauda, "Probing the importance of lipid diversity in cell membranes via molecular simulation," *Chemistry and physics of lipids*, vol. 192, pp. 12–22, 2015.

[8] B. Mostofian, T. Zhuang, X. Cheng, and J. D. Nickels, "Branched-chain fatty acid content modulates structure, fluidity, and phase in model microbial cell membranes," *The Journal of Physical Chemistry B*, vol. 123, no. 27, pp. 5814–5821, 2019.

[9] V. Lantez, I. Nikolaidis, M. Rechenmann, T. Vernet, and M. Noirclerc-Savoye, "Rapid automated detergent screening for the solubilization and purification of membrane proteins and complexes," *Engineering in Life Sciences*, vol. 15, 2015.

[10] J. P. Overington, B. Al-Lazikani, and A. L. Hopkins, "How many drug targets are there?," *Nature reviews Drug discovery*, vol. 5, no. 12, pp. 993–996, 2006.

[11] Y. Arinaminpathy, E. Khurana, D. M. Engelman, and M. B. Gerstein, "Computational analysis of membrane proteins: the largest class of drug targets," *Drug discovery today*, vol. 14, no. 23-24, pp. 1130–1135, 2009.

[12] J. K. Kim, D. C. Jang, H. M. Kang, K. J. Nam, M. H. Lee, J. K. Na, and K. Y. Choi, "Effects of light intensity and electrical conductivity level on photosynthesis, growth and functional material contents of lactuca indica l.'sunhyang'in hydroponics," *Journal of Bio-Environment Control*, vol. 30, no. 1, pp. 1–9, 2021.

[13] H. Gest, "Photosynthetic and quasi-photosynthetic bacteria," *FEMS microbiology letters*, vol. 112, no. 1, pp. 1–5, 1993.

[14] M. Koblížek, "The purple phototrophic bacteria: Hunter, cn, daldal, f., thurnauer, mc, beatty, jt: Advances in photosynthesis and respiration. vol. 28. springer, dordrecht 2009, isbn: 978-1-4020-8814-8. 1014 pp.," 2009.

[15] U. Kutschera and R. Khanna, "The origin of chloroplasts: Constantin s. merezhkowsky (1855–1921) and symbiogenesis," *Journal of Plant Biochemistry and Biotechnology*, vol. 31, no. 1, pp. 178–184, 2022.

[16] Waseemuddin, "Genomic studies in bacteria , mitochondria and chloroplast in relation with endosymbiotic theory," in *International Journal of Scientific and Research Publications*, 2016.

[17] L. Gu, "Optimizing the electron transport chain to sustainably improve photosynthesis," *Plant Physiology*, vol. 193, no. 4, pp. 2398–2412, 2023.

[18] R. A. Zamora, M. López-Ortiz, M. Sales-Mateo, C. Hu, R. Croce, R. A. Maniyara, V. Pruneri, M. I. Giannotti, and P. Gorostiza, "Light-and redox-dependent force spectroscopy reveals that the interaction between plastocyanin and plant photosystem i is favored when one partner is ready for electron transfer," *ACS nano*, vol. 16, no. 9, pp. 15155–15164, 2022.

[19] S.-J. Kim, C.-H. Lee, A. Hope, and W. S. Chow, "Inhibition of photosystems i and ii and enhanced back flow of photosystem i electrons in cucumber leaf discs chilled in the light," *Plant and cell physiology*, vol. 42, no. 8, pp. 842–848, 2001.

[20] L. Nikkanen, D. Solymosi, M. Jokel, and Y. Allahverdiyeva, "Regulatory electron transport pathways of photosynthesis in cyanobacteria and microalgae: Recent advances and biotechnological prospects," *Physiologia Plantarum*, vol. 173, no. 2, pp. 514–525, 2021.

[21] J. Asplund-Samuelsson and E. P. Hudson, "Wide range of metabolic adaptations to the acquisition of the calvin cycle revealed by comparison of microbial genomes," *PLOS Computational Biology*, vol. 17, no. 2, p. e1008742, 2021.

[22] N. J. Claassens, G. Scarinci, A. Fischer, A. I. Flamholz, W. Newell, S. Frielingsdorf, O. Lenz, and A. Bar-Even, "Phosphoglycolate salvage in a chemolithoautotroph using the calvin cycle," *Proceedings of the National Academy of Sciences*, vol. 117, no. 36, pp. 22452–22461, 2020.

[23] B. Robert, "Spectroscopic properties of antenna complexes from purple bacteria," *The purple phototrophic bacteria*, pp. 199–212, 2008.

[24] W. D. Swingley, R. E. Blankenship, and J. Raymond, "Evolutionary relationships among purple photosynthetic bacteria and the origin of proteobacterial photosynthetic systems," in *The purple phototrophic bacteria*, pp. 17–29, Springer, 2009.

[25] M. Jones, “Bacterial photosynthesis,” *Photobiological sciences online, American Society for Photobiology. http://www. photobiology. info/Jones. html*, 2009.

[26] V. V. Terentyev, A. A. Khorobrykh, and V. V. Klimov, “Photooxidation of mn-bicarbonate complexes by reaction centers of purple bacteria as a possible stage in the evolutionary origin of the water-oxidizing complex of photosystem ii,” *Photosynthesis: New Approaches to the Molecular, Cellular, and Organismal Levels*, pp. 85–132, 2015.

[27] M. T. Madigan, W. M. Sattley, Y. Kimura, and Z.-Y. Wang-Otomo, “Calcium and the ecology of photosynthesis in purple sulfur bacteria,” *Environmental Microbiology*, vol. 26, no. 2, p. e16591, 2024.

[28] J.-L. Soret, “Analyse spectrale: Sur le spectre d’absorption du sang dans la partie violette et ultra-violette,” *Compt. Rend*, vol. 97, pp. 1269–1270, 1883.

[29] G. Peralta, M. Seth, and T. Ziegler, “Magnetic circular dichroism of porphyrins containing m= ca, ni, and zn. a computational study based on time-dependent density functional theory,” *Inorganic chemistry*, vol. 46, no. 22, pp. 9111–9125, 2007.

[30] J. Harada, T. Mizoguchi, Y. Tsukatani, M. Noguchi, and H. Tamiaki, “A seventh bacterial chlorophyll driving a large light-harvesting antenna,” *Scientific reports*, vol. 2, no. 1, p. 671, 2012.

[31] M. T. Madigan, S. M. Resnick, M. L. Kempher, A. C. Dohnalkova, S. Takaichi, Z.-Y. Wang-Otomo, A. Toyoda, K. Kurokawa, H. Mori, and Y. Tsukatani, “Blastochloris tepida, sp. nov., a thermophilic species of the bacteriochlorophyll b-containing genus blastochloris,” *Archives of Microbiology*, vol. 201, pp. 1351–1359, 2019.

