





Hamburg University of Applied Sciences

Faculty of Life Sciences

Illuminated in Time: Evaluating a Reporter System and a Molecular Cage for Time-Resolved Serial Femtosecond X-ray Crystallography at the European X-ray Free-Electron Laser

Master Thesis

in the study program Pharmaceutical Biotechnology

submitted by

Marie Christin Heim

[redacted]

Hamburg,

4th of August 2026

1st Reviewer: Prof. Dr. Jörg Andrä (HAW Hamburg)

2nd Reviewer: Dr. Richard Bean (European XFEL)

The thesis was supervised and prepared at the European X-ray Free-Electron Facility in the SPB/SFX group.

I. Affidavit

I, Marie Christin Heim, hereby confirm that the present master's thesis was independently written by me without any external or unauthorised aids. All sources that have been used, whether verbatim or in essence, are indicated as such. I confirm that this thesis has not been submitted as part of another examination process nor in identical form.

[redacted]

Hamburg, 04.08.2026

II. Acknowledgements

I would like to take this opportunity to express my sincere gratitude to all those who have supported me in the preparation and completion of this thesis. First and foremost, I would like to thank my second supervisor, Dr. Richard Bean, who gave me the opportunity to join the SPB/SFX group and work at such a prestigious research facility.

My deepest gratitude goes to Dr. Raphaël de Wijn for including me in his research and supervising every step of this thesis. I am incredibly grateful for the patience and guidance I received, and for all the knowledge shared, from the very basics of crystallography to the details of time-resolved crystallography. Your support was essential to this thesis and made this journey a truly valuable experience.

I would also like to thank my first supervisor, Prof. Dr. Jörg Andrä, for supervising my thesis from the university side.

I owe special thanks to Dr. Tomas Popelar, Dr. Egor Sobolev, and Dr. Robin Schubert, who were directly involved in the success of this thesis by setting up and operating the nanosecond laser, helping with data processing, specifically detector geometry refinement, and supporting the use of the Raman microscope. I would also like to thank Dr. Diogo Melo and Dr. Iñaki de Diego for sharing their knowledge on data analysis and experimental methods beyond the scope of this thesis, as well as Dr. David von Stetten for the assistance in using the P13 beamline operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany) for collecting synchrotron data.

My thanks also go to the entire SPB/SFX group and XBI team for welcoming me into such a supportive environment, including me in active beamtime at the SPB/SFX instrument and introducing me to many different projects and their own work. Working at the European XFEL has been an exceptional experience I will always cherish.

Finally, I would like to thank my parents for supporting me through my studies and my partner Jonas for always being there for me, as well as Anna, Jana, Jana, Lisa and Malin for proofreading the thesis and the endless hours of studying together. Without you, university would not have been the same.

Table of Contents

I. Affidavit I

II. Acknowledgements II

III. List of Abbreviations V

IV. List of Tables VI

V. List of Figures VII

1. Abstract 1

2. Introduction 2

2.1. Aim of the study 3

3. Theoretical Background 4

3.1. Time-resolved serial femtosecond X-ray crystallography 4

3.1.1. Experimental workflow for TR-SFX experiments 6

3.2. TR-SFX at the European XFEL 8

3.3. DNA as a photon-sensitive reporter system 12

3.4. HPBC cage compound for the controlled release of oxygen 14

4. Materials and Methods 17

4.1. Materials 17

4.2. Preparation of crystallisation samples 22

4.3. UV spectrometry 22

4.4. Crystallisation of DNA oligomers 22

4.5. Agarose gel electrophoresis 23

4.6. UV irradiation 23

4.7. Raman spectrometry 23

4.8. Single crystal diffraction 24

4.9. Indexing of serial crystallography diffraction data 25

4.10. Phasing and refinement 25

4.10.1. Ligand identification and occupancy determination 26

5. Results 27
5.1. Preliminary tests for a DNA reporter system 27
5.1.1. DNA crystallisation screening 27
5.1.2. Thymine dimerisation tests 31
5.1.3. Collection of diffraction data 34
5.2. Data analysis of the HPBC oxygen cage 36
5.2.1. Indexing of serial crystallography diffraction data 36
5.2.2. Phasing and Refinement 40
6. Discussion 48
6.1. DNA reporter system 48
6.2. HPBC oxygen cage 51
7. Conclusion and Outlook 54
8. References 56
9. Appendix 61

III. List of Abbreviations

a.u.	Arbitrary Unit
bp	base pairs
C	Cytosine
CC	Correlation Coefficient
CPD	Cyclobutene Pyrimidine Dimers
DESY	*Deutsches Elektronen-Synchrotron*
DNA	Deoxyribonucleic Acid
EMBL	European Molecular Biology Laboratory
eV	Electron Volt
F_{calc}	Calculated structure factors
F_{obs}	Observed structure factors
His	Histidine
HPBC	(μ-peroxo)(μ-hydroxo)bis[bis(bipyridyl)cobalt(III)] I
IRU	Interaction Region Upstream
Mb	Myoglobin
PEG	Polyethylene glycol
SASE	Self-Amplified Spontaneous Emission
SFX	Serial Femtosecond X-ray Crystallography
SNR	Signal-to-Noise Ratio
SPB/SFX	Single Particles, Clusters, and Biomolecules & Serial Femtosecond Crystallography
T	Thymine
TR-SFX	Time-Resolved Serial Femtosecond X-ray Crystallography
UV	Ultraviolet
XFEL	X-ray Free-Electron Laser

IV. List of Tables

Table 1: Used devices & instruments 17
Table 2: Used laser and beamline 18
Table 3: Used consumables 18
Table 4: Used chemicals 19
Table 5: Used buffers 20
Table 6: Used DNA strands 21
Table 7: Used software 21
Table 8: Data collection and indexing statistics 37
Table 9: Refinement statistics 41
Table 10: Ligand occupancy and protein movement statistics 45
Table 11: Screened and collected unirradiated crystals. 62
Table 12: Screened and collected irradiated crystals. 62
Table 13: List of all analysed myoglobin structures 63

V. List of Figures

Figure 1: Schematic representation of Bragg's law 4
Figure 2: Schematic workflow for TR-SFX experiments 6
Figure 3: Crystallisation phase diagram 7
Figure 4: The European X-ray Free-Electron Laser 9
Figure 5: Schematic illustration of the undulator and photon tunnels of the European XFEL 9
Figure 6: The SPB/SFX instrument at the European XFEL 10
Figure 7: Pump-probed TR-SFX experiments with liquid jets 11
Figure 8: Inside of the IRU sample chamber 12
Figure 9: Structure of thymine photoproducts 13
Figure 10: Schematic structure of native DNA and with a thymine CPD 13
Figure 11: Structure of the HPBC oxygen cage 15
Figure 12: Different ligation states of myoglobin 16
Figure 13: Time duration until crystal growth dependent on crystallisation condition 28
Figure 14: Examples of regular- and irregular-shaped crystals 29
Figure 15: Crystal sizes and shapes dependent on the crystallisation condition 30
Figure 16: Agarose gel electrophoresis of crystals grown in buffers with 35% (w/v) PEG .. 31
Figure 17: UV-absorption spectra of irradiated and unirradiated samples 32
Figure 18: Raman spectra and change in the peak intensities 33
Figure 19: Single diffraction images of DNA crystals during screening 35
Figure 20: Offline processed diffraction data images 35
Figure 21: Electron density maps of the reference structures 40
Figure 22: Electron density from the shared-reservoir approach 43
Figure 23: Electron density from the mix-and-inject approach 44
Figure 24: Distance difference graphs 47
Figure 25: Collected UV-absorption spectra 61
Figure 26: Collected Raman spectra 61

1. Abstract

The invention of X-ray free-electron lasers revolutionised the field of time-resolved crystallography by utilising the ultrashort X-ray pulses to probe reactions down to the femtosecond timescale, permitting unprecedented insights into the dynamics of biological processes. At the European X-ray Free-Electron Laser, the megahertz repetition rate and high peak brilliance enable the collection of multiple time points within a single experiment at a high spatio-temporal resolution. However, beamtime is highly limited and requires extensive preparatory work and additional data-processing steps compared to traditional crystallography.

In this study, the preparatory work and data analysis are represented by analysing two distinct systems that both aim to enhance experiments performed at the Single Particles, Clusters, and Biomolecules & Serial Femtosecond Crystallography instrument at the European XFEL facility. Initial experiments on thymine dimerisation in DNA oligomers to validate the instrumental set-up for pump-probed experiments involved UV irradiation, subsequent testing for structural changes, as well as crystallisation and collection of diffraction data at a synchrotron source. The analysis of time-resolved serial femtosecond crystallography diffraction data was performed as part of an experiment on the applicability of an oxygen photocage by processing the diffraction patterns and refining the structures. While leakage from the photocage led to premature oxygenation of the test molecule, which could not be sufficiently compensated by a reduced contact time, the DNA reporter system showed structural changes upon UV irradiation in the initial tests and, with further improvement, could be implemented as a validation system improving the efficient use of beamtime.

2. Introduction

From its beginnings in the early 20th century, X-ray crystallography has been one of the most powerful techniques for determining the three-dimensional structure of macromolecules and has contributed to the understanding of biological processes as well as the development of new pharmaceuticals. Although additional methods like nuclear magnetic resonance or cryo-electron microscopy have been developed in the meantime, and the prediction of structures with artificial intelligence has recently become popular, X-ray crystallography remains the method of choice for high atomic resolution [1]. Since the first protein structures were solved in the 1950s, X-ray crystallography has evolved from a static picture of one structure derived from a single crystal to a dynamic visualisation of time-resolved molecular processes at atomic resolution [2–4]. While static structures provide detailed information about the arrangement of atoms, their biological function remains undisclosed. Conventional crystallography cannot capture reactions involving intermediate states such as enzymatic turnovers, ligand binding, allosteric regulation, or bond breaking. Meanwhile, time-resolved crystallography addresses this limitation by capturing a series of structural snapshots that reveal those molecular processes [4,5]. The high pulse repetition rate of X-ray free-electron lasers (XFELs) enables probing reactions down to the femtosecond range [4,6]. To ensure the best possible results from time-resolved serial femtosecond X-ray crystallography (TR-SFX), careful preparation and data analysis are needed. At the European XFEL, a unique megahertz repetition rate allows the collection of a full data set in a few minutes, enabling the mapping of different time points in one experiment at high temporal resolution [3,4]. However, beamtime at XFELs is highly limited and technically challenging. Therefore, robust validation systems for the instrumental setup and reliable reaction initiation are required.

2.1. Aim of the study

This study aimed to explore the required work steps of time-resolved serial femtosecond crystallography by investigating two distinct systems. The investigation of a photosensitive deoxyribonucleic acid (DNA) reporter system demonstrates preliminary steps, including crystallisation, activity tests and diffraction screening at a synchrotron. Evaluating diffraction data from an experiment with an oxygen photocage performed at the Single Particles, Clusters, and Biomolecules & Serial Femtosecond Crystallography (SPB/SFX) instrument at the European XFEL demonstrates the processing and analysis of TR-SFX diffraction data. Both systems hold potential advances for future experiments at the SPB/SFX instrument at the European XFEL. The DNA reporter system could serve as a validation system for the experimental setup, while the photocage could extend the applicability of TR-SFX to oxygen-dependent, light-insensitive systems.

3. Theoretical Background

3.1. Time-resolved serial femtosecond X-ray crystallography

Time-resolved serial femtosecond X-ray crystallography is based on the same principles as the conventional single-crystal approach. Incoming X-rays are elastically scattered when colliding with the electrons of an atom. Because each atom of a molecule acts as an individual scatterer in three-dimensional space, but the energy of the scattered X-rays equals the energy of the incoming beam, the scattered X-rays vary in phase. In a crystal, where the molecules are periodically arranged, constructive interference occurs if the difference in the path of the scattered X-rays equals the wavelength of the X-rays or a multiple of it. This relation is described by the Bragg equation:

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

where n is a positive integer, λ is the wavelength, d is the distance between parallel planes, and θ is the angle between the incoming and the scattered X-rays and the plane (Figure 1). Strong constructive interference can be measured as Bragg peaks, which are then used to reconstruct the three-dimensional structure of the molecules inside the crystal [7,8].

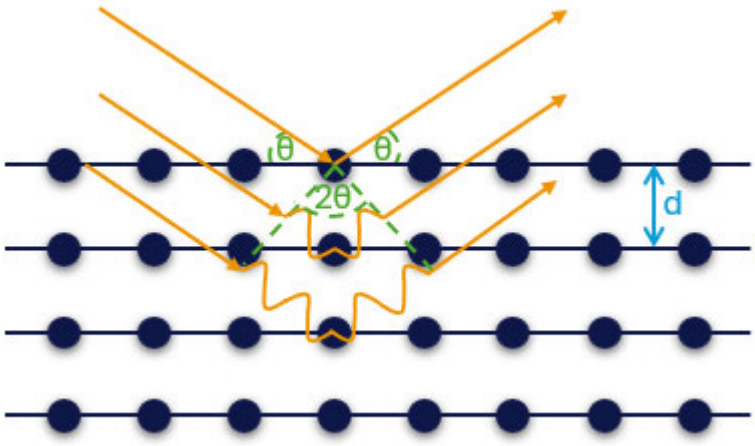


Figure 1: Schematic representation of Bragg's law. Bragg's law describes the conditions that must be met for constructive interference between the scattered X-rays. Parallel scattered X-rays have the same phase when the difference in the path length equals a multiple of the wavelength. Orange: X-rays, green: angle between the incoming and the scattered X-rays and the plane, blue: distance between the planes. Image adapted from **[7]**

The advances of X-ray sources to 3rd- and 4th-generation synchrotrons and X-ray free-electron lasers have enabled extremely bright X-rays, which increased the achievable spatial resolution of diffraction experiments. However, higher X-ray energy also causes more damage or, at XFELs, even destruction of the crystals. To overcome that limitation, rather than using a single

crystal to collect a full dataset, a large number of crystals are sequentially exposed to the X-ray beam. In serial crystallography, each diffraction pattern is therefore collected from a single crystal in a random orientation. At XFELs, the ultrashort femtosecond X-ray pulses outrun the time required for radiation-induced damage to form, and diffraction data is collected before the sample is influenced by radiation-induced structural changes. This concept, known as "diffraction before destruction", enables the collection of diffraction data at room temperature. The serial supply of crystals also diminishes the need for cryo-cooling, as induced damage does not accumulate in one crystal and probing at room temperature near physiological conditions is possible [3,8,9]. The high pulse repetition rate also requires a continuous replenishment of the sample at a speed that matches the pulse frequency. To achieve this, a variety of sample delivery methods have been developed, including liquid jets, drop-on-demand injectors, drop-on-tape drives, and fixed-target approaches such as microchips [3,5,9].

Together, short X-ray pulses, room-temperature data collection, and the "diffraction before destruction" approach opened the field of time-resolved macromolecular studies, down to the femtosecond scale [5,8,9]. TR-SFX combines the determination of macromolecular structures with reaction dynamics by collecting diffraction data at different time points after reaction initiation near atomic or at atomic resolution [3,4]. This high spatio-temporal resolution allows the exploration of short-lived transient species and intermediate states. However, for successful time-resolved experiments, precise timing of the investigated reaction is necessary. Initiators of a reaction can be temperature or pH jumps, mixing of the reaction components (mix-and-inject), or optical laser pulses for photosensitive reactions (pump-probe). Mix-and-inject approaches can be broadly applied and require a comparably simple set-up. However, the temporal resolution depends on the diffusion time of reaction components and is limited to a few milliseconds to several microseconds. Temperature and pH jumps enable the probing of reactions down to the nanosecond range but can induce physical stress to the molecules by inducing an unnatural environment [3,5,9–11]. Utilising an optical laser pulse to trigger a photosensitive system can initiate reactions in less than a few nanoseconds down to the femtosecond scale near physiological conditions. However, only about 0.5% of proteins are light sensitive, drastically limiting the applicability, and careful laser tuning is required to avoid multiphoton effects. Nevertheless, photoreactive systems are well suited for studying short-lived intermediates and broadening the applicability by utilising photolabile molecules, which bind reaction components, is under investigation [3,5,11].

3.1.1. Experimental workflow for TR-SFX experiments

With five XFEL facilities worldwide, accessible beamtime is limited, especially when specific experimental conditions can only be met at selected facilities. To minimise potential setbacks and get the best results possible, experiments need to be carefully planned, and robust workflows need to be established beforehand. The different steps included in TR-SFX experiments can be summarised in three steps, as visualised in Figure 2.