[32] A. Hiraishi, “Transfer of the bacteriochlorophyll b-containing phototrophic bacteria rhodopseudomonas viridis and rhodopseudomonas sulfoviridis to the genus blastochloris gen. nov.,” *International journal of systematic bacteriology*, vol. 47, no. 1, pp. 217–219, 1997.

[33] G. Drews and P. Giesbrecht, “Rhodopseudomonas viridis, nov. spec., ein neu isoliertes, obligat phototrophes bakterium,” *Archiv für Mikrobiologie*, vol. 53, pp. 255–262, 1966.

[34] M. Konorty, V. Brumfeld, A. Vermeglio, N. Kahana, O. Medalia, and A. Minsky, “Photosynthetic system in blastochloris viridis revisited,” *Biochemistry*, vol. 48, no. 22, pp. 4753–4761, 2009.

[35] T. Nogi, I. Fathir, M. Kobayashi, T. Nozawa, and K. Miki, “Crystal structures of photosynthetic reaction center and high-potential iron-sulfur protein from thermochromatium tepidum: thermostability and electron transfer,” *Proceedings of the National Academy of Sciences*, vol. 97, no. 25, pp. 13561–13566, 2000.

[36] A. G. McEwan, “Photosynthetic electron transport and anaerobic metabolism in purple non-sulfur phototrophic bacteria,” *Antonie Van Leeuwenhoek*, vol. 66, pp. 151–164, 1994.

[37] D. C. Youvan and B. L. Marrs, “Molecular mechanisms of photosynthesis,” *Scientific American*, vol. 256, no. 6, pp. 42–49, 1987.

[38] L. A. Staehelin and C. J. Arntzen, *Photosynthesis III: photosynthetic membranes and light harvesting systems*, vol. 19. Springer Science & Business Media, 2013.

[39] J. Williams, L. Steiner, and G. Feher, "Primary structure of the reaction center from rhodopseudomonas sphaeroides," *Proteins: Structure, Function, and Bioinformatics*, vol. 1, no. 4, pp. 312–325, 1986.

[40] G. J. Palenik, W. P. Jensen, and I.-H. Suh, "The history of molecular structure determination viewed through the nobel prizes," *Journal of Chemical Education*, vol. 80, no. 7, p. 753, 2003.

[41] J. Deisenhofer, O. Epp, K. Miki, R. Huber, and H. Michel, "Structure of the protein subunits in the photosynthetic reaction centre of rhodopseudomonas viridis at 3å resolution," *Nature*, vol. 318, no. 6047, pp. 618–624, 1985.

[42] Y. Lu, C. Goodson, R. E. Blankenship, and M. L. Gross, "Primary and higher order structure of the reaction center from the purple phototrophic bacterium blastochloris viridis: A test for native mass spectrometry," *Journal of proteome research*, vol. 17, no. 4, pp. 1615–1623, 2018.

[43] S. Scheuring, J. Seguin, S. Marco, D. Lévy, B. Robert, and J.-L. Rigaud, "Nanodissection and high-resolution imaging of the rhodopseudomonas viridis photosynthetic core complex in native membranes by afm," *Proceedings of the National Academy of Sciences*, vol. 100, no. 4, pp. 1690–1693, 2003.

[44] P. Qian, C. A. Siebert, P. Wang, D. P. Canniffe, and C. N. Hunter, "Cryo-em structure of the blastochloris viridis lh1–rc complex at 2.9 å," *Nature*, vol. 556, no. 7700, pp. 203–208, 2018.

[45] M. Okamura, M. Paddock, M. Graige, and G. Feher, "Proton and electron transfer in bacterial reaction centers," *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, vol. 1458, no. 1, pp. 148–163, 2000.

[46] C.-H. Chang, D. Tiede, J. Tang, U. Smith, J. Norris, and M. Schiffer, "Structure of rhodopseudomonas sphaeroides r-26 reaction center," *FEBS letters*, vol. 205, no. 1, pp. 82–86, 1986.

[47] A. Osyczka, K. V. Nagashima, S. Sogabe, K. Miki, K. Shimada, and K. Matsuura, "Comparison of the binding sites for high-potential iron- sulfur protein and cytochrome c on the tetraheme cytochrome subunit bound to the bacterial photosynthetic reaction center," *Biochemistry*, vol. 38, no. 48, pp. 15779–15790, 1999.

[48] V. Shuvalov and A. Yakovlev, "Coupling of nuclear wavepacket motion and charge separation in bacterial reaction centers," *FEBS letters*, vol. 540, no. 1-3, pp. 26–34, 2003.

[49] R. Shopes and C. Wraight, "Charge recombination from the p+ q- a state in reaction centers from rhodopseudomonas viridis," *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, vol. 893, no. 3, pp. 409–425, 1987.

[50] J. M. Ortega and P. Mathis, "Effect of temperature on the kinetics of electron transfer from the tetraheme cytochrome to the primary donor in rhodopseudomonas viridis," *FEBS letters*, vol. 301, no. 1, pp. 45–48, 1992.

[51] Y. Y. Sham, I. Muegge, and A. Warshel, "Simulating proton translocations in proteins: probing proton transfer pathways in the rhodobacter sphaeroides reaction center," *Proteins: Structure, Function, and Bioinformatics*, vol. 36, no. 4, pp. 484–500, 1999.

[52] N. S. Ponomarenko, L. Li, A. R. Marino, V. Tereshko, A. Ostafin, J. A. Popova, E. J. Bylina, R. F. Ismagilov, and J. R. Norris Jr, "Structural and spectropotentiometric analysis of blastochloris viridis heterodimer mutant reaction center," *Biochimica et Biophysica Acta (BBA)-Biomembranes*, vol. 1788, no. 9, pp. 1822–1831, 2009.

[53] A. Rutherford and P. Faller, "Photosystem ii: evolutionary perspectives," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 358, no. 1429, pp. 245–253, 2003.

[54] J. Breton, J.-L. Martin, A. Migus, A. Antonetti, and A. Orszag, "Femtosecond spectroscopy of excitation energy transfer and initial charge separation in the reaction center of the photosynthetic bacterium rhodopseudomonas viridis," *Proceedings of the National Academy of Sciences*, vol. 83, no. 14, pp. 5121–5125, 1986.

[55] J.-L. Martin, J. Breton, A. Hoff, A. Migus, and A. Antonetti, "Femtosecond spectroscopy of electron transfer in the reaction center of the photosynthetic bacterium rhodopseudomonas sphaeroides r-26: Direct electron transfer from the dimeric bacteriochlorophyll primary donor to the bacteriopheophytin acceptor with a time constant of 2.8$\pm$0.2 psec," *Proceedings of the National Academy of Sciences*, vol. 83, no. 4, pp. 957–961, 1986.

[56] W. Zinth, M. Nuss, M. Franz, W. Kaiser, and H. Michel, "Femtosecond studies of the reaction center of rhodopseudomonas viridis: The very first dynamics of the electron-transfer processes," in *Antennas and Reaction Centers of Photosynthetic Bacteria: Structure, Interactions and Dynamics*, pp. 286–291, Springer, 1985.

[57] M. Leonova, T. Y. Fufina, L. Vasilieva, and V. Shuvalov, "Structure-function investigations of bacterial photosynthetic reaction centers," *Biochemistry (Moscow)*, vol. 76, pp. 1465–1483, 2011.

[58] V. C. Nashine, S. Hammes-Schiffer, and S. J. Benkovic, "Coupled motions in enzyme catalysis," *Current opinion in chemical biology*, vol. 14, no. 5, pp. 644–651, 2010.

[59] A. J. Hoff and J. Deisenhofer, "Photophysics of photosynthesis. structure and spectroscopy of reaction centers of purple bacteria," *Physics reports*, vol. 287, no. 1-2, pp. 1–247, 1997.

[60] A. Ishizaki, T. R. Calhoun, G. S. Schlau-Cohen, and G. R. Fleming, "Quantum coherence and its interplay with protein environments in photosynthetic electronic energy transfer," *Physical Chemistry Chemical Physics*, vol. 12, no. 27, pp. 7319–7337, 2010.

[61] M. Mahalingam, K. Martínez-Mayorga, M. F. Brown, and R. Vogel, "Two protonation switches control rhodopsin activation in membranes," *Proceedings of the National Academy of Sciences*, vol. 105, no. 46, pp. 17795–17800, 2008.

[62] K. Henzler-Wildman and D. Kern, "Dynamic personalities of proteins," *Nature*, vol. 450, no. 7172, pp. 964–972, 2007.

[63] H. Frauenfelder, G. Chen, J. Berendzen, P. W. Fenimore, H. Jansson, B. H. McMahon, I. R. Stroe, J. Swenson, and R. D. Young, "A unified model of protein dynamics," *Proceedings of the National Academy of Sciences*, vol. 106, no. 13, pp. 5129–5134, 2009.

[64] D. N. LeBard and D. V. Matyushov, "Protein–water electrostatics and principles of bioenergetics," *Physical Chemistry Chemical Physics*, vol. 12, no. 47, pp. 15335–15348, 2010.

[65] D. N. LeBard, D. R. Martin, S. Lin, N. W. Woodbury, and D. V. Matyushov, "Protein dynamics to optimize and control bacterial photosynthesis," *Chemical Science*, vol. 4, no. 11, pp. 4127–4136, 2013.

[66] S. Schmidt, T. Arlt, P. Hamm, H. Huber, T. Nägele, J. Wachtveitl, M. Meyer, H. Scheer, and W. Zinth, "Energetics of the primary electron transfer reaction revealed by ultrafast spectroscopy on modified bacterial reaction centers," *Chemical physics letters*, vol. 223, no. 1-2, pp. 116–120, 1994.

[67] Y. Fujihashi, M. Higashi, and A. Ishizaki, "Intramolecular vibrations complement the robustness of primary charge separation in a dimer model of the photosystem ii reaction center," *The Journal of Physical Chemistry Letters*, vol. 9, no. 17, pp. 4921–4929, 2018.

[68] A. Sirohiwal, F. Neese, and D. A. Pantazis, "Protein matrix control of reaction center excitation in photosystem ii," *Journal of the American Chemical Society*, vol. 142, no. 42, pp. 18174–18190, 2020.

[69] N. P. Pawlowicz, R. Van Grondelle, I. H. Van Stokkum, J. Breton, M. R. Jones, and M. L. Groot, "Identification of the first steps in charge separation in bacterial photosynthetic reaction centers of rhodobacter sphaeroides by ultrafast mid-infrared spectroscopy: electron transfer and protein dynamics," *Biophysical journal*, vol. 95, no. 3, pp. 1268–1284, 2008.

[70] S. Lin, P. R. Jaschke, H. Wang, M. Paddock, A. Tufts, J. P. Allen, F. I. Rosell, A. G. Mauk, N. W. Woodbury, and J. T. Beatty, "Electron transfer in the rhodobacter sphaeroides reaction center assembled with zinc bacteriochlorophyll," *Proceedings of the National Academy of Sciences*, vol. 106, no. 21, pp. 8537–8542, 2009.

[71] M. M. Gomari, N. Saraygord-Afshari, M. Farsimadan, N. Rostami, S. Aghamiri, and M. M. Farajollahi, "Opportunities and challenges of the tag-assisted protein purification techniques: Applications in the pharmaceutical industry," *Biotechnology advances*, vol. 45, p. 107653, 2020.

[72] M. Du, Z. Hou, L. Liu, Y. Xuan, X. Chen, L. Fan, Z. Li, and B. Xu, "1progress, applications, challenges and prospects of protein purification technology," *Frontiers in Bioengineering and Biotechnology*, vol. 10, p. 1028691, 2022.

[73] T. Baier, D. Kros, R. C. Feiner, K. J. Lauersen, K. M. Müller, and O. Kruse, "Engineered fusion proteins for efficient protein secretion and purification of a human growth factor from the green microalga chlamydomonas reinhardtii," *ACS synthetic biology*, vol. 7, no. 11, pp. 2547–2557, 2018.

[74] F. Mancia and J. Love, "High-throughput expression and purification of membrane proteins," *Journal of structural biology*, vol. 172, no. 1, pp. 85–93, 2010.

[75] M. H. Mohd Salleh, A. Fitri Peli, M. S. Ngalimat, and K. J. Sim, "A mini literature review on current advancements in protein purification techniques," in *Biology and Life Sciences Forum*, vol. 20, p. 12, MDPI, 2022.

[76] A. de Marco, E. Deuerling, A. Mogk, T. Tomoyasu, and B. Bukau, "Chaperone-based procedure to increase yields of soluble recombinant proteins produced in e. coli," *BMC biotechnology*, vol. 7, pp. 1–9, 2007.

[77] A. Stetsenko and A. Guskov, "An overview of the top ten detergents used for membrane protein crystallization," *Crystals*, vol. 7, no. 7, p. 197, 2017.

[78] L. H. Urner, F. Junge, F. Fiorentino, T. J. El-Baba, D. Shutin, G. Nölte, R. Haag, and C. V. Robinson, "Rationalizing the optimization of detergents for membrane protein purification," *Chemistry–A European Journal*, vol. 29, no. 30, p. e202300159, 2023.

[79] M. Woubshete, S. Cioccolo, and B. Byrne, "Advances in membrane mimetic systems for manipulation and analysis of membrane proteins: Detergents, polymers, lipids and scaffolds," *ChemPlusChem*, p. e202300678, 2024.

[80] D. Linke, "Detergents: an overview," *Methods in enzymology*, vol. 463, pp. 603–617, 2009.