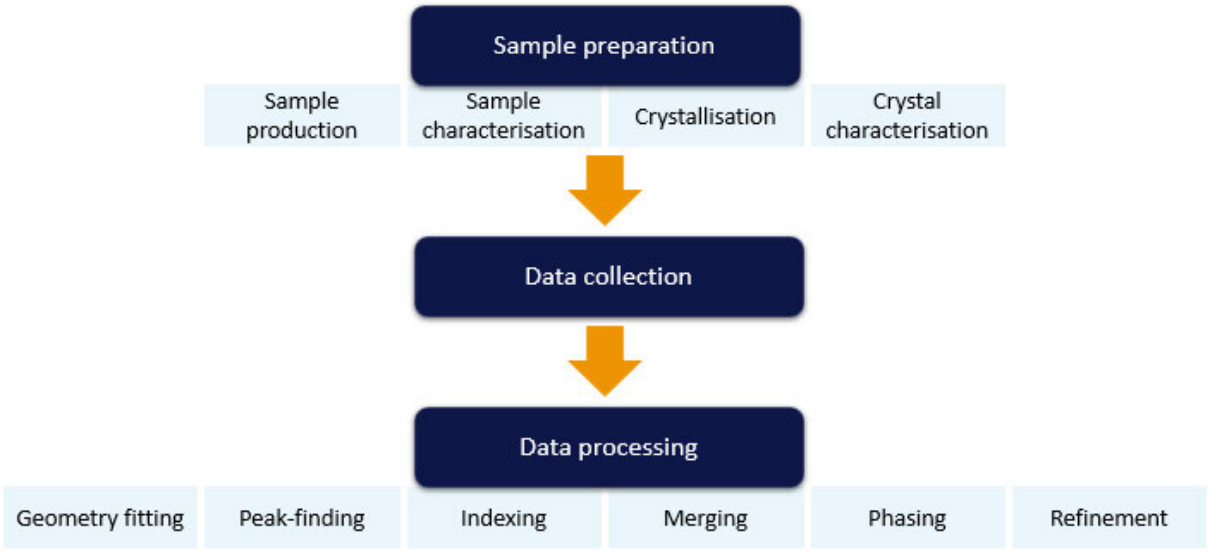


Figure 2: Schematic workflow for TR-SFX experiments. TR-SFX experiments can be summarised into three steps, each of them including several considerations and preparations. Establishing a robust workflow, from sample production to data analysis, is essential for obtaining reliable results.

Sample preparation includes the initial steps of sample production and purification, both of which can already be very time-consuming, as well as the characterisation of the sample, e.g. its concentration and activity regarding the reaction of interest, crystallisation of the sample, including a broad crystallisation screening and refining promising conditions, as well as the characterisation of those crystals by validating the crystal content, determining their intrinsic order and screening for diffraction. One of the most crucial steps in sample preparation is finding conditions from which crystals with high intrinsic order can be obtained, as low order decreases diffraction resolution [11–13].

Although crystallisation has been performed for many years, it is still not thoroughly understood. The screening of conditions is therefore often based on commercial screening kits investigating a large number of varying conditions known to crystallise molecules. If crystals grow, the found conditions are then optimised [13,14]. The process of crystallisation and adjustment of the crystallisation conditions can be visualised with the phase diagram (Figure 3). The phase diagram represents different phases of soluble, metastable, nucleation

and precipitation that the molecule of interest can be in, based on its concentration and the concentration of the crystallisation agent. By varying both concentrations, the solution is shifted through the different phases. Reaching the nucleation zone, the molecules form nucleation points from which crystals can grow. Due to the formation of a solid state, the concentration of molecules in the solution decreases, and the solution moves into the metastable zone where crystals continue to grow [14,15].

One commonly employed method to shift a solution through the phases is vapour diffusion. An undersaturated sitting or hanging drop with the molecule of interest in the crystallisation buffer is placed next to a reservoir with a higher concentration of the buffer components. Due to the different concentrations, water from the drop diffuses to the reservoir, increasing the concentration of crystallisation agent and molecule of interest until the nucleation or precipitation zone is reached. If the molecules nucleate or precipitate, the concentration of the molecule of interest in solution decreases, and the drop reaches the metastable phase. If nucleation points have formed, crystals grow until the solution reaches the solubility curve and an equilibrium between dissolution and crystal growth is reached (Figure 3) [14,15].

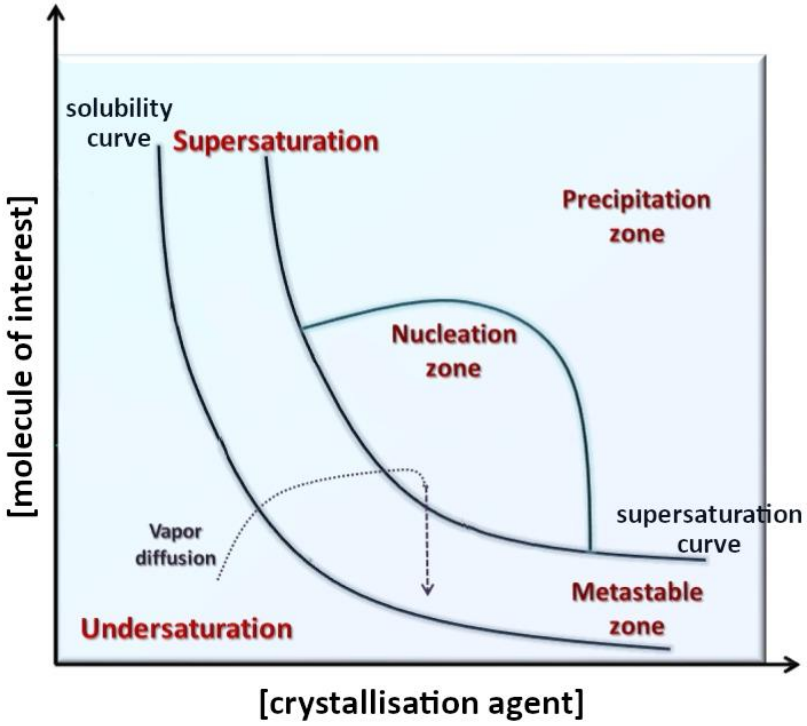


Figure 3: Crystallisation phase diagram. The phase diagram displays different zones of solution conditions. In the undersaturated zone, molecules stay in solution. In the nucleation zone, molecules form nucleation points from which crystals can grow. In the metastable zone, crystals can grow if nucleation points are present. In the precipitation zone, high concentrations of the molecules lead to the precipitation of the molecule of interest, making it unavailable for crystallisation. With vapour diffusion (purple dashed line), all molecules in solution are concentrated, shifting the crystallisation solution through the phase diagram to reach the nucleation phase. When molecules nucleate, the concentration of molecules in solution decreases, shifting the solution into the metastable phase. Image adapted from **[14]**

Evaluating the morphology of crystals can give an initial hint of their quality, as good internal order should produce sharp-edged crystals with a regular form. However, screening for diffraction at a synchrotron will tell whether the intrinsic order is well enough to diffract to high resolution [12,13].

Once a system is characterised, considerations like sample delivery method and reaction triggering are relevant and realistic probing delays must be taken into account before the beamtime [11]. After diffraction data has been collected, it needs to be processed. Data processing includes refining the detector geometry, peak-finding, indexing and merging, followed by the phasing and refinement of the structure [11,16]. Refinement of the detector geometry is needed to accurately match the position of the found peaks to predicted ones during indexing. For peak-finding adjustments, such as peak sizes in pixels, the background definition or acceptable signal-to-noise ratios (SNR) need to be considered to determine true Bragg peaks. During indexing, the position of the found peaks on the detector is used to determine the reciprocal crystal lattice and unit cell parameters by matching the found peaks to predicted positions for each diffraction pattern. In the merging step, the individual reflections from all processed frames are combined into structure factors [17,18]. Structure factors, described as $F(h, k, l)$ where h, k, l are coordinates in the reciprocal lattice, include the amplitudes I and phases of the scattered X-rays from which the position of the scatterer could be determined. However, collected diffraction data only hold information on the amplitude. The observed structure factors derived from the data processing (F_{obs}) therefore miss the phase information [19,20]. A method to obtain initial phases is molecular replacement, in which the phases of related solved structures are used to describe the position of the atoms from which the amplitudes arose [20,21]. During structural refinement, the positions of the atoms are used to calculate new phases, from which new structure factors (F_{calc}) and new atomic positions are calculated, ultimately finding a structure that matches the observed structure factors best [11,16].

3.2. TR-SFX at the European XFEL

The European X-ray Free-Electron Laser facility was the world's first high-repetition-rate hard XFEL that started commissioning in 2016. Today, it continues to be one of the world's most advanced X-ray sources, delivering X-rays at the world's highest peak brilliance at a unique megahertz repetition rate, in which X-ray pulses are delivered in trains [6,22,23].



Figure 4: The European X-ray Free-Electron Laser. The European XFEL runs underground for 3.4 km, starting at the campus of the *Deutsches Elektronen-Synchrotron* (DESY) in Hamburg and ending in the experimental hutches at the user facility in Schenefeld. Image copyright: European XFEL / Luftaufnahmen 2017/18/21: FHH, Landesbetrieb Geoinf. und Vermessung

The European XFEL stretches over a total distance of 3.4 km (Figure 4). Within the first 30 cm, the electron gun "GUN5", newly installed at the end of 2025, generates free-electron bunches by striking a cathode with an intense laser beam and freeing the electrons via the photoelectric effect [24]. From there, electrons are injected into the 1.7 km long linear accelerator, in which they are accelerated to a maximum of 17.5 gigaelectron volts (GeV). After acceleration, the electrons are directed into one of the three self-amplified spontaneous emission undulator tunnels (SASE), where a magnetic field forces them into a curved trajectory. At each bend, photons are emitted and amplified through the process, generating X-ray pulses between 1 to over 20 kiloelectron volts (keV) [23,25,26].

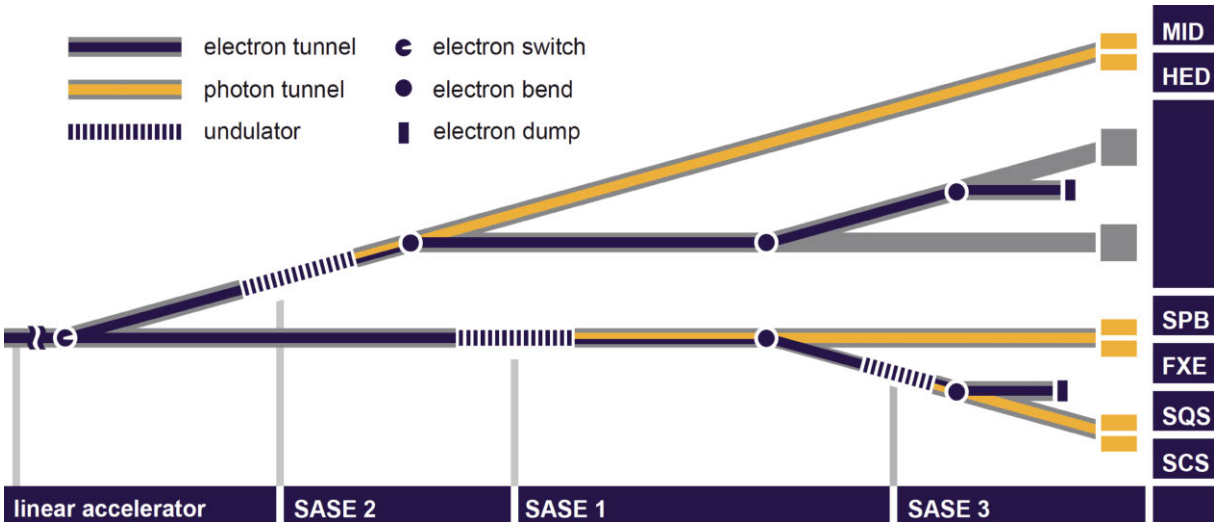


Figure 5: Schematic illustration of the undulator and photon tunnels of the European XFEL. From the linear accelerator, an electron switch leads electron bunches either into SASE1 or SASE2, where photons are generated in the undulators. From SASE1, electrons are led into SASE3, where additional photons are generated. Electrons get dumped, and photons are led to the experimental hutches **[25]**.

At the end of SASE1, the Single Particles, Clusters, and Biomolecules & Serial Femtosecond Crystallography instrument is located (Figure 5). One of the main purposes of this instrument is to exploit the fast pulse repetition rate for time-resolved SFX experiments and capture reaction intermediates down to the femtosecond range. At SPB/SFX, the X-ray beam is generated with a photon energy range from 3-16 keV and focused to a spot size between 0.1 and a few µm, allowing sample sizes in the same size range that otherwise would be difficult to measure at synchrotron beamlines due to high background noise [3,6]. The attenuating, focusing, and cleaning of the beam are conducted with multiple attenuators, mirrors and slits located upstream, reaching from the SASE1 tunnel through an optics hutch into the experimental hutch. A detailed description can be found in the initial installation report of the SPB/SFX instrument by Mancuso *et. al* [6].

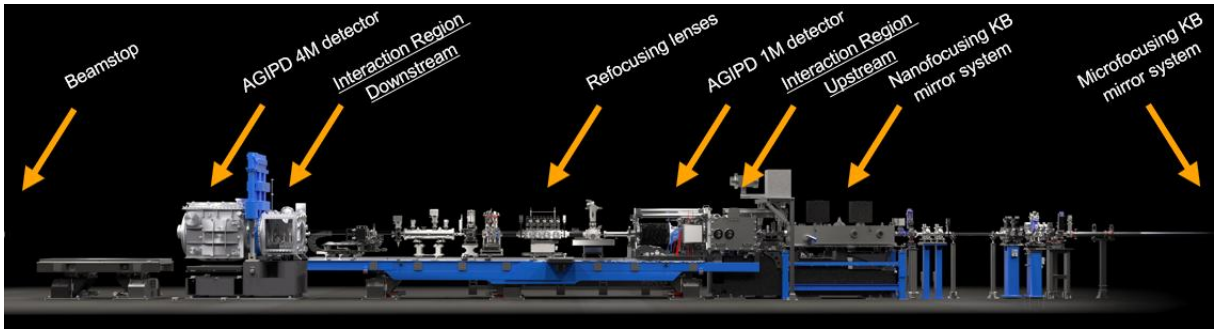


Figure 6: The SPB/SFX instrument at the European XFEL. Inside the experimental hutch, the X-ray beam is focused with either one of two mirror systems before reaching the interaction region upstream (IRU). Currently, experiments are performed solely at the IRU where data is collected with the AGIPD 1M detector. Once installation of the new interaction region downstream and the AGIPD 4M are done, experiments can also be performed there, after refocusing the X-ray beam from the IRU. Residual X-rays are captured behind the instrument with a beamstop. The image has kindly been provided by Dr. Raphaël de Wijn from SPB/SFX.

At the SPB/SFX instrument, displayed in Figure 6, experiments are currently only performed at the interaction region upstream (IRU) as the new downstream interaction region is under installation. The main employed sample delivery method for TR-SFX is liquid jets, as they are currently the only method that can replenish the sample fast enough to match the high X-ray repetition rate. However, due to the pulse delivery in trains, a large amount of the sample is not probed when a continuous jet is flown [3,4,6]. Therefore, additional methods like drop-on-demand injection and sample spread on a spinning disc are under investigation. While liquid jets can be employed for mix-and-inject as well as pump-probed experiments, reaction initiation by mixing does not allow a time resolution below a few milliseconds, and initiation by an optical pulse is the only suitable method to reach higher time resolutions (chapter 3.1).

The SPB/SFX instrument is equipped with different kinds of optical pump lasers with repetition rates ranging from 10 Hz to 4.5 MHz [27]. At the femto- and picosecond scale, the irradiated area does not significantly move, and the pump and probe pulses hit at the same position in the jet [27]. In the nanosecond range and slower, however, the pump laser needs to irradiate a defined area of the jet, which continues to move downstream until it is probed by the X-rays. Therefore, the delay time can be adjusted by varying the offset between the pump and the probing pulse, with respect to the jet's speed (Figure 7 A). To reach different delays, the synchronisation between the pump pulse and the X-rays is adjusted. Depending on the pump laser's pulse duration, focus, and repetition rate, specific time points in an activated reaction can be collected along with the native molecule state, resulting in pump-"on" and pump-"off" data. However, due to photon propagation through the jet, "off" data collected immediately after the "on" data can be contaminated and also show signs of laser triggering. Therefore, multiple "off" states are collected to capture the true dark state as well as the intermediates (Figure 7 B) [27]. The different "off" states can later be subtracted from "on" data to highlight changes in electron densities during refinement, which can specifically be helpful if only small structural changes or a low quantum yield are expected [4,9,27].