[81] E. van den Brink-van der Laan, V. Chupin, J. A. Killian, and B. de Kruijff, "Stability of kcsa tetramer depends on membrane lateral pressure," *Biochemistry*, vol. 43, no. 14, pp. 4240–4250, 2004.

[82] D. L. Parton, J. W. Klingelhoefer, and M. S. Sansom, "Aggregation of model membrane proteins, modulated by hydrophobic mismatch, membrane curvature, and protein class," *Biophysical journal*, vol. 101, no. 3, pp. 691–699, 2011.

[83] E. Sparr, W. L. Ash, P. V. Nazarov, D. T. Rijkers, M. A. Hemminga, D. P. Tieleman, and J. A. Killian, "Self-association of transmembrane α-helices in model membranes: importance of helix orientation and role of hydrophobic mismatch," *Journal of Biological Chemistry*, vol. 280, no. 47, pp. 39324–39331, 2005.

[84] J. A. Killian, "Hydrophobic mismatch between proteins and lipids in membranes," *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, vol. 1376, no. 3, pp. 401–416, 1998.

[85] C. Ostermeier and H. Michel, "Crystallization of membrane proteins," *Current opinion in structural biology*, vol. 7, no. 5, pp. 697–701, 1997.

[86] B. Ismail and S. S. Nielsen, "Basic principles of chromatography," *Food analysis*, vol. 27, pp. 473–498, 2010.

[87] M. F. Vitha, *Chromatography: principles and instrumentation*, vol. 185. John Wiley & Sons, 2016.

[88] J. M. Miller, *Chromatography: concepts and contrasts*. John Wiley & Sons, 2005.

[89] D. Winters, M. Tran, D. Yoo, and K. W. Walker, "Development of biorad ngc and ge äkta pure systems for highly automated three column protein purification employing tandem affinity, buffer exchange and size exclusion chromatography," *Protein Expression and Purification*, vol. 165, p. 105497, 2020.

[90] R. P. Scott, "Principles and practice of chromatography," *Chrom-ed book series*, vol. 1, 2003.

[91] A. M. Davis, S. J. Teague, and G. J. Kleywegt, "Application and limitations of x-ray crystallographic data in structure-based ligand and drug design," *Angewandte Chemie International Edition*, vol. 42, no. 24, pp. 2718–2736, 2003.

[92] J. Holcomb, N. Spellmon, Y. Zhang, M. Doughan, C. Li, and Z. Yang, "Protein crystallization: Eluding the bottleneck of x-ray crystallography," *AIMS biophysics*, vol. 4, no. 4, p. 557, 2017.

[93] R. Giegé, "A historical perspective on protein crystallization from 1840 to the present day," *The FEBS journal*, vol. 280, no. 24, pp. 6456–6497, 2013.

[94] C. Stohrer, S. Horrell, S. Meier, M. Sans, D. von Stetten, M. Hough, A. Goldman, D. C. Monteiro, and A. R. Pearson, "Homogeneous batch micro-crystallization of proteins from ammonium sulfate," *Acta Crystallographica Section D: Structural Biology*, vol. 77, no. 2, pp. 194–204, 2021.

[95] P. D. Shaw Stewart, S. A. Kolek, R. A. Briggs, N. E. Chayen, and P. F. Baldock, "Random microseeding: a theoretical and practical exploration of seed stability and seeding techniques for successful protein crystallization," *Crystal Growth & Design*, vol. 11, no. 8, pp. 3432–3441, 2011.

[96] A. McPherson and B. Cudney, "Optimization of crystallization conditions for biological macromolecules," *Acta Crystallographica Section F: Structural Biology Communications*, vol. 70, no. 11, pp. 1445–1467, 2014.

[97] T. Bergfors, "Seeds to crystals," *Journal of structural biology*, vol. 142, no. 1, pp. 66–76, 2003.

[98] Z. S. Derewenda, "The use of recombinant methods and molecular engineering in protein crystallization," *Methods*, vol. 34, no. 3, pp. 354–363, 2004.

[99] L. Griffin and A. Lawson, "Antibody fragments as tools in crystallography," *Clinical & Experimental Immunology*, vol. 165, no. 3, pp. 285–291, 2011.

[100] C. Hunte and H. Michel, "Crystallisation of membrane proteins mediated by antibody fragments," *Current opinion in structural biology*, vol. 12, no. 4, pp. 503–508, 2002.

[101] A. Warke and C. Momany, "Addressing the protein crystallization bottleneck by cocrystallization," *Crystal Growth and Design*, vol. 7, no. 11, pp. 2219–2225, 2007.

[102] G. Rhodes, *Crystallography made crystal clear: a guide for users of macromolecular models*. Elsevier, 2010.

[103] S. K. Talapatra, M. R. Penny, S. T. Hilton, and F. Kozielski, "Design, 3d printing and validation of a novel low-cost high-capacity sitting-drop bridge for protein crystallization," *Journal of Applied Crystallography*, vol. 52, no. 1, pp. 171–174, 2019.

[104] A. Zabara, T. G. Meikle, R. Trenker, S. Yao, J. Newman, T. S. Peat, F. Separovic, C. E. Conn, M. J. Call, M. E. Call, *et al.*, "Lipidic cubic phase-induced membrane protein crystallization: interplay between lipid molecular structure, mesophase structure and properties, and crystallogenesis," *Crystal Growth & Design*, vol. 17, no. 11, pp. 5667–5674, 2017.

[105] E. M. Landau and J. P. Rosenbusch, "Lipidic cubic phases: a novel concept for the crystallization of membrane proteins," *Proceedings of the National Academy of Sciences*, vol. 93, no. 25, pp. 14532–14535, 1996.

[106] M. W. Martynowycz, A. Shiriaeva, X. Ge, J. Hattne, B. L. Nannenga, V. Cherezov, and T. Gonen, "Microed structure of the human adenosine receptor determined from a single nanocrystal in lcp," *Proceedings of the National Academy of Sciences*, vol. 118, no. 36, p. e2106041118, 2021.

[107] P. Krawinski, L. Smithers, L. van Dalsen, C. Boland, N. Ostrovitsa, J. Pérez, and M. Caffrey, "7.10 mag. a novel host monoacylglyceride for in meso (lipid cubic phase) crystallization of membrane proteins," *Crystal Growth & Design*, vol. 24, no. 7, pp. 2985–3001, 2024.

[108] F. Nüsslin, "Wilhelm conrad röntgen: The scientist and his discovery," *Physica Medica*, vol. 79, pp. 65–68, 2020.

[109] W. R. Hendee, "X rays in medicine," *Physics Today*, vol. 48, no. 11, pp. 51–56, 1995.

[110] P. J. Withers, C. Bouman, S. Carmignato, V. Cnudde, D. Grimaldi, C. K. Hagen, E. Maire, M. Manley, A. Du Plessis, and S. R. Stock, "X-ray computed tomography," *Nature Reviews Methods Primers*, vol. 1, no. 1, p. 18, 2021.