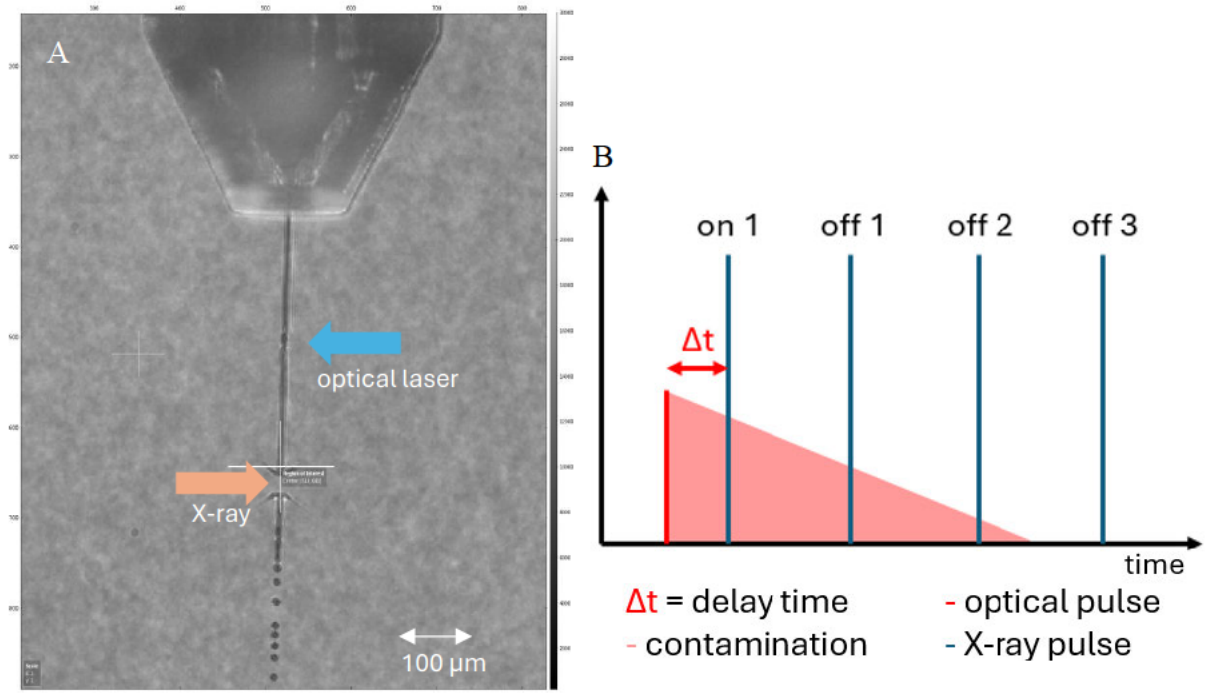


Figure 7: Pump-probed TR-SFX experiments with liquid jets. A: The reaction of a sample delivered by a liquid jet is triggered by an optical laser pulse (blue) and probed with an X-ray pulse (orange) after a defined delay time. B: Schematic illustration of the pulse repetition during a pump-probed TR-SFX experiment. The sample is triggered with an optical pulse (red) and collected as "on 1" data with an X-ray pulse (blue) after a defined delay time (Δt). Following X-ray pulses are used to collect "off" data. Sample probed immediately after the "on" state can be contaminated (light red) by the optical pulse, although not directly illuminated. Therefore, multiple "off" states are collected with increasing time to the optical pulse and decreasing contamination. Image A has kindly been provided by Dr. Mohammad Vakili.

During a pump-probed TR-SFX experiment, the sample is stored outside the sample chamber and provided through a nozzle into the sample catcher. There, the light-sensitive reaction is triggered by an optical laser pulse and probed with the X-rays (Figure 8 A). The scattered X-rays are then detected by the AGIPD 1M detector (Figure 8 B). The collected data is processed with the *EXtra-Xwiz* streamline, which combines hit-finding, indexing, and merging into a single streamlined workflow by utilising programs from the *CrystFEL* suite [17,28].

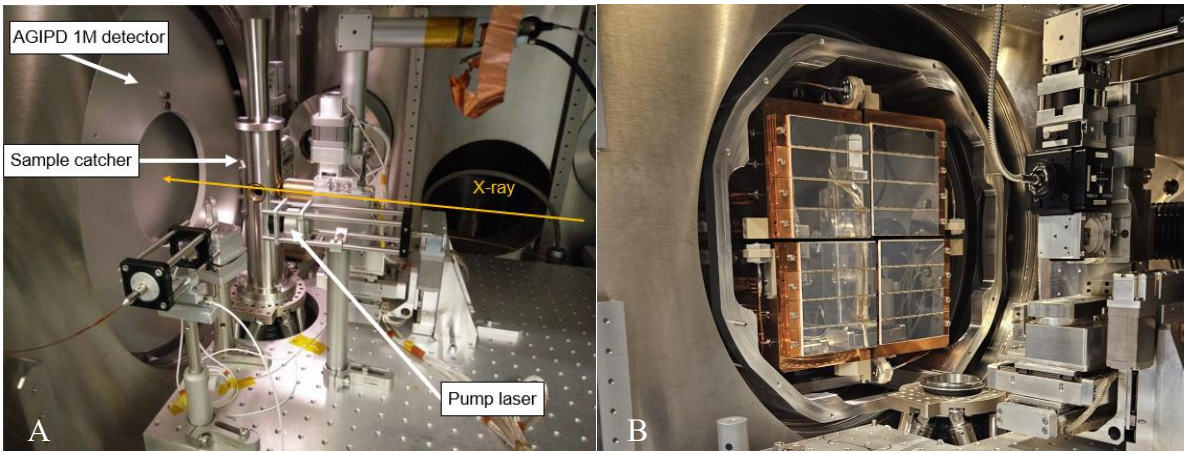


Figure 8: Inside of the IRU sample chamber. A: Sample chamber set-up for pump-probed TR-SFX experiments. A nozzle providing the liquid jet is inserted into the sample catcher. Inside the catcher, the sample is hit by the pump laser and the X-rays. Scattered X-rays are detected by the AGIPD 1M detector, located behind the gate valve. B: The AGIPD 1M detector is composed of 16 individual modules mounted on four quadrants movable in translation. Image A has kindly been provided by Dr. Raphaël de Wijn from SPB/SFX.

3.3. DNA as a photon-sensitive reporter system

Besides the careful planning of an experiment, validation of the experimental set-up is important to increase confidence in the results. Since data processing can be a time-consuming step, a validation system that does not require long processing and is easy to assess would be preferred.

Irradiation of DNA with ultraviolet (UV) light can lead to the formation of photoproducts in adjacent bases. The most abundant photoproducts are thymine-thymine (T-T) cyclobutene pyrimidine dimers (CPD) followed by thymine-cytosine (T-C) CPDs as well as T-C and T-T 6-4 photoproducts [29,30]. The structures of the T-T CPD and T-T (6-4) photoproducts are displayed in Figure 9. The formation of CPDs is a reversible reaction under further UV irradiation, establishing an equilibrium after a given radiation dose. Whereas the formation of (6-4) photoproducts is irreversible, leading to a constant increase or formation of a further product called Dewer isomer [30].

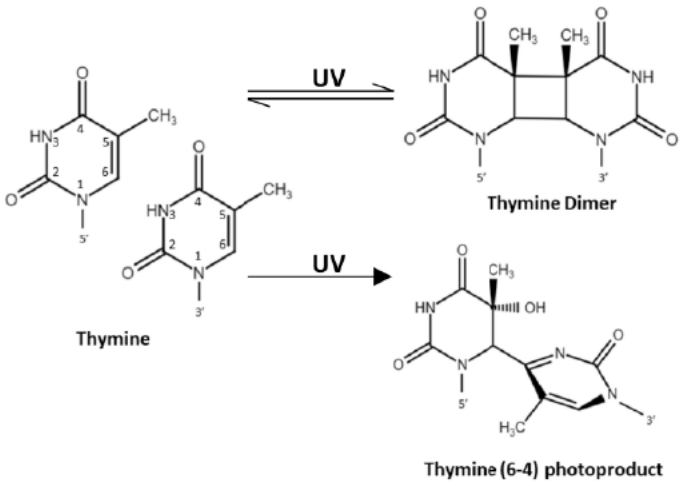


Figure 9: Structure of thymine photoproducts. Adjacent thymine bases form photoproducts upon UV irradiation. Most common is the reversible reaction to thymine cyclobutene pyrimidine dimers. The T-T (6-4) photoproduct is formed less frequently. However, the formation is irreversible **[30]**.

Photoproducts are formed because the C5=C6 double bond in pyrimidine is highly sensitive to UV irradiation and dissociates upon energy absorption. The close proximity in the DNA's base stacking then emphasises the formation of two single bonds between adjacent bases [29]. Park *et al.* investigated the structural changes of a DNA decamer when a T-T CPD is induced by synchrotron diffraction [31]. They found that in the dimer structure, the DNA unwinds by about 10° and a 30° bend is induced compared to the native structure (Figure 10). The change in the structure was also visible in a change in the unit cell parameters by 1 to 3 Å [31]. Since the cell parameters are calculated during the indexing of the diffraction data, the observed changes between the native and the dimerised DNA could potentially be exploited for a time-efficient reporter system that does not require phasing and refinement of the whole structure.

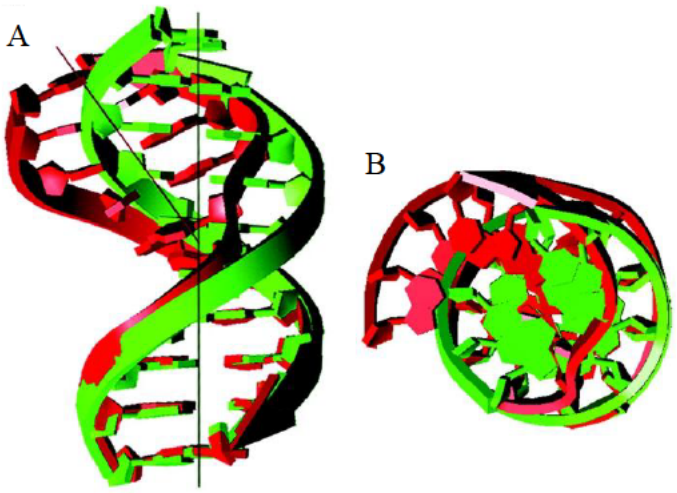


Figure 10: Schematic structure of native DNA and with a thymine CPD. The formation of photoproducts leads to the deformation of the DNA's native linear structure by inducing a kink. A: side view of a DNA oligo in its native (green) form and with a thymine CPD (red), B: top view **[31]**

3.4. HPBC cage compound for the controlled release of oxygen

As mentioned in chapter 3.1, mix-and-inject experiments cannot explore reactions on the sub-millisecond scale, whereas photoreactive systems can be probed at a femtosecond scale. However, only about 0.5% of all proteins are photoreactive [3,5]. To probe the remaining 99.5% with high temporal precision beyond the limits of ligand diffusion-based approaches, photocages, which have been studied across various scientific fields for several decades, are under investigation for their applicability in TR-SFX [5].

Photocages are photolabile protecting groups that are covalently bound to a reactant, thereby blocking it from chemical and physical reactions. Upon irradiation with a light pulse, the photocage is cleaved from the reactant and releases the active molecule. The cleavage of the photocage and the subsequent dissociation of the reactant determine the rate of reaction initiation [5,32]. For the high temporal resolution needed in TR-SFX at a nano- or femtosecond time scale, utilised photocages need to liberate the active compound on a faster time scale than the reaction of interest occurs in a sufficiently high concentration that the structural changes of the reaction are detectable [5,33,34]. To achieve a high reactant concentration, the cages need a sufficiently high quantum yield to release enough molecules while keeping a sufficiently low photocage concentration, so that all photocages can be irradiated. Furthermore, the absorption wavelength of the photocage should lie in a spectral region where the other reactants do not absorb light, ensuring selective excitation of the photocage [5]. In addition, the cages should be chemically resistant over time, and at the temperature at which the reaction is performed, so premature cleavage and reaction initiation are avoided [33,34]. However, many cages are sensitive to pH changes, and engineering the chemical stability often comes with a decrease in the quantum yield [5,32].

One of the most life-essential molecules is diatomic oxygen, which is involved in multiple biological processes, many of them occurring on a faster timescale than what is possible with mix-and-inject approaches [32,34]. To study oxygen-dependent processes, cobalt-based cages have been developed. The (μ-peroxo)(μ-hydroxo)bis[bis(bipyridyl)cobalt(III)] I (HPBC) complex (Figure 11) is one of them. HPBC cages diatomic oxygen between two bis(bipyridyl)cobalt(III) complexes, and releases the oxygen upon irradiation with 300-400 nm at a reported quantum yield of 0.03-0.06 in aqueous solution in less than 40 ns [32–34]. It has been reported that the cage degrades in aqueous solution, however, degradation occurred on a time scale ranging from several hours at room temperature to more than 24 h at 10°C [32]. Although HPBC cages have been characterised for use in time-resolved

crystallography, no studies reporting their application in time-resolved crystallography have been published at the time of writing, to the best of the authors' knowledge.

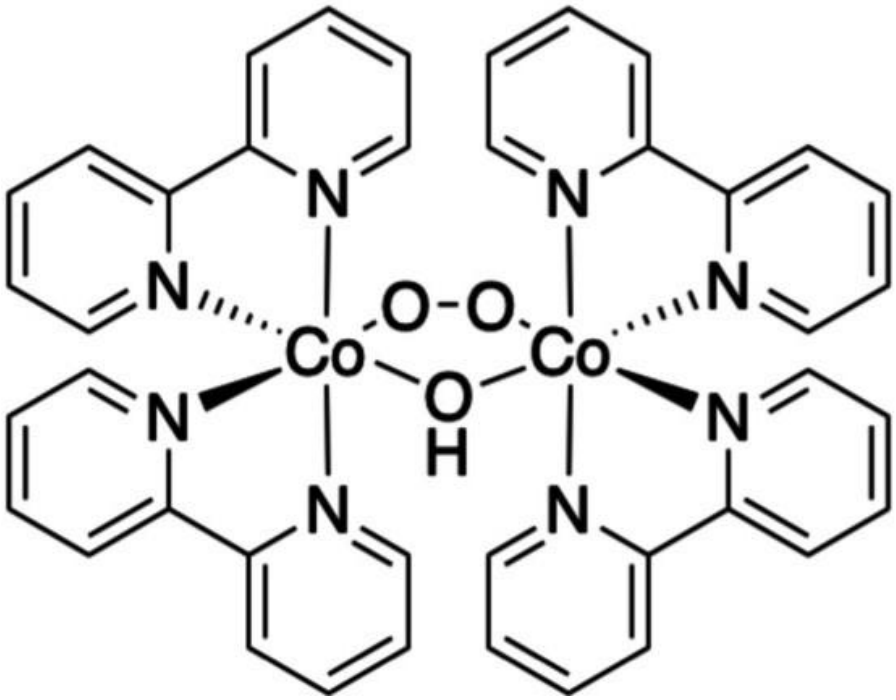


Figure 11: Structure of the HPBC oxygen cage. HPBC is a cobalt-based photocage that binds diatomic oxygen in its centre. Upon irradiation between 300 and 400 nm, the bond is cleaved, and diatomic oxygen is released. Image adapted from **[5]**.

To test the applicability of HBPC I in TR-SFX, a well-known system that binds oxygen is needed. Myoglobin (Mb) was the first protein whose three-dimensional structure was determined by X-ray crystallography in the 1950s [2]. Since then, it has served as a model system and has been extensively studied to investigate its structure with different ligands, as well as the dynamics of ligand binding and dissociation [35–38]. E.g. Vojtěchovský *et al*. solved the structures of deoxyMb, oxyMb, *aquomet*-Mb and carbonmonoxyMb (MbCO)at a resolution of 1.2 Å and extensively studied the spatial arrangement in the heme pocket of the different ligation states. In the study, it was shown that with different ligands, the iron of the heme moves in and out of the heme plane, and the distance between the iron and the histidine (His) to which the heme is bound changes (Figure 12 A) [36]. In addition, time-resolved dissociation studies have shown how the overall protein moves upon ligand dissociation (Figure 12 B), and how irradiation from pump lasers can influence the heme [37–39]. The intense research on myoglobin makes it an ideal test system to study HPBC in a TR-SFX experiment.