[111] R. Lawaczeck, V. Arkadiev, F. Diekmann, and M. Krumrey, "Monochromatic x-rays in digital mammography," *Investigative radiology*, vol. 40, no. 1, pp. 33–39, 2005.

[112] I. Jandejsek, J. Jakubek, M. Jakubek, P. Prucha, F. Krejci, P. Soukup, D. Turecek, D. Vavrik, and J. Zemlicka, "X-ray inspection of composite materials for aircraft structures using detectors of medipix type," *Journal of Instrumentation*, vol. 9, no. 05, p. C05062, 2014.

[113] D. Mery, D. Saavedra, and M. Prasad, "X-ray baggage inspection with computer vision: A survey," *IEEE Access*, vol. 8, pp. 145620–145633, 2020.

[114] C. Giacovazzo, *Fundamentals of crystallography*, vol. 7. Oxford university press, USA, 2002.

[115] J. Patterson and R. George, "William henry bragg, man and scientist, nobel laureate and first professor of physics, university of adelaide 1886–1909," 2018.

[116] W. H. Bragg and W. L. Bragg, "The reflection of x-rays by crystals," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 88, no. 605, pp. 428–438, 1913.

[117] B. Madhavi, "Development editor note: Biomolecular crystallography," in *Journal of Biomolecular Research & Therapeutics*, 2021.

[118] B. Rupp, *Biomolecular crystallography: principles, practice, and application to structural biology*. Garland Science, 2009.

[119] D. Haas and M. Rossmann, "Crystallographic studies on lactate dehydrogenase at–75° c," *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, vol. 26, no. 7, pp. 998–1004, 1970.

[120] K. H. Nam, "Serial x-ray crystallography," *Crystals*, 2022.

[121] A. M. Wolff, I. D. Young, R. G. Sierra, A. S. Brewster, M. W. Martynowycz, E. Nango, M. Sugahara, T. Nakane, K. Ito, A. Aquila, *et al.*, "Comparing serial x-ray crystallography and microcrystal electron diffraction (microed) as methods for routine structure determination from small macromolecular crystals," *IUCrJ*, vol. 7, no. 2, pp. 306–323, 2020.

[122] J. S. Fraser, H. Van Den Bedem, A. J. Samelson, P. T. Lang, J. M. Holton, N. Echols, and T. Alber, "Accessing protein conformational ensembles using room-temperature x-ray crystallography," *Proceedings of the National Academy of Sciences*, vol. 108, no. 39, pp. 16247–16252, 2011.

[123] R. Neutze, R. Wouts, D. Van der Spoel, E. Weckert, and J. Hajdu, "Potential for biomolecular imaging with femtosecond x-ray pulses," *Nature*, vol. 406, no. 6797, pp. 752–757, 2000.

[124] A. R. Pearson and P. Mehrabi, "Serial synchrotron crystallography for time-resolved structural biology," *Current Opinion in Structural Biology*, vol. 65, pp. 168–174, 2020.

[125] K. H. Nam, "Guide to serial synchrotron crystallography," *Current Research in Structural Biology*, p. 100131, 2024.

[126] H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, *et al.*, "Femtosecond x-ray protein nanocrystallography," *Nature*, vol. 470, no. 7332, pp. 73–77, 2011.

[127] V. Šrajer and M. Schmidt, "Watching proteins function with time-resolved x-ray crystallography," *Journal of physics D: Applied physics*, vol. 50, no. 37, p. 373001, 2017.

[128] D. R. Hekstra, K. I. White, M. A. Socolich, R. W. Henning, V. Šrajer, and R. Ranganathan, "Electric-field-stimulated protein mechanics.," *Nature*, vol. 540, no. 7633, pp. 400–405, 2016.

[129] A. M. Wolff, E. Nango, I. D. Young, A. S. Brewster, M. Kubo, T. Nomura, M. Sugahara, S. Owada, B. A. Barad, K. Ito, *et al.*, "Mapping protein dynamics at high spatial resolution with temperature-jump x-ray crystallography," *Nature chemistry*, vol. 15, no. 11, pp. 1549–1558, 2023.

[130] S. Ghosh, D. Zorić, P. Dahl, M. Bjelčić, J. Johannesson, E. Sandelin, P. Borjesson, A. Björling, A. Banacore, P. Edlund, *et al.*, "A simple goniometer-compatible flow cell for serial synchrotron x-ray crystallography," *Journal of Applied Crystallography*, vol. 56, no. 2, pp. 449–460, 2023.

[131] S. Boutet and G. J. Williams, "The coherent x-ray imaging (cxi) instrument at the linac coherent light source (lcls)," *New Journal of Physics*, vol. 12, no. 3, p. 035024, 2010.

[132] T. Ishikawa, H. Aoyagi, T. Asaka, Y. Asano, N. Azumi, T. Bizen, H. Ego, K. Fukami, T. Fukui, Y. Furukawa, *et al.*, "A compact x-ray free-electron laser emitting in the sub-ångström region," *Nature photonics*, vol. 6, no. 8, pp. 540–544, 2012.

[133] H. Weise, W. Decking, *et al.*, "Commissioning and first lasing of the european xfel," *Proc. FEL'17*, pp. 9–13, 2017.

[134] E. Prat, R. Abela, M. Aiba, A. Alarcon, J. Alex, Y. Arbelo, C. Arrell, V. Arsov, C. Bacellar, C. Beard, *et al.*, "A compact and cost-effective hard x-ray free-electron laser driven by a high-brightness and low-energy electron beam," *Nature Photonics*, vol. 14, no. 12, pp. 748–754, 2020.

[135] J. H. Beale, M. E. Sharpe, *et al.*, "Batch crystallisation methods for serial crystallography experiments," *Acta Crystallographica Section A: Foundations and Advances*, vol. 77, pp. C934–C934, 2021.

[136] P. Berntsen, R. Sharma, M. Kusel, B. Abbey, and C. Darmanin, "Lipidico injection protocol for serial crystallography measurements at the australian synchrotron," *J. Vis. Exp*, vol. 163, p. e61650, 2020.

[137] J. Park and K. H. Nam, "Sample delivery systems for serial femtosecond crystallography at the pal-xfel," in *Photonics*, vol. 10, p. 557, MDPI, 2023.

[138] I. Martiel, H. M. Müller-Werkmeister, and A. E. Cohen, "Strategies for sample delivery for femtosecond crystallography," *Biological Crystallography*, vol. 75, no. 2, pp. 160–177, 2019.

[139] A. G. Leslie, "The integration of macromolecular diffraction data," *Acta Crystallographica Section D: Biological Crystallography*, vol. 62, no. 1, pp. 48–57, 2006.

[140] W. Kabsch, "xds," *Acta Crystallographica Section D: Biological Crystallography*, vol. 66, no. 2, pp. 125–132, 2010.

[141] T. A. White, "Post-refinement method for snapshot serial crystallography," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1647, p. 20130330, 2014.