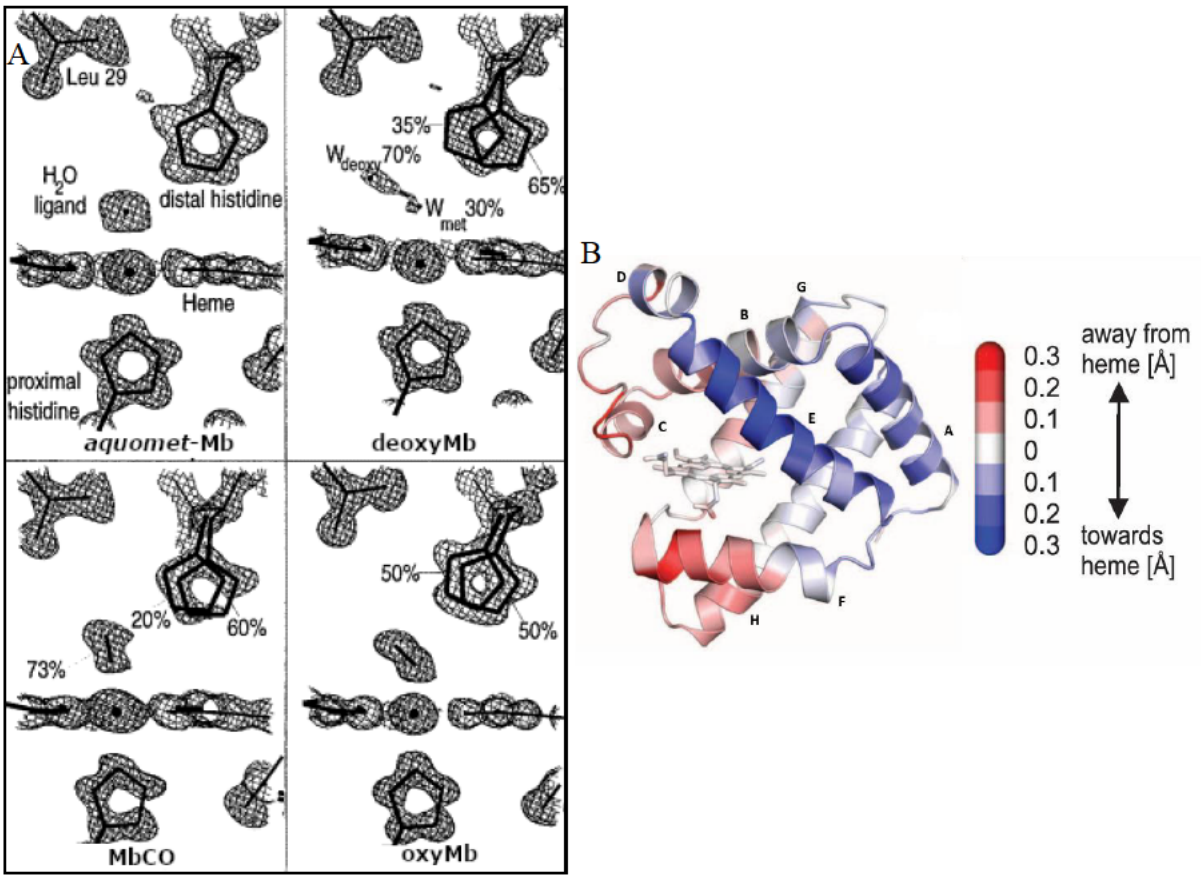


Figure 12: Different ligation states of myoglobin. A: Myoglobin can be found in four different states. *Aquomet*-Mb: water is bound to the heme, deoxyMb: one unbound water is found in the heme pocket and the iron moves out of the heme plane, displayed together with an 30% *aquomet*-Mb contamination, MbCO: carbon monoxide is bound to the heme, deoxyMb: oxygen is bound to the heme. All electron densities are displayed at a r.m.s.d of 1.80. Image adapted from **[36]**. B: Overall protein movement upon ligand dissociation. After carbon monoxide was photolysed from the heme group, the protein helices moved towards (blue) or away (red) from the heme. The strongest movement is found in the E moving towards the heme and F helices moving away from the heme. Image adapted from **[38]**

4. Materials and Methods

4.1.Materials

Table 1: Used devices & instruments. Listed below are the devices and instruments used in this study.

Device/Instrument	Specification	Article number	Supplier
Centrifuge	Centrifuge 5435, 24-place lab standard	5435	Eppendorf
Cryo Loops	18 mm Mounted CryoLoop™, 0.05-0.1 mm, 0.025-0.05 mm	-	Hampton Research
Dewar	CX100B-11M, 4.1 l	-	Worthington Industries
Electrophoresis chamber	neoChoice electrophoresis chamber Maxi	7-0219	NeoLab
Electrophoresis power supply	300 V, 2000 mA, 300 W	EV3020	Consort bvba
Gel imager	UVp Transilluminator	FI-26X	Analytik Jena
Heat block	ThermoMixer® C	5382000015	Eppendorf
Incubator cabinet	-10°C to +60°C	MIR-254-PE	PHC Europe
Micro balance	Secura® Semi Micro Balances	SECURA225D-1S	Sartorius
Milligram balance	Entris® milligram Balance with Internal Adjustment	BCE423I-1S	Sartorius
Pipetting robot	NT8®, 10 nl – 1.5 µl	V3	Formulatrix
Raman microscope	LabRam Odyssey, Confocal Raman & High-Resolution Spectrometer	-	Horiba
Sapphire slit	Sapphire Window Assembly Kit	334641	Beckman Coulter

Device/Instrument	Specification	Article number	Supplier
Stereo microscope	SDF PLAPO 1XPF objective, U-CMAD3 camera	SZX16	EVIDENT
ultrapure water system	Barnstead™ GenPure™ Pro	50131952	Thermo Fisher Scientific
UV/VIS Spectrophotometer	NanoDrop One^{C} microvolume-UV/VIS Spectrophotometer	ND-ONE-W	Thermo Fisher Scientific
Vortex	Vortex Genie 2	SI-0236	Scientific Industries

Table 2: Used laser and beamline. Listed below are the laser and beamline used in this study. Further specifications are given in the referenced papers.

Beamline or laser type	Wavelength	Photon or pulse energy	Specification
P13 EMBL beamline [40]	3.1-0.7 Å	4-17.5 keV	EIGER 16M detector
optical parametric oscillator (OPO) ns laser [27]	210-2200 nm	Up to 5 mJ	Up to 20 Hz repetition rate

Table 3: Used consumables. Listed below are the consumables used in this study.

Product	Specification	Article number	Supplier
Crystallisation plate	96-wells, 2 sub-wells, sitting drop standard	-	SWISSCI
Sealing foil	ClearVue™ Sheets, UV illumination friendly	MD6-01S	Molecular Dimensions

Table 4: Used chemicals. Listed below are the chemicals used in this study.

Name	Specification	Article number	Supplier
Agarose	DNase-, RNase free	3810	Carl Roth
DNA loading dye	Purple 6x,	B7024S	New England Biolabs
Ethidium bromide	0.025%, dropper bottle	HO47	Carl Roth
Ethylenediaminetetraacetic acid disodium salt (EDTA)	0.5 M pH 8.0, molecular biology grade	BU-105	Jena Bioscience
HEPES	1 M pH 7.5, molecular biology grade	BU-106-75	Jena Bioscience
HEPES sodium salt	PUFFERAN®, ≥ 99%	7020	Carl Roth
2-(*N*-Morpholino)-ethansulfonacid (MES)	PUFFERAN®, ≥ 99%	4256	Carl Roth
Paraffin oil	100%	HR3-411	Hampton Research
Polyethene glycol (PEG) 400	100%, Crystallisation grade	HR2-603	Hampton Research
PEG 2000 monomethyl ether (MME)	50% (w/v), Crystallisation grade	HR2-613	Hampton Research
PEG 3350	50% (w/v), BioUltra	88276	SIGMA Life Science
PEG 4000	50% (w/v), BioUltra	95904	SIGMA Life Science
PEG 6000	50 % w/v	CSS-255	Jena bioscience
PEG 8000	50% (w/v), BioUltra, for molecular biology	81268	SIGMA Life Science
Sodium acetate	for molecular biology, ≥ 99.0%	71183	SIGMA Life Science

Name	Specification	Article number	Supplier
Sodium chloride	≥ 99% Ph. Eur., USP	P029	Carl Roth
TRIS hydrochloride	PUFFERAN® ≥99 %, p.a.	9090	Carl Roth

Table 5: Used buffers. Listed below are the names and compositions of the buffers used in this study.

Buffer		Components	Concentrations
DNA storage buffer		Tris-HCl	10 mM
		EDTA	1 mM
			pH 8.0
10-mer crystallisation buffer			
	10.1	NaCl	0.2 M
		HEPES	0.1 M
		PEG 3350	25% (w/v)/ 35% (w/v)/ 45% (w/v)
			pH 7.5
	10.2	Sodium acetate	0.1 M
		PEG 4000	25% (w/v)/ 35% (w/v)/ 45% (w/v)
			pH 4.6
	10.3	HEPES sodium salt	0.1 M
		PEG 4000	25% (w/v)/ 35% (w/v)/ 45% (w/v)
			pH 7.5
20-mer crystallisation buffer			
	20.1	MES	0.1 M
		PEG 2000 MME	25% (w/v)/ 35% (w/v)/ 45% (w/v)
			pH 6.5
	20.2	MES	0.1 M
		PEG 8000	25% (w/v)/ 35% (w/v)/ 45% (w/v)
			pH 6.5
	20.3	HEPES sodium salt	0.1 M
		PEG 6000	25% (w/v)/ 35% (w/v)/ 45% (w/v)
			pH 7.5

Table 6: Used DNA strands. Listed below are the single-stranded DNA sequences used in this study.

Oligo	Sequence	Supplier
10-mer	5'-GCGTAATGCG-3'	Eurofins Genomics
	3'-CGCATTACGC-5'	Eurofins Genomics
20-mer	5'-GAGCTGCGTAATCGGTCGTG-3'	Eurofins Genomics
	3'-CTCGACGCATTAGCCAGCAC-5'	Eurofins Genomics

Table 7: Used software. Listed below is the software used in this study. Scientific, non-commercial programmes are cited with their reference paper.

Software	Version	Provider
coot [41]	0.8.9.1	MRC Laboratory of Molecular Biology
CrystFEL [28]	12:0	DESY
EXtra-Xwiz [17]	Alpha	European XFEL
LabSpec 6	6.7.2.11	Horiba
phenix [16]	1.21.1	Phenix
PRECiV Core	2.2.33723.0	EVIDENT Europe
XDS [42]	20250430	Dectris/Max Planck Society

4.2. Preparation of crystallisation samples

The single-stranded DNA oligomers (Table 6) were resuspended to 6 mM in DNA storage buffer (Table 5). Annealing was performed by mixing the single-stranded 10-mers and 20-mers to a final concentration of 3 mM, respectively, heating the samples to 80°C for 15 min in a heat block (eppendorf), followed by incubation at room temperature overnight. The next day, 2 µl of each sample was used to measure the concentrations via UV spectrometry (chapter 4.3). The double-stranded oligos were stored at room temperature.

4.3. UV spectrometry

UV absorbance spectra were recorded to measure DNA concentration and to analyse the dimerisation of thymine. 2 µl of each sample was pipetted onto the measurement stage of the NanoDrop (Thermo Fisher Scientific). The absorbance was measured from 220 nm to 350 nm. To compare the dimerisation, the spectra were normalised to the 230 nm absorption, and the 257-to-230 nm ratio was calculated to evaluate the dimerisation.

4.4. Crystallisation of DNA oligomers

Crystallisation of the double-stranded 10-mer and 20-mer DNA oligos (Table 6) was performed by sitting-drop vapour diffusion in 96-well crystallisation plates (Table 3). The plates were prepared by pipetting 50 µl of the respective crystallisation buffer (Table 5) into the reservoir well. DNA samples were diluted with crystallisation buffers to a final concentration of 1.5 mM, 2.0 mM, 2.25 mM and 2.5 mM and dispensed onto the plate by the NT8® pipette robot (Formulatrix) in 300 nl for 1.5 mM, 2.0 mM and 2.25 mM and 450 nl for 2.5 mM. The plates were sealed with adhesive crystallisation foil (Table 3) and stored in a temperature-controlled cabinet (PHC Europe) at 20°C. The formation of crystals and their morphology were evaluated weekly using a stereo microscope (EVIDENT). In addition, the size of the crystals was measured with *PRECiV Core* image processing software (Table 7).

4.5. Agarose gel electrophoresis

To validate that the crystals consist of DNA rather than buffer components, agarose gel electrophoresis was performed. For each crystallisation buffer, multiple crystals from different DNA-to-buffer ratios were fished under the stereo microscope (EVIDENT) and washed in their respective crystallisation buffers (Table 5). After washing, the crystals from buffers with varying DNA concentrations were dissolved together in 26 µl of the DNA-storage buffer (Table 5). 1 µl of each dissolved sample was used to measure the DNA concentration (chapter 4.3). The remaining 25 µl were mixed with 5 µl of 6x loading dye (Table 4), and 25 µl of each sample were loaded onto a 3% agarose gel containing one drop of 0.025% ethidium bromide solution (Table 4). The gel was run at a constant 100 V for one hour and illuminated with the UVp transilluminator (Analytik Jena) to visualise the bands.

4.6. UV irradiation

To dimerise the adjacent thymine bases, samples were irradiated with UV light at 257 nm using the OPO nanosecond laser (Table 2) at 0.2 mW/cm². Crystals were irradiated by placing the respective well of the crystallisation plate under the laser. Liquid samples were irradiated by pipetting 2 µl into a crystallisation well, next to a reservoir well filled with 50 µl DNA storage buffer (Table 5), which was then sealed with adhesive crystallisation foil to minimise evaporation (Table 3) and placed under the laser beam. The samples were irradiated for 3, 30, 60 or 90 minutes, resulting in 36, 360, 720 or 1080 mJ/cm², respectively. To assess whether the thymine bases dimerised, UV-absorption spectra were measured (chapter 4.3), and Raman spectra (chapter 4.7) were acquired.

4.7. Raman spectrometry

To collect Raman spectra, 0.5 µl of a DNA sample was dried on a sapphire slit (Beckman Coulter) and excited by the 633 nm laser of the LabRam Odyssey Raman microscope (Horiba) at 100% laser power for 30 s. The Raman spectra were recorded from 1150- 1750 cm^{-1} with a 600 nm grating using a 100× objective. Each spectrum was measured as a duplicate. Background subtraction and baseline correction, as well as normalisation to 100 counts at the 1370 cm^{-1} were performed using *LabSpec 6* (Table 7). The collection and processing of the Raman spectra was performed with the assistance of the Sample Environment and Characterisation group of the European XFEL.

4.8. Single crystal diffraction

Crystals were grown under selected conditions (chapter 4.4) and irradiated with UV light to induce dimerisation (chapter 4.6). Single irradiated and unirradiated crystals were fished individually, cryocooled and stored in liquid nitrogen dewars (Worthington Industries). Crystals grown in buffers containing 25% PEG (w/v) (Table 5) were washed in paraffin oil (Table 4) to prevent formation of ice crystals during cryocooling. Diffraction data were collected at the P13 beamline of the European Molecular Biology Laboratory (EMBL) at the *Deutsches Electronen Synchrotorn* (DESY) (Table 2) at a beam energy of 13.2 keV (0.9393 Å). The collection was conducted in air, while the mounted crystal was cooled by a liquid nitrogen stream. All crystals were screened for diffraction by taking 4 images with 1.00° rotation at 2.58 Å detector cut-off resolution for a total exposure time of 0.4 s. If the screening showed a potential diffraction signal, full data sets of 360 or 3600 images, with a respective rotation of 1.00° or 0.10°, were collected at 1.2, 1.5 or 2.58 Å. Collected diffraction data were automatically processed with *XDS* (Table 7) using standard settings of the P13 beamline [40]. Offline processing with *XDS* (Table 7) was done using the provided default settings, with the exclusion of untrustworthy resolution ranges [42]. No space group or unit cell parameters were provided for the processing, as they were unknown for the system under study.

4.9. Indexing of serial crystallography diffraction data

Diffraction data collected at the SPB/SFX instrument of the European XFEL were indexed using the automated *EXtra-Xwiz alpha* streamline (Table 6). For hit-finding and indexing *peakfinder8* and *xgandalf* were utilised, both integrated in *indexamajig* from *CrystFEL*. Splitting of pumped samples and merging of reflections were done by *partialator* of *CrystFEL*. Figures of merit were created by *compare_hkl*. In a first instance, all datasets were processed at the detector's resolution limit of 1.5 Å. The completeness, signal-to-noise ratio, $CC_{1/2}$, and R_{split} were evaluated, and the datasets were reprocessed at resolution cutoffs that resulted in completeness >99%, SNR >1, and CC1/2 >30% in the highest-resolution shell.

4.10. Phasing and refinement

The initial structure models were obtained by molecular replacement using the program *Phaser* [43] integrated in the *Phenix* suite (Table 6). The models were phased using the fast mode with 100% sequence identity and inclusion of all possible alternative space groups within the same point group during the search. As a reference model for phasing the myoglobin structures, the Protein Data Bank entry 1WLA from recombinant wild-type horse heart myoglobin was used.

Refinement of the initial model was conducted with *phenix.refine* using cartesian simulated annealing, reciprocal- and real-space as well as B-factor refinement in the first step. Afterwards, further refinement was performed in a repetitive cycle of real-space, rotamer and Ramachandran adjustments in *coot* (Table 6) and reciprocal-, real-space and B-factor refinement with *penix.refine*. Once a structure was refined to sufficient quality, hydrogens and other solvent molecules were added and adjusted with *phenix.refine* and *coot,* respectively. After each round of refinement, the structure's quality was validated using the *MolProbity* statistics integrated in *phenix.refine*.

4.10.1. Ligand identification and occupancy determination

To determine the state of the myoglobin and the ligand's occupancy, multiple approaches were used:

I. Isomorphous difference maps

If the refined difference density map (F_{obs}-F_{calc}) did not show a distinctively identifiable positive electron density in the heme pocket, isomorphous difference maps were created using the 'Isomorphous difference map' feature in *phenix* (Table 6). The observed structure factors of the refined deoxyMb structure were subtracted from the observed structure factors of the refined triggered data. The refined deoxyMb served as the structure for phasing.

II. *phenix* occupancy refinement

A single diatomic oxygen molecule or a water and a diatomic oxygen molecule as alternating conformations were modelled into the density maps in *coot* (Table 6) and the structures were refined using *phenix.refine* with occupancy and B-factor refinement to calculate the occupancy of the ligands.

III. Heme geometry

The geometry of the heme was examined in *coot* (Table 6) to evaluate the refined occupancies. The distance between the proximal histidine (His93) and the iron of the heme, as well as the distance between the iron and the centre of the porphyrin ring's nitrogen atoms were measured.

IV. Helix movement

In addition to the heme geometry, the overall movement of the myoglobin's tertiary structure was assessed. The distance of the atomic position of each C_α -atom to the centre of the porphyrin ring's nitrogen atoms was calculated for every structure. The distances were then subtracted by the respective distances of the deoxyMb to compare the helix movements towards or away from the heme.

5. Results

5.1. Preliminary tests for a DNA reporter system

5.1.1. DNA crystallisation screening

To find a condition in which the DNA crystallises into diffracting crystals in a reasonable time, a screening of buffers with different salts and crystallisation agents was performed. The tested conditions were derived from a prior screening of several buffers and aimed to improve those that had previously yielded promising crystals. In total, 78 conditions were tested. For both the 10-mer and 20-mer, three buffers (Table 5) with polyethene glycol concentrations of 25% (w/v), 35% (w/v), and 40% (w/v) and DNA concentrations of 1.5 mM, 2.0 mM, 2.25 mM and 2.5 mM were tested. The preparation of the DNA samples was conducted according to chapter 4.2 and yielded 26.39 µg/µl (3.998 mM) and 47.76 µg/µl (3.618 mM) for the 10- and 20-mer, respectively. Crystallisation was performed as described in chapter 4.4. All conditions were tested in duplicate.

Increasing the DNA and PEG concentration generally reduced the time until crystals were visible, though the magnitude of the decrease depended on the buffer. The longest incubation time was recorded for the 20-mer in buffer 20.2 at 25% (w/v) PEG (Figure 13, 20.2), while the shortest was one day for the 20.1 buffer at 40% (w/v) PEG (Figure 13, 20.1). On average, the incubation time was 61.5 days at 25% (w/v) PEG, 31.9 days at 35% (w/v) PEG, and 21.8 days at 40% (w/v) PEG.

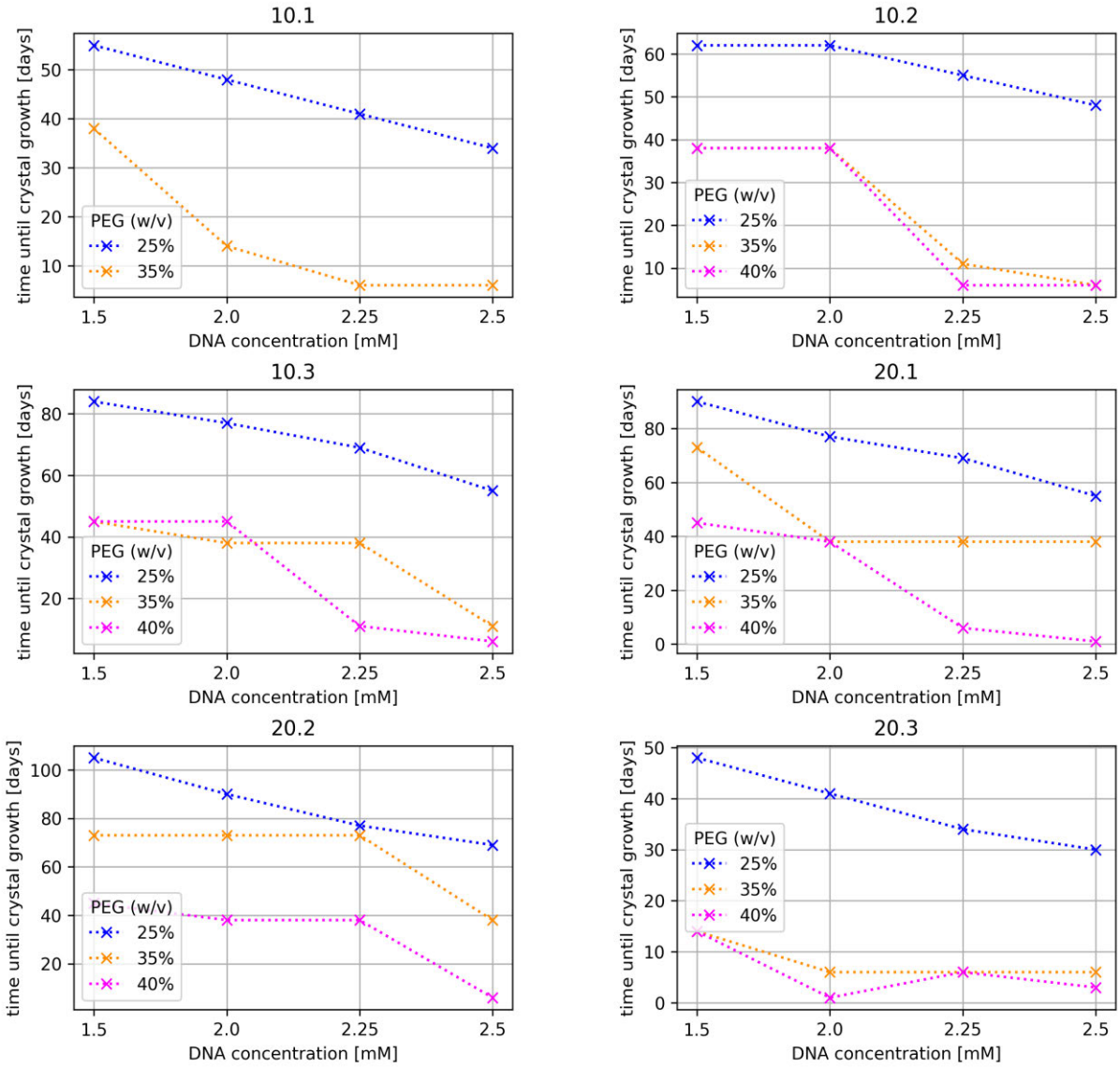


Figure 13: Time duration until crystal growth dependent on crystallisation condition. Displayed is the time required for crystals to become visible using a stereo microscope in days vs the DNA concentration in mM, depending on the crystallisation buffer and PEG concentration in % (w/v). 10.x graphs show the crystallisation of 10-mer oligos, 20.x display the 20-mers. The compositions of each crystallisation buffer 10.1-20.3 are listed in Table 5. The buffers were tested with 25% (w/v) PEG (blue), 35% (w/v) PEG (orange), and 40% (w/v) PEG (pink).

In addition to the crystal formation, the morphology of the crystals was assessed by measuring their sizes and evaluating their shapes (chapter 4.4). Crystals were considered regular-shaped when they showed a symmetric appearance with straight edges (Figure 14 A, B). Plates and crystals with a rough surface or bent sides were considered irregular-shaped (Figure 14 C-F).

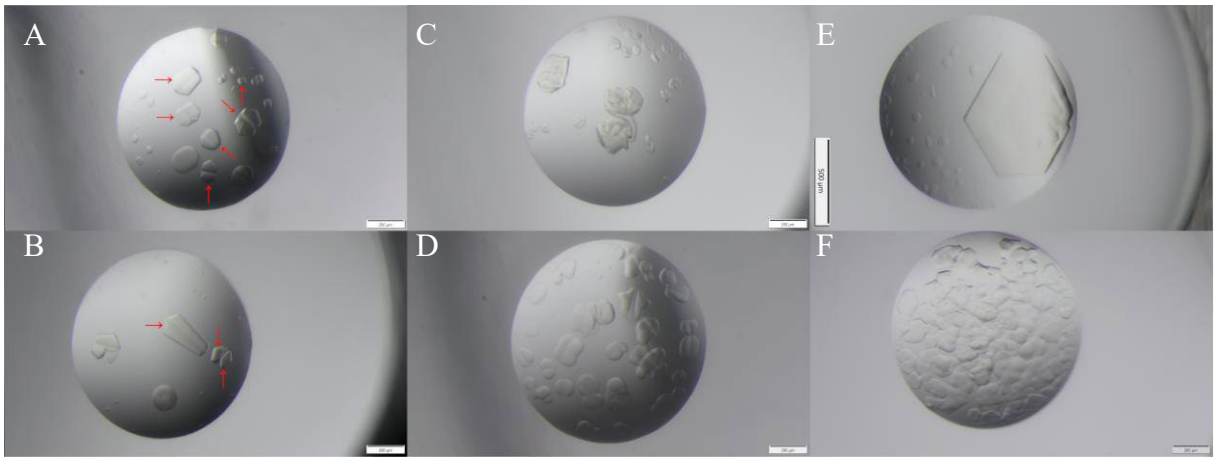


Figure 14: Examples of regular- and irregular-shaped crystals. Regular-shaped crystals are marked with →. A: Regular-shaped crystals grown in buffer 10.1 with 35% (w/v) PEG at 2 mM DNA, B: regular-shaped crystals grown in buffer 10.2 with 35% (w/v) PEG at 2 mM DNA, C: irregular-shaped crystals grown in buffer 10.2 with 40% (w/v) PEG at 2.25 mM DNA, D: irregular-shaped crystals grown in buffer 20.1 with 35% (w/v) PEG at 2.5 mM DNA, E: plates grown in buffer 20.2 with 25% (w/v) PEG at 2.0 mM DNA, F: overgrown crystals in buffer 20.2 with 40% (w/v) PEG at 2.0 mM DNA.

Comparing crystal shapes based on the PEG concentration in the buffers shows that at 40% (w/v) PEG, 5 out of 20 buffers produced crystals with regular shapes. At 35% (w/v) PEG, 14 out of 24 conditions and at 25% (w/v) PEG, 13 out of 24 conditions yielded regular-shaped crystals. Comparison by the DNA concentrations shows that 3, 8, 9, and 12 of 17 conditions at 2.5 mM, 2.25 mM, 2.0 mM, and 1.5 mM DNA, respectively, resulted in regular-shaped crystals (Figure 15). The crystal size ranges did not show a general trend correlated with the buffer conditions.

Though crystals grew the fastest in buffers with 40% (w/v) PEG, only one of the six buffers produced regular-shaped crystals. Therefore, only crystals from 25% and 35% (w/v) PEG were used for further experiments. Due to the long incubation time at 25% (w/v) PEG and at a DNA concentration of 1.5 mM, additional crystals were produced at 35% (w/v) PEG at 2.0 mM DNA.

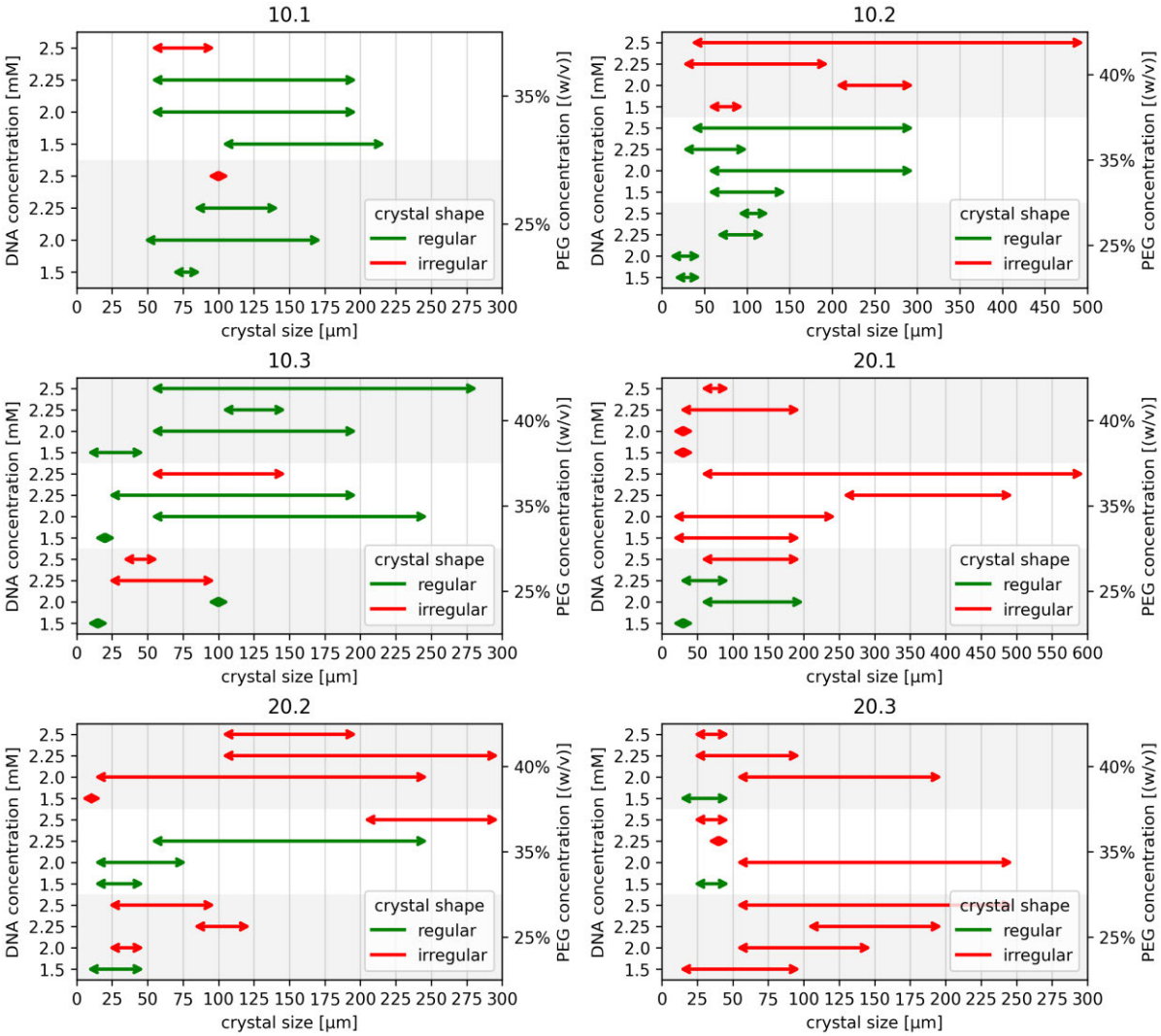


Figure 15: Crystal sizes and shapes dependent on the crystallisation condition. Multiple crystals of each condition were measured along their longest axis. The size ranges in µm are plotted against the DNA and PEG concentrations in mM and (w/v), respectively. 10.x graphs show the morphology of 10-mer oligos, 20.x display the 20-mers. The compositions of each crystallisation buffer 10.1-20.3 are listed in Table 5. Green: regular-shaped crystals, red: irregular-shaped crystals.

To verify that the crystals consisted of DNA, crystals from the 35% (w/v) PEG buffers were collected and dissolved in DNA storage buffer (Table 5). The concentration of the solutions was then measured (chapter 4.3), followed by agarose gel electrophoresis (chapter 4.5). The 10-mer and 20-mer stock solutions served as positive controls, while ultrapure water served as a negative control. In all samples, DNA was measured with the NanoDrop (Thermo Fisher Scientific), and all solutions showed faint bands on the agarose gel (Figure 16). Therefore, all samples could be confirmed as DNA crystals. Besides the 20.2 sample, which showed the weakest band while measuring the highest sample concentration of 3.03 µg/µl, the band intensities correlated with the measured DNA concentrations.