[142] G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, *et al.*, "Dials: implementation and evaluation of a new integration package," *Acta Crystallographica Section D: Structural Biology*, vol. 74, no. 2, pp. 85–97, 2018.

[143] T. A. White, R. A. Kirian, A. V. Martin, A. Aquila, K. Nass, A. Barty, and H. N. Chapman, "Crystfel: a software suite for snapshot serial crystallography," *Journal of applied crystallography*, vol. 45, no. 2, pp. 335–341, 2012.

[144] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, *et al.*, "Highly accurate protein structure prediction with alphafold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.

[145] A. J. McCoy, M. D. Sammito, and R. J. Read, "Implications of alphafold2 for crystallographic phasing by molecular replacement," *Acta Crystallographica Section D: Structural Biology*, vol. 78, no. 1, pp. 1–13, 2022.

[146] S.-K. Son, H. N. Chapman, and R. Santra, "Multiwavelength anomalous diffraction at high x-ray intensity," *Physical review letters*, vol. 107, no. 21, p. 218102, 2011.

[147] K. Yamashita, N. Kuwabara, T. Nakane, T. Murai, E. Mizohata, M. Sugahara, D. Pan, T. Masuda, M. Suzuki, T. Sato, *et al.*, "Experimental phase determination with selenomethionine or mercury-derivatization in serial femtosecond crystallography," *IUCrJ*, vol. 4, no. 5, pp. 639–647, 2017.

[148] M. S. Hunter, C. H. Yoon, H. DeMirci, R. G. Sierra, E. H. Dao, R. Ahmadi, F. Aksit, A. L. Aquila, H. Ciftci, S. Guillet, *et al.*, "Selenium single-wavelength anomalous diffraction de novo phasing using an x-ray-free electron laser," *Nature communications*, vol. 7, no. 1, p. 13388, 2016.

[149] S. Panjikar, V. Parthasarathy, V. S. Lamzin, M. S. Weiss, and P. A. Tucker, "On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination," *Acta Crystallographica Section D: Biological Crystallography*, vol. 65, no. 10, pp. 1089–1097, 2009.

[150] P. Emsley and K. Cowtan, "Coot: model-building tools for molecular graphics," *Acta crystallographica section D: biological crystallography*, vol. 60, no. 12, pp. 2126–2132, 2004.

[151] D. Sehnal, S. Bittrich, M. Deshpande, R. Svobodová, K. Berka, V. Bazgier, S. Velankar, S. K. Burley, J. Koča, and A. S. Rose, "Mol* viewer: modern web app for 3d visualization and analysis of large biomolecular structures," *Nucleic acids research*, vol. 49, no. W1, pp. W431–W437, 2021.

[152] O. Carugo, "B-factor accuracy in protein crystal structures," *Biological Crystallography*, vol. 78, no. 1, pp. 69–74, 2022.

[153] R. A. LASKOWSKI, N. FURNHAM, and J. M. THORNTON, "The ramachandran plot and protein structure validation," in *Biomolecular forms and functions: a celebration of 50 years of the ramachandran map*, pp. 62–75, World Scientific, 2013.

[154] I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall III, J. Snoeyink, J. S. Richardson, *et al.*, "Molprobity: all-atom contacts and structure validation for proteins and nucleic acids," *Nucleic acids research*, vol. 35, no. suppl_2, pp. W375–W383, 2007.

[155] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, *et al.*, "Phenix: a comprehensive python-based system for macromolecular structure solution," *Acta Crystallographica Section D: Biological Crystallography*, vol. 66, no. 2, pp. 213–221, 2010.

[156] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, *et al.*, "Macromolecular structure determination using x-rays, neutrons and electrons: recent developments in phenix," *Acta Crystallographica Section D: Structural Biology*, vol. 75, no. 10, pp. 861–877, 2019.

[157] C. C. Project *et al.*, "The ccp4 suite: programs for protein crystallography," *Acta crystallographica. Section D, Biological crystallography*, vol. 50, no. Pt 5, pp. 760–763, 1994.

[158] G. M. Sheldrick and T. R. Schneider, "[16] shelxl: High-resolution refinement," in *Methods in enzymology*, vol. 277, pp. 319–343, Elsevier, 1997.

[159] A. Vallejos, G. Katona, and R. Neutze, "Appraising protein conformational changes by resampling time-resolved serial x-ray crystallography data," *Structural Dynamics*, vol. 11, no. 4, 2024.

[160] R. Henderson and J. Moffat, "The difference fourier technique in protein crystallography: errors and their treatment," *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, vol. 27, no. 7, pp. 1414–1420, 1971.

[161] G. Brändén and R. Neutze, "Advances and challenges in time-resolved macromolecular crystallography," *Science*, vol. 373, no. 6558, p. eaba0954, 2021.

[162] U. Raviv, R. Asor, A. Shemesh, A. Ginsburg, T. Ben-Nun, Y. Schilt, Y. Levartovsky, and I. Ringel, "Insight into structural biophysics from solution x-ray scattering," *Journal of Structural Biology*, p. 108029, 2023.

[163] H. D. Mertens and D. I. Svergun, "Combining nmr and small angle x-ray scattering for the study of biomolecular structure and dynamics," *Archives of biochemistry and biophysics*, vol. 628, pp. 33–41, 2017.

[164] H. Kawamukai, S. Takishita, K. Shimizu, D. Kohda, K. Ishimori, and T. Saio, "Conformational distribution of a multidomain protein measured by single-pair small-angle x-ray scattering," *The Journal of Physical Chemistry Letters*, vol. 15, no. 3, pp. 744–750, 2024.

[165] L. Makowski, J. Bardhan, D. Gore, J. Lal, S. Mandava, S. Park, D. Rodi, N. Ho, C. Ho, and R. Fischetti, "Waxs studies of the structural diversity of hemoglobin in solution," *Journal of molecular biology*, vol. 408, no. 5, pp. 909–921, 2011.

[166] J. Choudhury, K. Yonezawa, A. Anu, N. Shimizu, and B. Chaudhuri, "Saxs/waxs data of conformationally flexible ribose binding protein," *Data in Brief*, vol. 52, p. 109932, 2024.

[167] Y. Lee, K. Y. Oang, D. Kim, and H. Ihee, "A comparative review of time-resolved x-ray and electron scattering to probe structural dynamics," *Structural Dynamics*, vol. 11, no. 3, 2024.

[168] B. Chu and B. S. Hsiao, "Small-angle x-ray scattering of polymers," *Chemical reviews*, vol. 101, no. 6, pp. 1727–1762, 2001.

[169] L. C. Johansson, D. Arnlund, T. A. White, G. Katona, D. P. DePonte, U. Weierstall, R. B. Doak, R. L. Shoeman, L. Lomb, E. Malmerberg, *et al.*, "Lipidic phase membrane protein serial femtosecond crystallography," *Nature methods*, vol. 9, no. 3, pp. 263–265, 2012.