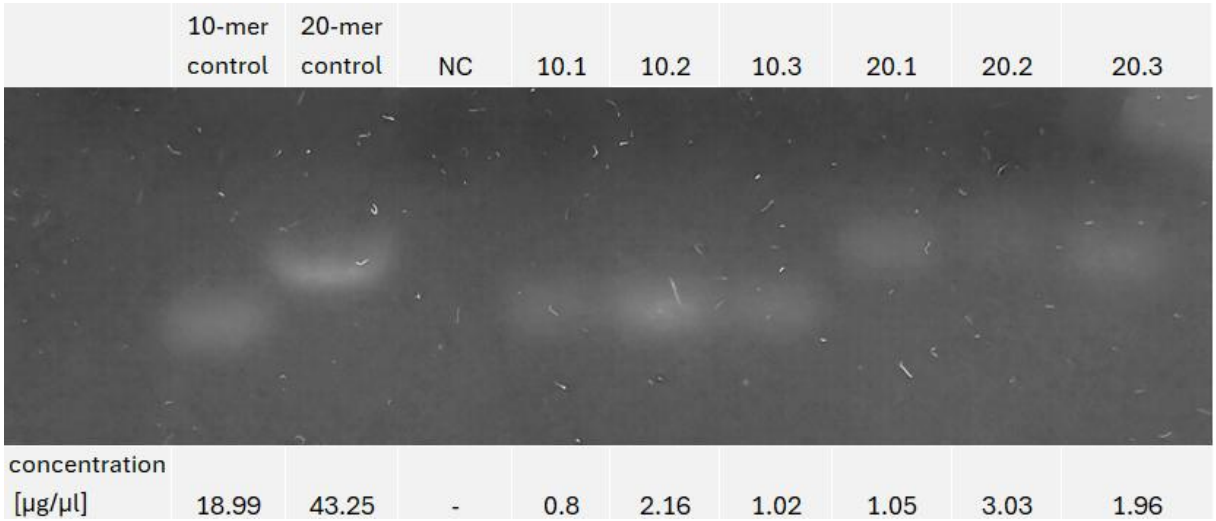


Figure 16: Agarose gel electrophoresis of crystals grown in buffers with 35% (w/v) PEG. Crystals from each DNA concentration of the respective buffer with 35% (w/v) PEG were collected, washed and dissolved together. Samples from left to right: 50 bp ladder, 10-mer stock solution, 20-mer stock solution, ultrapure H_2O, dissolved 10-mer crystals from buffer 10.1, 10.2 and 10.3, dissolved 20-mer crystals from buffer 20.1, 20.2 and 20.3.

5.1.2. Thymine dimerisation tests

A new 10-mer solution was prepared (chapter 4.2) with a concentration of 20.94 µg/µl (3.17 mM). To dimerise the adjacent thymine bases, samples were irradiated with UV light at 257 nm with total doses of 36, 360, 720, or 1080 mJ/cm². For each dose, three 2 µl samples were prepared as described in chapter 4.6 and analysed by recording UV-absorption spectra (chapter 4.3) in triplets and Raman spectra (chapter 4.7) in doublets. In addition, Raman spectra of an unirradiated crystal, of one crystal irradiated at 36 mJ/cm² and one irradiated at 1080 mJ/cm² were collected. However, due to limited sample amounts, establishing a workflow to collect comparable results from crystals was not possible.

In the normalised and averaged UV-absorption spectra (Figure 17 A, B), the samples exposed to the higher doses of 720 and 1080 mJ/cm² showed noticeable decreases at the absorption maximum of 257.5 nm. Specifically, compared with the unirradiated control, they decreased by 0.58% ± 0.74% and 1.91% ± 1.38%, respectively (Figure 17 C). Meanwhile, the changes in the samples irradiated at 3.6 and 36 mJ/cm² were 0.19% ± 0.28% and -0.22% ± 0.16%, respectively, which fell within the standard deviation of the unirradiated sample of 0.24%.

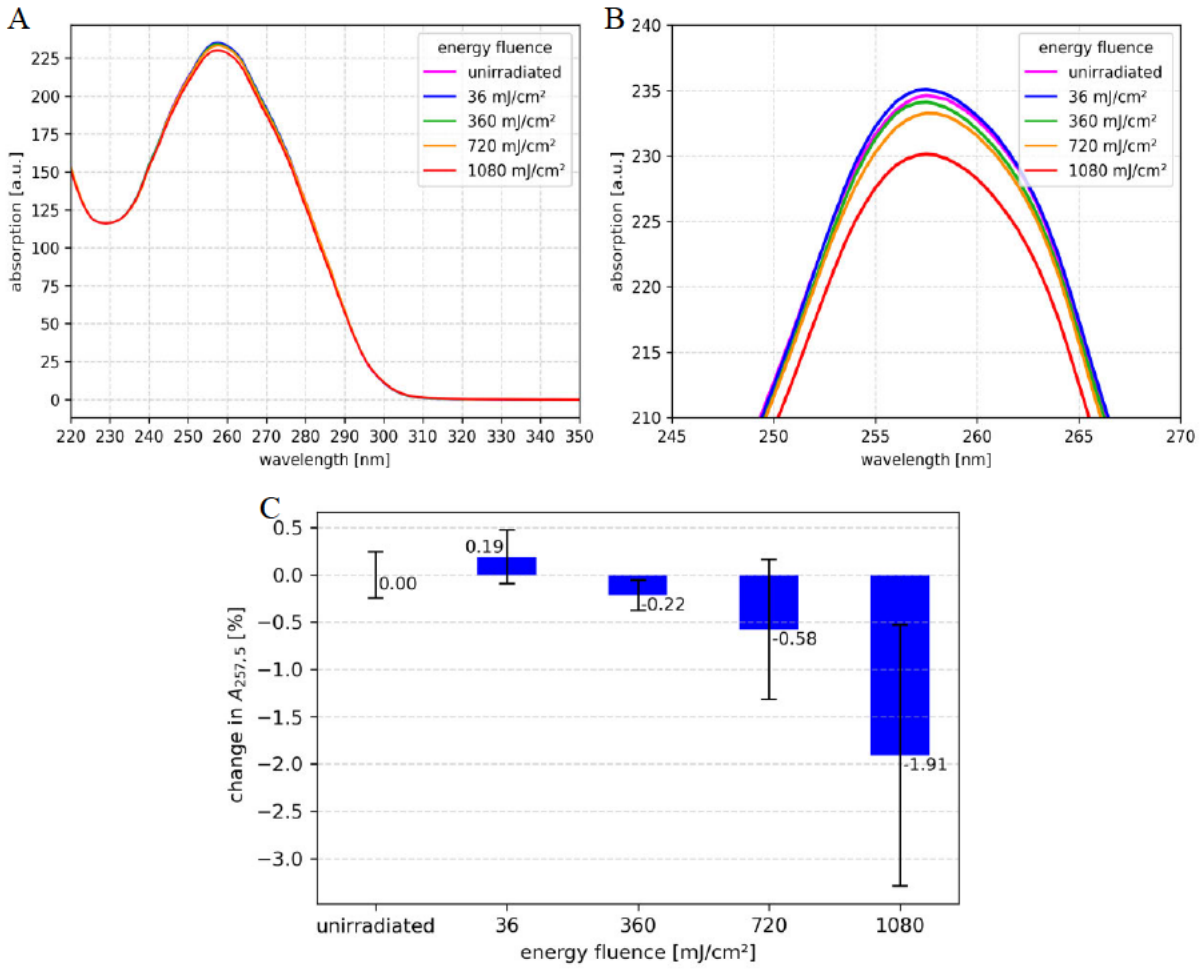


Figure 17: UV-absorption spectra of irradiated and unirradiated samples. 2 µl of DNA in the DNA storage buffer were irradiated with 36 mJ/cm² (blue), 360 mJ/cm² (green), 720 mJ/cm² (orange) or 1080 mJ/cm² (red), UV spectra from 220-350 nm were taken normalised to A_{230}. Unirradiated DNA (magenta) was measured as a reference. All samples were measured as triplets. A: 220-350 nm spectra, B: close-up of the UV spectra in the 245-270 nm range, C: change in $A_{257,5}$ compared to $A_{257.5_unirradiated}$ with standard deviation. Original data can be found in the appendix (Figure 25).

In the Raman spectra, the overall intensities of the samples irradiated at 360 mJ/cm² or higher decreased, whereas at 36 mJ/cm² the overall intensities increased compared to the unirradiated sample (Figure 18, A). At 1250 cm^{-1}, 1470 cm^{-1}, and 1675 cm^{-1}, strong changes in the peak intensities were visible, while at 1420 cm^{-1} and 1570 cm^{-1}, medium changes were visible. At the 1375 cm^{-1} peak, only the 1080 mJ/cm² irradiated samples showed a weak visible change. The changes in intensity in the thymine-associated peaks dependent on the irradiation dose are plotted in Figure 18 B. The figure shows that at the 1250 cm^{-1} and 1675 cm^{-1} peaks, the intensities first increase with an irradiation dose of 36 mJ/cm² and then decrease at 360 mJ/cm². At higher doses, the intensities increased again but stayed below the level of the unirradiated sample. The 1375 cm^{-1} peak showed the opposite behaviour, however, the changes were overall small within ten counts.

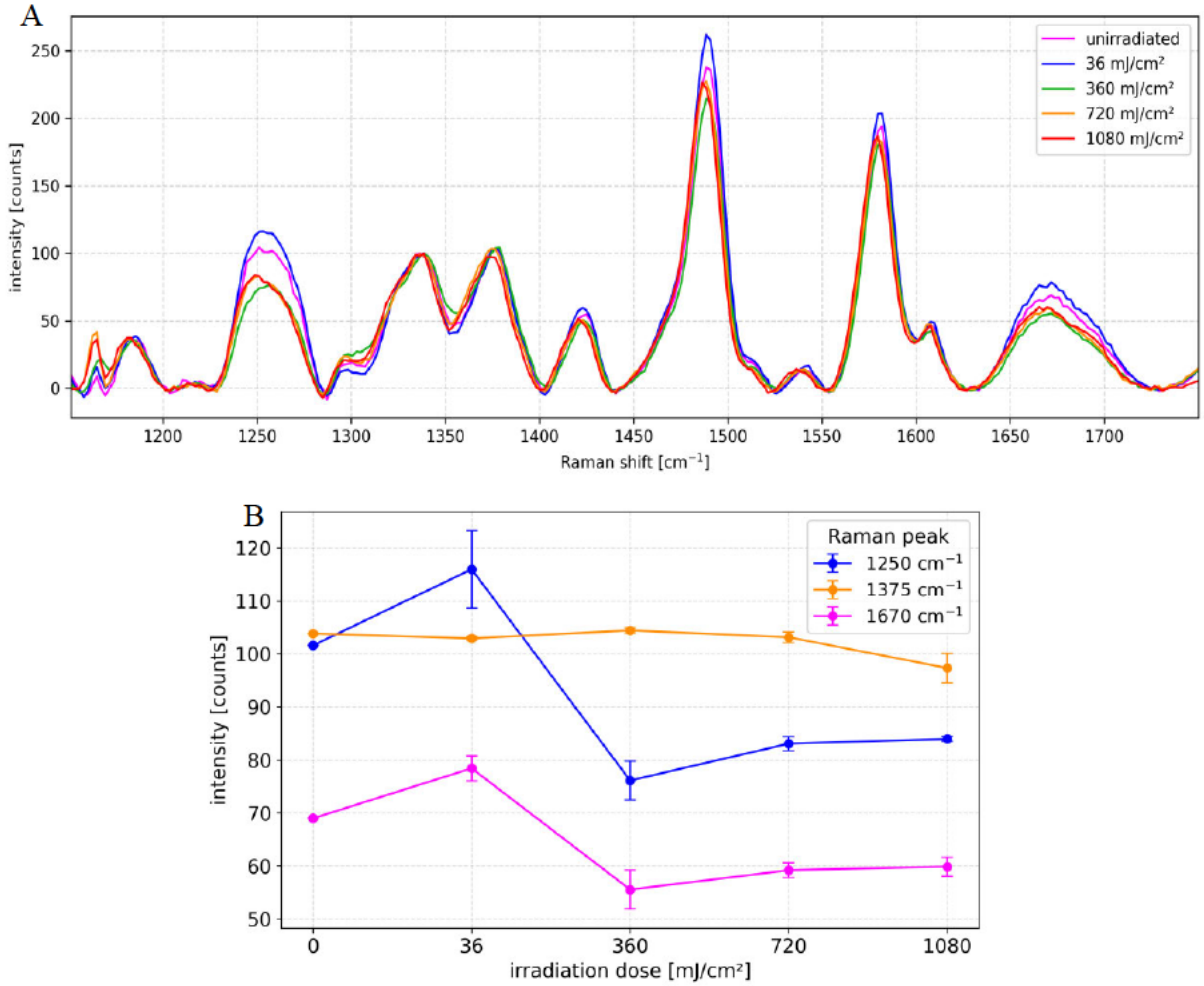


Figure 18: Raman spectra and change in the peak intensities. Raman spectra were collected in duplicates, baseline- and background-subtracted, and normalised. A: Full collected spectra of unirradiated (magenta) and irradiated at 36 mJ/cm² (blue), 360 mJ/cm² (green), 720 mJ/cm² (orange) and 1080 mJ/cm² (red) 10-mer DNA samples. B: Intensities of thymine-associated peaks at 1250 cm^{-1} (blue), 1375 cm^{-1} (orange) and 1670 cm^{-1} (magenta) plotted against the irradiation dose. Meaurements of all Raman samples can be found in the appebdix (Figure 26).

5.1.3. Collection of diffraction data

48 DNA crystals from various conditions were fished to collect diffraction data from synchrotron radiation, with 29 unirradiated and 19 irradiated crystals in a size range of 10-50 µm. All crystals were screened for diffraction, and full-rotation data sets were acquired from 18 crystals (chapter 4.8). The used crystals and acquisition conditions are listed in Table 11 and Table 12 for unirradiated and irradiated crystals, respectively.

During the screening, the crystals showed little to no Bragg peaks. Instead, single diffraction images showed rings, arches, or clouds around 4 Å (Figure 19 A-C), a few strong peaks around 25 Å (Figure 19 D), or combinations of these features. No automatic processing of the collected data sets was possible (chapter 4.8). The data set at 1.5 Å from a crystal irradiated for 60 min (Table 12) was used for additional offline processing with *XDS*, as the single diffraction image (Figure 20 A) showed the most diffraction spots among all data sets. Excluding signal from 4.1-3.8 Å, 905 out of 7603 hits (11.9%) were indexed (Figure 20 B). The suggested cell parameters with the highest quality of fit were a = 183.6 Å, b = 216.2 Å, c = 370.7 Å, $\alpha = 100.6°$, $\beta = 93.6°$ and $\gamma = 108.9°$ in the P1 space group. Due to the low indexing rate, no further processing was possible.

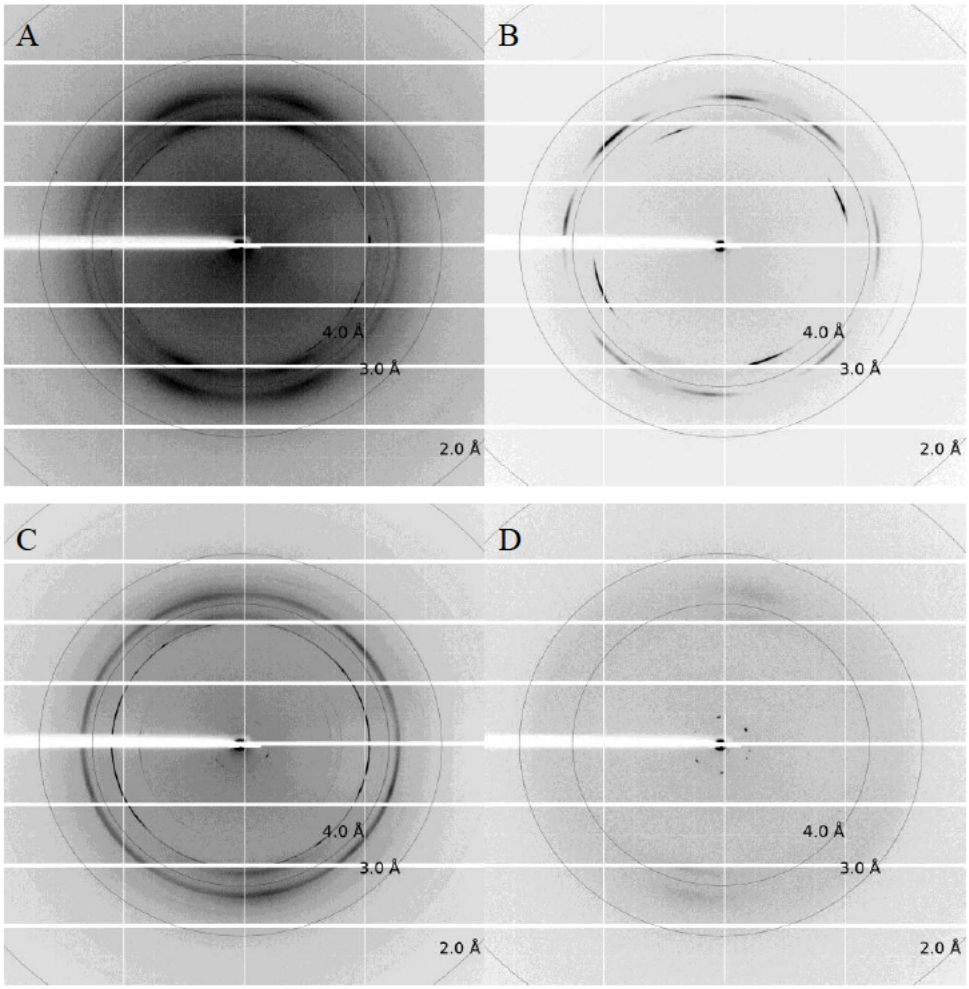


Figure 19: Single diffraction images of DNA crystals during screening. Displayed are single diffraction images of four different crystals collected during the screening and full-rotation collection. All crystals were grown under the same conditions in buffer 10.1 at 35% (w/v) PEG at 2.0 mM DNA. A: 60 min irradiated crystal with X-like, diffuse diffraction around 4 Å, B: unirradiated crystal with diffraction arches around 4 Å, C: 60 min irradiated crystal with diffraction rings around 4 Å, D: 90 min irradiated crystal with diffraction spots around 25 Å.

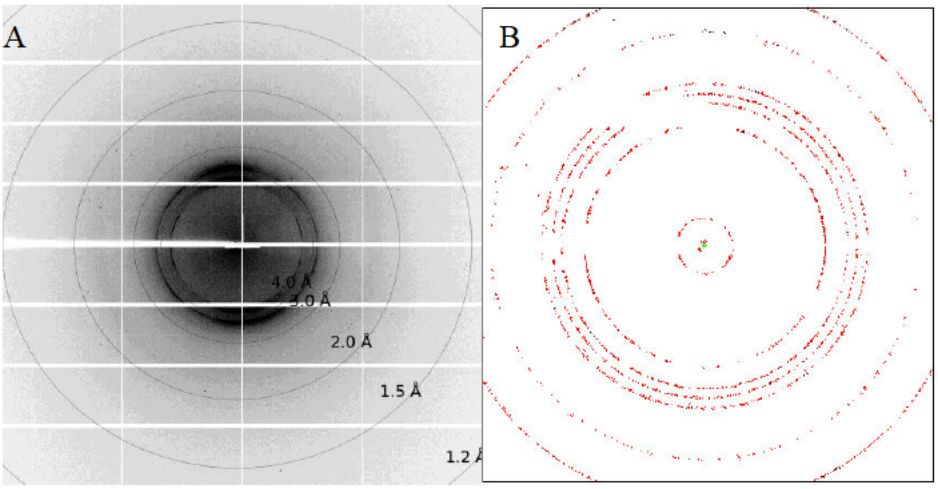


Figure 20: Offline processed diffraction data images. The diffraction data were collected during a full-rotation data collection of a 10-mer crystal grown in buffer 10.1 at 35% (w/v) PEG at 2.0 mM DNA, irradiated for 60 minutes. A: Single diffraction image, two rings and diffused clouds are visible around 4 Å, while Bragg peaks go up to 1.5 Å. B: Indexed (blue) vs. non-indexed (red) hits, with exclusion of 4.1-3.8 Å data.

5.2. Data analysis of the HPBC oxygen cage

Two distinct preparation strategies were used to assess the applicability of the HPBC oxygen cage. In the first experiment, deoxyMb and HPBC were mixed in advance and delivered from a shared reservoir through a liquid jet. One data set without laser triggering was collected as a reference. A second set was pumped with 300 mJ/cm^2 at 343 nm and probed after 100 ns. In the second experiment, the deoxyMb and HPBC were delivered by a mix-and-inject approach with a mixing time of 15 ms, before triggering the reaction with the pump laser at 343 nm (chapter 3.1). One reference data set without laser triggering was collected. Triggered data sets were collected at 194 mJ/cm^2 after 10 ns, 194 mJ/cm^2 after 4 µs, and 79 mJ/cm2 after 4 µs. Each dataset with laser triggering was probed four times, collecting data from one pulse timed with the pump laser, called "on 1", and three pulses without triggering named "off 1", "off 2" and "off 3" (chapter 3.2). In addition, one reference set of pure deoxyMb and one of oxyMb were collected. All datasets were processed as described in chapter 4.9.

5.2.1. Indexing of serial crystallography diffraction data

After the initial indexing, the mix-and-inject data triggered with 194 mJ/cm^2 and 4 µs probing delay could not be further processed, as less than 10,000 frames per structure were collected, resulting in low data quality and incomplete reflections even in low-resolution shells (Table 8, p. 39). The oxyMb and deoxyMb references showed the best data quality with the highest overall $CC_{1/2}$ of 97.56% and 97.8%, and the lowest overall R_{split} of 13.78 and 13.47, respectively, while cut at 1.60 Å (Table 8, p. 37). Compared to that, the references without laser triggering had to be cut at 1.7 Å to maintain a $CC_{1/2}$ above 30% in the highest-resolution shell of 31.78 and 45.09 for the shared-reservoir and mix-and-inject approaches, respectively (Table 8, pp. 37-38). The trigger data sets from the shared-reservoir approach and the mix-and-inject deoxyMb + HPBC, triggered with 79 mJ/cm^2 and probed after 4 µs, were both cut at 1.7 Å (Table 8, pp. 37, 39). However, the mix-and-inject data showed a higher correlation, ranging from 36.15% to 47.67% in the highest-resolution shell, than the shared-reservoir data, which ranged from 25.29% to 35.87%, and a lower R_{split}, from 56.58 to 63.16 in the highest-resolution shell, compared with 78.72 to 81.49, respectively. In addition, the data from the lower laser fluence of 79 mJ/cm^2 and a later probing time of 4 µs showed higher data quality than the data from the higher laser fluence of 194 mJ/cm^2 and a 10 ns probing delay (Table 8, p. 38) and could therefore be processed at a higher resolution.

Table 8: Data collection and indexing statistics. All data sets were indexed at the detector's cut-off at 1.5 Å and subsequently reprocessed at a resolution cut-off giving an SNR above 1 and $CC_{1/2}$ above 0.3 in the highest resolution shell. Given are the overall statistics and the statistics for the highest-resolution shell in parentheses.

Data set	oxyMb	deoxyMb	shared reservoir				
			deoxyMb + HPBC, no laser	deoxyMb + HPBC, 300 mJ/cm², 100 ns			
				on 1	off 1	off 2	off 3
Instrument	SPB/SFX instrument at European XFEL		SPB/SFX instrument at European XFEL	SPB/SFX instrument at European XFEL			
Wavelength (Å)	1.31		1.03	1.03			
Pulse energy (mJ)	1.7-2.2	2.2-2.5	1.3	1.9			
Pulse length (fs)	25		25	25			
Detector	AGIPD 1M		AGIPD 1M	AGIPD 1M			
Temperature (K)	293		293	293			
Space group	$P2_1$		$P2_1$	$P2_1$			
a, b, c (Å)	35.70, 28.79, 63.41		36.05, 29.11, 64.03	36.02, 29.11, 64.04			
α, β, γ (°)	90.0, 105.46, 90.0		90.0, 105.48, 90.0	90.0, 105.48, 90.0			
Resolution (Å)	20.0-1.60	20.0-1.60	20.0-1.70	20.0-1.70			
No. of indexed frames	212495	190089	12124	9953	10138	9796	10012
Completeness (%)	99.94 (99.41)	100.0 (100.0)	100 (100)	100 (99.96)	100 (100)	100 (100)	100 (100)
***CC** (%)**	99.38 (73.40)	99.44 (74.69)	90.21 (69.45)	91.09 (71.43)	91.86 (72.67)	90.34 (71.20)	91.61 (63.53)
$CC_{1/2}$ (%)	97.56 (36.88)	97.80 (38.68)	68.60 (31.78)	72.98 (35.87)	68.93 (33.95)	72.29 (25.29)	70.90 (34.24)
Redundancy	376.7 (65.8)	372.9 (65.0)	23.4 (14.5)	21.0 (13.0)	21.0 (13.1)	20.1 (12.5)	20.9(13.1)
R_{split}	13.78 (70.01)	13.47 (90.59)	47.88 (76.38)	45.15 (79.29)	47.0 (79.44)	45.27 (81.49)	46.56 (78.72)
$I/\sigma(I)$	5.75 (1.38)	5.41 (1.02)	2.63 (1.37)	2.46 (1.08)	2.44 (1.13)	2.48 (1.09)	2.47 (1.13)
Wilson B-factor (Å²)	20.29	22.90	13.77	15.21	15.57	15.05	15.06

Data set	mix-and-inject				
	deoxyMb + HPBC, no laser	**deoxyMb + HPBC, 194 mJ/cm^2, 10 ns**			
		on 1	**off 1**	**off 2**	**off 3**
Instrument	SPB/SFX instrument at European XFEL	SPB/SFX instrument at European XFEL			
Wavelength (Å)	1.31	1.31			
Pulse energy (mJ)	2.2-2.5	2.2-2.5			
Pulse length (fs)	25	25			
Detector	AGIPD 1M	AGIPD 1M			
Temperature (K)	293	293			
Space group	$P2_1$	$P2_1$			
a, b, c (Å)	35.70, 28.73, 63.48	35.69, 28.75, 63.53			
α, β, γ (°)	90.0, 105.42, 90.0	90.0, 105.38, 90.0			
Resolution (Å)	20.0-1.70	20.0-2.00			
No. of indexed frames	44763	23498	32085	39627	40055
Completeness (%)	99.97 (99.70)	100 (100)	99.96 (99. 57)	99.93 (99.27)	99.97 (99.63)
***CC** (%)**	97.97 (78.84)	97.45 (84.18)	97.05 (86.85)	97.43 (91.04)	97.67 (89.53)
***CC*$_{1/2}$ (%)**	92.26 (45.09)	90.42 (54.87)	89.0 (60.55)	90.36 (70.77)	91.18 (66.89)
Redundancy	83.9 (39.5)	55.0 (32.9)	74.0 (43.1)	92.2 (53.5)	92.6 (54.0)
***R*$_{split}$**	24.12 (62.71)	26.46 (54.36)	25.48 (46.22)	23.71 (41.54)	22.69 (42.68)
***I*/ σ (*I*)**	3.58 (1.64)	3.55 (2.01)	3.96 (2.46)	4.26 (2.54)	4.36 (2.67)
Wilson B-factor (Å²)	22.01	26.36	20.36	19.75	19.18

Data set	mix-and-inject							
	deoxyMb + HPBC, 194 mJ/cm^2, 4 µs				deoxyMb + HPBC, 79 mJ/cm^2, 4 µs			
	on 1	off 1	off 2	off 3	on 1	off 1	off 2	off 3
Instrument	SPB/SFX instrument at European XFEL				SPB/SFX instrument at European XFEL			
Wavelength (Å)	1.31				1.31			
Pulse energy (mJ)	2.2-2.5				2.2-2.5			
Pulse length (fs)	25				25			
Detector	AGIPD 1M				AGIPD 1M			
Temperature (K)	293				293			
Space group	$P2_1$				$P2_1$			
a, b, c (Å)	35.71, 28.75, 63.53				35.71, 28.78, 63.54			
α, β, γ (°)	90.0, 105.41, 90.0				90.0, 105.38, 90.0			
Resolution (Å)	20.0-1.50				20.0-1.70			
No. of indexed frames	1491	748	8950	1786	47020	47562	48276	43258
Completeness (%)	63.08 (1.34)	43.9 (0.33)	93.41 (26.69)	59.28 (1.03)	100 (99.96)	100 (99.92)	100 (99.96)	100 (100)
CC* (%)	71.82 (nan)	59.39 (nan)	88.14 (58.79)	65.42 (nan)	97.87 (72.87)	97.85 (80.35)	98.17 (78.89)	98.05 (75.54)
***CC*$_{1/2}$ (%)**	34.75 (nan)	21.41 (nan)	63.52 (20.89)	27.22 (nan)	92.55 (39.92)	91.93 (36.15)	91.86 (47.67)	92.99 (45.18)
Redundancy	1.5 (2.1)	0.8 (2.1)	11.1 (2.6)	1.3 (2.1)	89.3 (43.1)	90.6 (43.6)	92.0 (44.2)	80.9 (39.1)
***R*$_{split}$**	69.81 (nan)	80.16 (nan)	51.62 (-273.7)	75.56 (nan)	23.12 (56.58)	63.16 (59.48)	24.31 (59.48)	22.62 (60.7)
***I*/ σ (*I*)**	3.927 (-0.63)	3.976 (-2.32)	1.992 (-0.7)	5.475 (-6.98)	3.648 (1.8)	3.613 (1.65)	3.761 (1.86)	3.761 (1.86)
Wilson B-factor (Å^2)	-	-	-	-	22.49	22.64	22.44	23.87

5.2.2. Phasing and Refinement

In a first instance, all structures were refined without adding a ligand or water to the heme pocket (chapter 4.10). During refinement, no structure showed distinct electron density for the last two C-terminal amino acids (glutamine 152, glycine 153). Therefore, their occupancies were set to zero in all structures. The refinement statistics of all final structures are given in Table 9.

After the initial refinement, the deoxyMb structure showed a positive density in the heme pocket, which could be fitted with a single water molecule at 100% occupancy (Figure 21 A, B). The data taken as a reference for oxyMb showed a density above the heme's iron (Figure 21 C). Occupancy refinement of an oxygen modelled into the density set the occupancy to 59% at R_{free} = 20.02%. However, adding the oxygen resulted in negative electron densities (Figure 21 D). In addition, the distance between the iron and the proximal histidine of 2.14 Å corresponded to the expected distance for *aquomet*-Mb, rather than 2.06 Å for oxyMb (Table 10) [44]. H_2O modelled at 100% occupancy into the density resulted in a lower R_{free} of 19.87% with no negative densities (Figure 21 E). Therefore, the structure was not further used as a reference for oxyMb.

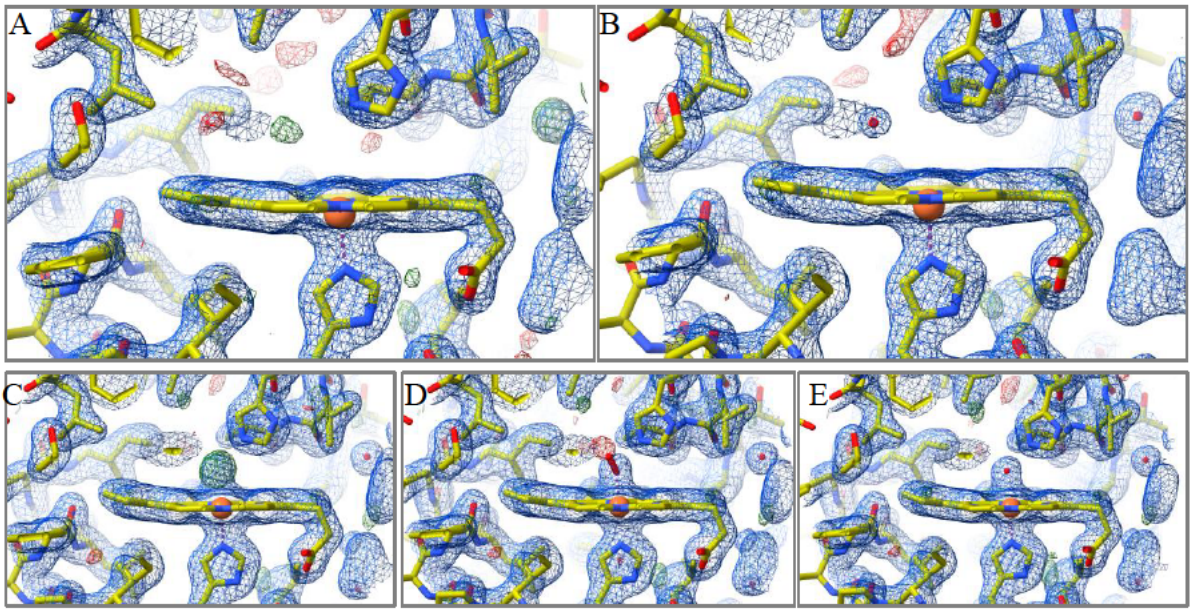


Figure 21: Electron density maps of the reference structures. Shown are the electron density maps of deoxyMb and oxyMb after the initial refinement and after a ligand or solvent molecule were added to the heme pocket. Blue: 2mFo-DFc map at an r.m.s.d of 1.50, green: positive mFo-DFc densities at +3σ, red: negative mFo-DFc densities at -3σ. A: deoxyMb after the initial refinement, B: deoxyMb final refinement with one H_2O molecule in the heme pocket, C: oxyMb reference structure after the initial refinement, D: oxyMb reference structure with a bound O_2 molecule occupancy refined to 59%, E: final refinement with a bound H_2O.