[170] L. C. Johansson, D. Arnlund, G. Katona, T. A. White, A. Barty, D. P. DePonte, R. L. Shoeman, C. Wickstrand, A. Sharma, G. J. Williams, *et al.*, "Structure of a photosynthetic reaction centre determined by serial femtosecond crystallography," *Nature communications*, vol. 4, no. 1, p. 2911, 2013.

[171] M. L. Chiu, P. Nollert, M. e. C. Loewen, H. Belrhali, E. Pebay-Peyroula, J. P. Rosenbusch, and E. M. Landau, "Crystallization in cubo: general applicability to membrane proteins," *Acta Crystallographica Section D: Biological Crystallography*, vol. 56, no. 6, pp. 781–784, 2000.

[172] P. Gast, P. Hemelrijk, and A. Hoff, "Determination of the number of detergent molecules associated with the reaction center protein isolated from the photosynthetic bacterium rhodopseudomonas viridis: Effects of the amphiphilic molecule 1, 2, 3-heptanetriol," *FEBS letters*, vol. 337, no. 1, pp. 39–42, 1994.

[173] P. A. Timmins, J. Hauk, T. Wacker, and W. Welte, "The influence of heptane-1, 2, 3-triol on the size and shape of ldao micelles implications for the crystallisation of membrane proteins," *FEBS letters*, vol. 280, no. 1, pp. 115–120, 1991.

[174] R. Andersson, C. Safari, P. Båth, R. Bosman, A. Shilova, P. Dahl, S. Ghosh, A. Dunge, R. Kjeldsen-Jensen, J. Nan, *et al.*, "Well-based crystallization of lipidic cubic phase microcrystals for serial x-ray crystallography experiments," *Acta Crystallographica Section D: Structural Biology*, vol. 75, no. 10, pp. 937–946, 2019.

[175] R. Dods, P. Båth, D. Arnlund, K. R. Beyerlein, G. Nelson, M. Liang, R. Harimoorthy, P. Berntsen, E. Malmerberg, L. Johansson, *et al.*, "From macrocrystals to microcrystals: a strategy for membrane protein serial crystallography," *Structure*, vol. 25, no. 9, pp. 1461–1468, 2017.

[176] A. B. Wohri, W. Y. Wahlgren, E. Malmerberg, L. C. Johansson, R. Neutze, and G. Katona, "Lipidic sponge phase crystal structure of a photosynthetic reaction center reveals lipids on the protein surface," *Biochemistry*, vol. 48, no. 41, pp. 9831–9838, 2009.

[177] C. Wickstrand, R. Dods, A. Royant, and R. Neutze, "Bacteriorhodopsin: Would the real structural intermediates please stand up?," *Biochimica et Biophysica Acta (BBA)-General Subjects*, vol. 1850, no. 3, pp. 536–553, 2015.

[178] G. Selikhanov, T. Fufina, L. Vasilieva, C. Betzel, and A. Gabdulkhakov, "Novel approaches for the lipid sponge phase crystallization of the rhodobacter sphaeroides photosynthetic reaction center," *IUCrJ*, vol. 7, no. 6, pp. 1084–1091, 2020.

[179] M. Stowell, T. McPhillips, D. Rees, S. Soltis, E. Abresch, and G. Feher, "Light-induced structural changes in photosynthetic reaction center: implications for mechanism of electron-proton transfer," *Science*, vol. 276, no. 5313, pp. 812–816, 1997.

[180] L. Li, D. Mustafi, Q. Fu, V. Tereshko, D. L. Chen, J. D. Tice, and R. F. Ismagilov, "Nanoliter microfluidic hybrid method for simultaneous screening and optimization validated with crystallization of membrane proteins," *Proceedings of the National Academy of Sciences*, vol. 103, no. 51, pp. 19243–19248, 2006.

[181] C. R. D. Lancaster and H. Michel, "The coupling of light-induced electron transfer and proton uptake as derived from crystal structures of reaction centres from rhodopseudomonas viridis modified at the binding site of the secondary quinone, qb," *Structure*, vol. 5, no. 10, pp. 1339–1359, 1997.

[182] J. Breton, "Absence of large-scale displacement of quinone qb in bacterial photosynthetic reaction centers," *Biochemistry*, vol. 43, no. 12, pp. 3318–3326, 2004.

[183] J. Koepke, E.-M. Krammer, A. R. Klingen, P. Sebban, G. M. Ullmann, and G. Fritzsch, "ph modulates the quinone position in the photosynthetic reaction center from rhodobacter sphaeroides in the neutral and charge separated states," *Journal of Molecular Biology*, vol. 371, no. 2, pp. 396–409, 2007.

[184] R. H. Baxter, B.-L. Seagle, N. Ponomarenko, and J. R. Norris, "Cryogenic structure of the photosynthetic reaction center of blastochloris viridis in the light and dark," *Acta Crystallographica Section D: Biological Crystallography*, vol. 61, no. 5, pp. 605–612, 2005.

[185] A. W. Roszak, V. Moulisová, A. D. Reksodipuro, A. T. Gardiner, R. Fujii, H. Hashimoto, N. W. Isaacs, and R. J. Cogdell, "New insights into the structure of the reaction centre from blastochloris viridis: evolution in the laboratory," *Biochemical Journal*, vol. 442, no. 1, pp. 27–37, 2012.

[186] L. Li, S. Nachtergaele, A. M. Seddon, V. Tereshko, N. Ponomarenko, and R. F. Ismagilov, "Simple host- guest chemistry to modulate the process of concentration and crystallization of membrane proteins by detergent capture in a microfluidic device," *Journal of the American Chemical Society*, vol. 130, no. 43, pp. 14324–14328, 2008.

[187] P. Qian, D. J. Swainsbury, T. I. Croll, J. H. Salisbury, E. C. Martin, P. J. Jackson, A. Hitchcock, P. Castro-Hartmann, K. Sader, and C. N. Hunter, "Cryo-em structure of the monomeric rhodobacter sphaeroides rc–lh1 core complex at 2.5 å," *Biochemical Journal*, vol. 478, no. 20, pp. 3775–3790, 2021.

[188] J. Lieske, M. Cerv, S. Kreida, D. Komadina, J. Fischer, M. Barthelmess, P. Fischer, T. Pakendorf, O. Yefanov, V. Mariani, *et al.*, "On-chip crystallization for serial crystallography experiments and on-chip ligand-binding studies," *IUCrJ*, vol. 6, no. 4, pp. 714–728, 2019.

[189] A. B. Wöhri, G. Katona, L. C. Johansson, E. Fritz, E. Malmerberg, M. Andersson, J. Vincent, M. Eklund, M. Cammarata, M. Wulff, *et al.*, "Light-induced structural changes in a photosynthetic reaction center caught by laue diffraction," *science*, vol. 328, no. 5978, pp. 630–633, 2010.

[190] R. Dods, P. Båth, D. Morozov, V. A. Gagnér, D. Arnlund, H. L. Luk, J. Kübel, M. Maj, A. Vallejos, C. Wickstrand, *et al.*, "Ultrafast structural changes within a photosynthetic reaction centre," *Nature*, vol. 589, no. 7841, pp. 310–314, 2021.

[191] K. M. Dalton, J. B. Greisman, and D. R. Hekstra, "A unifying bayesian framework for merging x-ray diffraction data," *Nature Communications*, vol. 13, no. 1, p. 7764, 2022.

[192] M. H. Vos, F. Rappaport, J.-C. Lambry, J. Breton, and J.-L. Martin, "Visualization of coherent nuclear motion in a membrane protein by femtosecond spectroscopy," *Nature*, vol. 363, no. 6427, pp. 320–325, 1993.

[193] M. H. Vos, M. R. Jones, C. N. Hunter, J. Breton, and J.-L. Martin, "Coherent nuclear dynamics at room temperature in bacterial reaction centers.," *Proceedings of the National Academy of Sciences*, vol. 91, no. 26, pp. 12701–12705, 1994.

[194] C. Rischel, D. Spiedel, J. P. Ridge, M. R. Jones, J. Breton, J.-C. Lambry, J.-L. Martin, and M. H. Vos, "Low frequency vibrational modes in proteins: changes induced by point-mutations in the protein-cofactor matrix of bacterial reaction centers," *Proceedings of the National Academy of Sciences*, vol. 95, no. 21, pp. 12306–12311, 1998.

[195] V. Palaniappan, C. C. Schenck, and D. F. Bocian, "Low-frequency near-infrared-excitation resonance raman spectra of (m) h202l mutant reaction centers from rhodobacter sphaeroides. implications for the structural, vibronic, and electronic properties of the bacteriochlorin cofactors," *The Journal of Physical Chemistry*, vol. 99, no. 46, pp. 17049–17058, 1995.

[196] H. Wang, S. Lin, J. P. Allen, J. C. Williams, S. Blankert, C. Laser, and N. W. Woodbury, "Protein dynamics control the kinetics of initial electron transfer in photosynthesis," *Science*, vol. 316, no. 5825, pp. 747–750, 2007.

[197] H. Wang, S. Lin, E. Katilius, C. Laser, J. P. Allen, J. C. Williams, and N. W. Woodbury, "Unusual temperature dependence of photosynthetic electron transfer due to protein dynamics," *The Journal of Physical Chemistry B*, vol. 113, no. 3, pp. 818–824, 2009.

[198] U. Weierstall, D. James, C. Wang, T. A. White, D. Wang, W. Liu, J. C. Spence, R. Bruce Doak, G. Nelson, P. Fromme, *et al.*, "Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography," *Nature communications*, vol. 5, no. 1, p. 3309, 2014.

[199] D. DePonte, U. Weierstall, K. Schmidt, J. Warner, D. Starodub, J. Spence, and R. Doak, "Gas dynamic virtual nozzle for generation of microscopic droplet streams," *Journal of Physics D: Applied Physics*, vol. 41, no. 19, p. 195505, 2008.

[200] R. D. Miller, O. Paré-Labrosse, A. Sarracini, and J. E. Besaw, "Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin in the multiphoton regime and biological relevance," *Nature communications*, vol. 11, no. 1, p. 1240, 2020.

[201] R. Neutze and R. D. Miller, "Energetic laser pulses alter outcomes of x-ray studies of proteins," 2024.

[202] P. Nogly, T. Weinert, D. James, S. Carbajo, D. Ozerov, A. Furrer, D. Gashi, V. Borin, P. Skopintsev, K. Jaeger, *et al.*, "Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser," *Science*, vol. 361, no. 6398, p. eaat0094, 2018.

[203] G. Nass Kovacs, J.-P. Colletier, M. L. Grünbein, Y. Yang, T. Stensitzki, A. Batyuk, S. Carbajo, R. B. Doak, D. Ehrenberg, L. Foucar, *et al.*, "Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin," *Nature communications*, vol. 10, no. 1, p. 3177, 2019.

[204] T. R. Barends, A. Gorel, S. Bhattacharyya, G. Schirò, C. Bacellar, C. Cirelli, J.-P. Colletier, L. Foucar, M. L. Grünbein, E. Hartmann, *et al.*, "Influence of pump laser fluence on ultrafast myoglobin structural dynamics," *Nature*, vol. 626, no. 8000, pp. 905–911, 2024.

[205] K. Pande, C. D. Hutchison, G. Groenhof, A. Aquila, J. S. Robinson, J. Tenboer, S. Basu, S. Boutet, D. P. DePonte, M. Liang, *et al.*, "Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein," *Science*, vol. 352, no. 6286, pp. 725–729, 2016.

[206] A. Hosseinizadeh, N. Breckwoldt, R. Fung, R. Sepehr, M. Schmidt, P. Schwander, R. Santra, and A. Ourmazd, "Few-fs resolution of a photoactive protein traversing a conical intersection," *Nature*, vol. 599, no. 7886, pp. 697–701, 2021.

[207] D. Arnlund, L. C. Johansson, C. Wickstrand, A. Barty, G. J. Williams, E. Malmerberg, J. Davidsson, D. Milathianaki, D. P. DePonte, R. L. Shoeman, *et al.*, "Visualizing a protein quake with time-resolved x-ray scattering at a free-electron laser," *Nature methods*, vol. 11, no. 9, pp. 923–926, 2014.

[208] M. Shelby, D. Gilbile, T. Grant, C. Seuring, B. Segelke, W. He, A. Evans, T. Pakendorf, P. Fischer, M. Hunter, *et al.*, "A fixed-target platform for serial femtosecond crystallography in a hydrated environment," *IUCrJ*, vol. 7, no. 1, pp. 30–41, 2020.

[209] M. Andersson, E. Malmerberg, S. Westenhoff, G. Katona, M. Cammarata, A. B. Wöhri, L. C. Johansson, F. Ewald, M. Eklund, M. Wulff, J. Davidsson, and R. Neutze, "Structural dynamics of light-driven proton pumps.," *Structure (London, England : 1993)*, vol. 17, pp. 1265–75, 9 2009.

[210] R. K. Clayton and B. J. Clayton, "Molar extinction coefficients and other properties of an improved reaction center preparation from rhodopseudomonas viridis," *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, vol. 501, no. 3, pp. 478–487, 1978.

[211] M. L. Grünbein, M. Stricker, G. Nass Kovacs, M. Kloos, R. B. Doak, R. L. Shoeman, J. Reinstein, S. Lecler, S. Haacke, and I. Schlichting, "Illumination guidelines for ultrafast pump–probe experiments by serial femtosecond crystallography," *Nature methods*, vol. 17, no. 7, pp. 681–684, 2020.