Table 9: Refinement statistics. Listed are the refinement statistics of all final myoglobin structures. * Others atoms include the heme structure, the oxygen ligand and additional solvent molecules.

Data set	oxyMb	deoxyMb	shared reservoir				
			deoxyMb + HPBC, no laser	deoxyMb + HPBC, 300 mJ/cm², 100 ns			
				on 1	off 1	off 2	off 3
Resolution Range (Å)	20.0-1.60	20.0-1.60	20.0-1.70	20.0-1.70			
No. of reflections (working set)	17510	17494	13176	13199	13212	13193	13561
No. of reflections (test set)	917	916	694	690	698	700	716
R_{work} (%)	18.40	19.02	23.07	23.20	22.72	22.72	22.53
R_{free} (%)	19.87	21.80	26.08	26.20	27.77	26.41	26.11
No. of atoms							
Protein	1242	1242	1242	1242	1242	1242	1242
Water	65	64	94	51	82	91	69
Others*	46	50	49	49	51	49	51
Protein residues	153	153	153	153	153	153	153
R.m.s. deviations							
Bond length (Å)	0.007	0.007	0.007	0.014	0.008	0.008	0.012
Bond angles (°)	0.886	1.040	0.835	1.183	0.982	0.999	1.003
B-factors (Å²)							
Avg. B-factors	23.99	24.03	16.89	17.93	18.40	18.03	19.06
Protein	23.71	27.32	16.65	17.94	18.20	17.51	18.90
Water	31.57	32.91	21.55	20.21	22.79	26.15	22.94
Ramachandran plot (%)							
Favoured	98.01	99.34	99.34	99.32	98.68	98.65	98.68
Allowed	1.99	0.66	0.66	0.68	1.32	1.35	1.32
Outliers	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Rotamer outlier (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Clashscore	2.81	2.40	6.08	3.25	6.08	6.90	1.62

Data set	mix-and-inject								
	deoxyMb + HPBC, no laser	deoxyMb + HPBC, 194 mJ/cm², 10 ns				deoxyMb + HPBC, 79 mJ/cm², 4 µs			
		on 1	off 1	off 2	off 3	on 1	off 1	off 2	off 3
Resolution Range (Å)	20-1.70	20.0-2.00				20.0-1.70			
No. of reflections (working set)	13257	8339	8251	8233	8219	13328	13229	13296	13361
No. of reflections (test set)	710	435	430	434	435	699	695	705	705
R_{work} (%)	19.87	18.32	17.79	17.64	17.93	19.54	20.29	19.68	19.08
R_{free} (%)	22.89	25.86	24.39	22.78	22.06	21.80	21.62	22.22	23.18
No. of atoms									
Protein	1242	1242	1242	1242	1242	1242	1242	1242	1242
Water	89	59	75	71	72	66	62	63	69
Others*	49	50	49	50	51	50	44	44	49
Protein residues	153	153	153	153	153	153	153	153	153
R.m.s. deviations									
Bond length (Å)	0.008	0.007	0.009	0.006	0.008	0.006	0.006	0.006	0.006
Bond angles (°)	0.958	0.922	0.918	0.932	0.970	0.980	0.913	0.978	1.003
B-factors (Å²)									
Avg. B-factors	23.13	30.51	24.27	21.22	21.55	24.00	23.67	23.93	25.84
Protein	22.61	30.28	23.85	20.94	21.07	23.60	23.50	23.79	25.53
Water	23.38	37.09	32.89	28.16	30.32	32.08	30.62	30.63	33.66
Ramachandran plot (%)									
Favoured	98.01	93.50	91.06	98.01	99.34	99.34	98.01	98.01	98.01
Allowed	1.99	6.50	8.94	1.99	0.66	0.66	1.99	1.99	1.99
Outliers	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Rotamer outlier (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Clashscore	6.42	8.11	4.87	1.61	2.84	3.20	2.41	2.03	2.84

All structures from a shared reservoir of deoxyMb and the HPBC oxygen cage showed a distinct density near the heme iron, where bound oxygen could be modelled (Figure 22, Table 13). The refinement of the occupancy of those structures resulted in 60% occupancy in the "on 1" structure to 92% in the reference structure without laser triggering (Table 10).

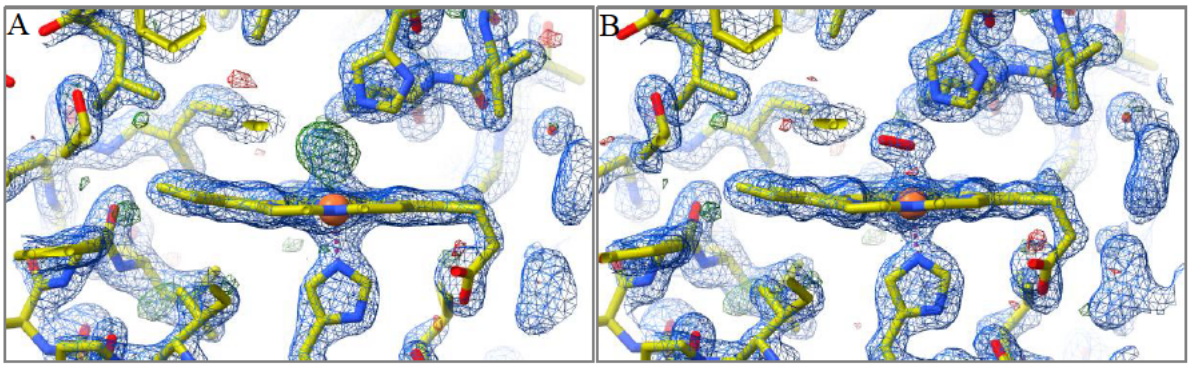


Figure 22: Electron density from the shared-reservoir approach. Shown exemplarily is the electron density of the dark deoxyMb + HPBC data without laser triggering from the shared-reservoir approach. A: electron density before a ligand was added to the heme, B: final electron density with an O_2 at 92% occupancy. Blue: 2mFo-DFc map at an r.m.s.d of 1.50, green: positive mFo-DFc densities at +3σ, red: negative mFo-DFc densities at -3σ.

The densities in the mix-and-inject approach were less unambiguous than those in the shared-reservoir approach and stretched from the deoxy water position towards the centre of the heme (Figure 23 A.1, B.1, Table 13). For a clearer view on the ligand density, isomorphous difference maps were calculated (chapter 4.10.1 I) by subtracting the deoxyMb density from the triggered data (Figure 23 A.2, B.2). The isomorphous difference maps showed no positive or negative density around the deoxyMb water, indicating that all structures of the mixing-and-inject approach had a similar electron density to the deoxyMb structure at that position. Meanwhile, positive densities of varying intensities were visible above the heme, indicating that all structures had an additional density at the bound position. In addition, some structures showed a positive density above the histidine, indicating movement towards the heme (Figure 23 A 2).

With alternating conformation refinement of the deoxy water and bound oxygen (chapter 4.10.1 II), the oxygen occupancies varied from the lowest 15% and 18% in the deoxyMb + HPBC, 10 ns, 194 mJ/cm^2 "off 2" and deoxyMb + HPBC, 4 µs, 79 mJ/cm^2 "off 2", respectively, to 39% in the data without laser triggering (Table 10). However, the refined occupancies did not correlate with the displayed electron densities in all structures. In the deoxyMb + HPBC, 4 µs, 79 mJ/cm^2 "off 2" data with 18% oxygen occupancy, the electron density was larger than in the "on 1" at 25% of the same data set (Table 13). To get a clearer

view of the oxygen occupancies, the geometry of the heme and the overall movement of the proximal and distal helix E & F were measured for all structures (chapter 4.10.1 III, IV).

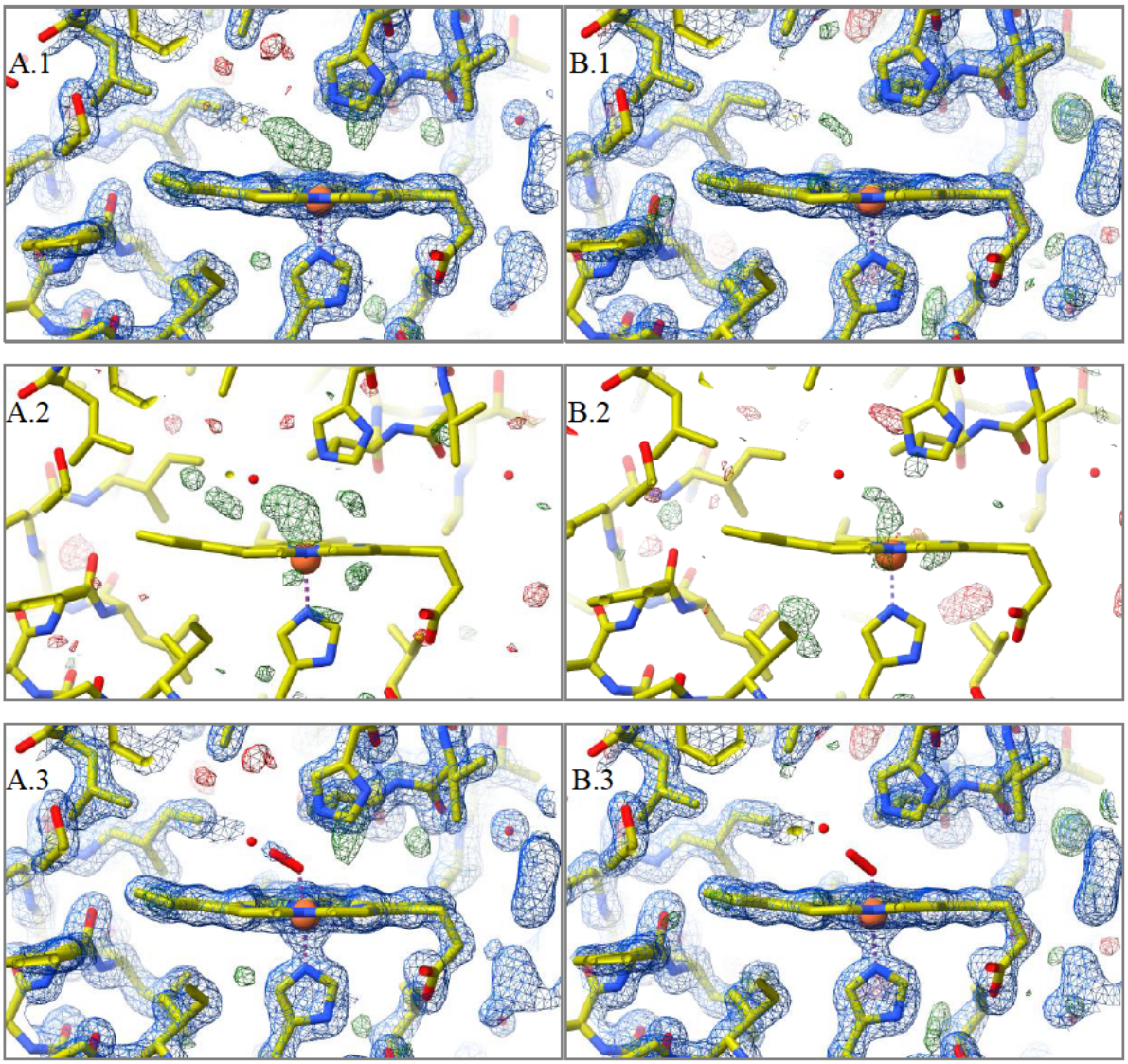


Figure 23: Electron density from the mix-and-inject approach. Shown exemplarily are 1: the electron densities before a ligand and solvent molecules were added to the heme pocket, 2: isomorphous difference maps calculated by subtracting the refined deoxyMb structure factors, 3: the electron densities of the final structures. A: the dark deoxyMb + HPBC data without laser triggering from the mix-and-inject approach, B: deoxyMb + HPBC, 79 mJ/cm^2, 4µs, on 1. Blue: 2mFo-DFc map at an r.m.s.d of 1.50, green: positive mFo-DFc densities at +3σ, red: negative mFo-DFc densities at -3σ.

Table 10: Ligand occupancy and protein movement statistics. Listed are oxygen occupancies calculated by the *phenix.refine* program, the distance between the heme's iron and the proximal histidine, the distance measured between the heme's iron and the heme's N-plane as well as the mean movement of helix E and F helices compared to their position in the deoxyMb structure.

Structure		Refined oxygen occupancy [%]	Fe-to-His distance [Å]	Fe-to-heme N-plane distance [Å]	Mean helix movement [Å]	
					E	F
deoxyMb		-	2.13	-0.29	-	-
oxy-/*aquomet*-Mb		-	2.14	-0.07	-	-
shared reservoir						
deoxyMb + HPBC, no laser		92	2.11	-0.05	0.17	-0.2
deoxyMb + HPBC, 300 mJ/cm^2, 100 ns	on 1	60	2.07	-0.08	0.13	-0.18
	off 1	64	2.08	-0.08	0.15	-0.18
	off 2	77	2.05	-0.11	0.15	-0.17
	off 3	75	2.05	-0.11	0.14	-0.18
mix-and-inject						
deoxyMb + HPBC, no laser		39	2.09	-0.22	0.03	-0.07
deoxyMb + HPBC, 194 mJ/cm^2, 10 ns	on 1	30	2.07	-0.17	0.02	-0.08
	off 1	33	2.06	-0.20	0.05	-0.1
	off 2	15	2.07	-0.19	0.06	-0.08
	off 3	35	2.06	-0.15	0.04	-0.11
deoxyMb + HPBC, 79 mJ/cm^2, 4 µs	on 1	25	2.12	-0.24	0.02	-0.04
	off 1	30	2.14	-0.23	0.01	-0.03
	off 2	18	2.13	-0.25	0.01	-0.02
	off 3	30	2.07	-0.26	0.00	-0.04

Overall, the measured distances did not correlate with the calculated occupancies and were inconsistent within a single structure. In the triggered data from the shared-reservoir approach, the distance between the iron and the His decreased with increasing occupancy, whereas the Fe-to-N-plane distance increased as well at higher occupancies. Meanwhile, the reference data with 92% occupancy showed a decreased Fe-to-N-plane distance while the Fe-His distance increased compared to the triggered data. In the mix-and-inject approach, the reference data at 39% occupancy showed longer distances for both the Fe-to-N-plane and Fe-His distances than the data triggered with 194 mJ/cm^2, though those were calculated to lower occupancies. In the data triggered with 79 mJ/cm^2, the oxygen occupancies of "off 1" and "off 3" were both

calculated to 30%. Yet in both structures, longer Fe-to-N-plane distances were measured than in the “on 1” at 194 mJ/cm^2, which was also set to an oxygen occupancy of 30% (Table 10).

In the difference distance calculations, the E & F Helix showed stronger movement in the shared-reservoir approach than in the mix-and-inject data, and stronger movements in the 194 mJ/cm^2 data than in the 79 mJ/cm^2 data (Figure 24). In the untriggered data of the shared reservoir, the E helix moved away from the heme with a mean helix movement of 0.17 Å, whereas the F helix moved closer to the heme with a mean movement of -0.2 Å. In comparison, the E helix of the reference data of the mix-and-inject approach moved with a mean value of 0.03 Å and the F helix with a mean of -0.07 (Figure 24 A). In the triggered data from the shared-reservoir, distinct movements of the two helices are visible (Figure 24 B), with the E helices moving around 0.15 Å away and the F helices around 0.18 Å towards the heme. In the mix-and-inject data triggered with 194 mJ/cm^2, the F helices showed movement of about 0.1 Å towards the heme. The first residues of the E helices (serine 58-threonine 70), which are more closely located to the heme, moved away from the heme while the residues alanine 71 to lysine 77 moved closer, resulting in a mean movement of 0.06-0.02 Å. Meanwhile, in the 79 mJ/cm^2 triggered data, all movements lay below 0.05 Å, and graphically no distinct movements were visible.

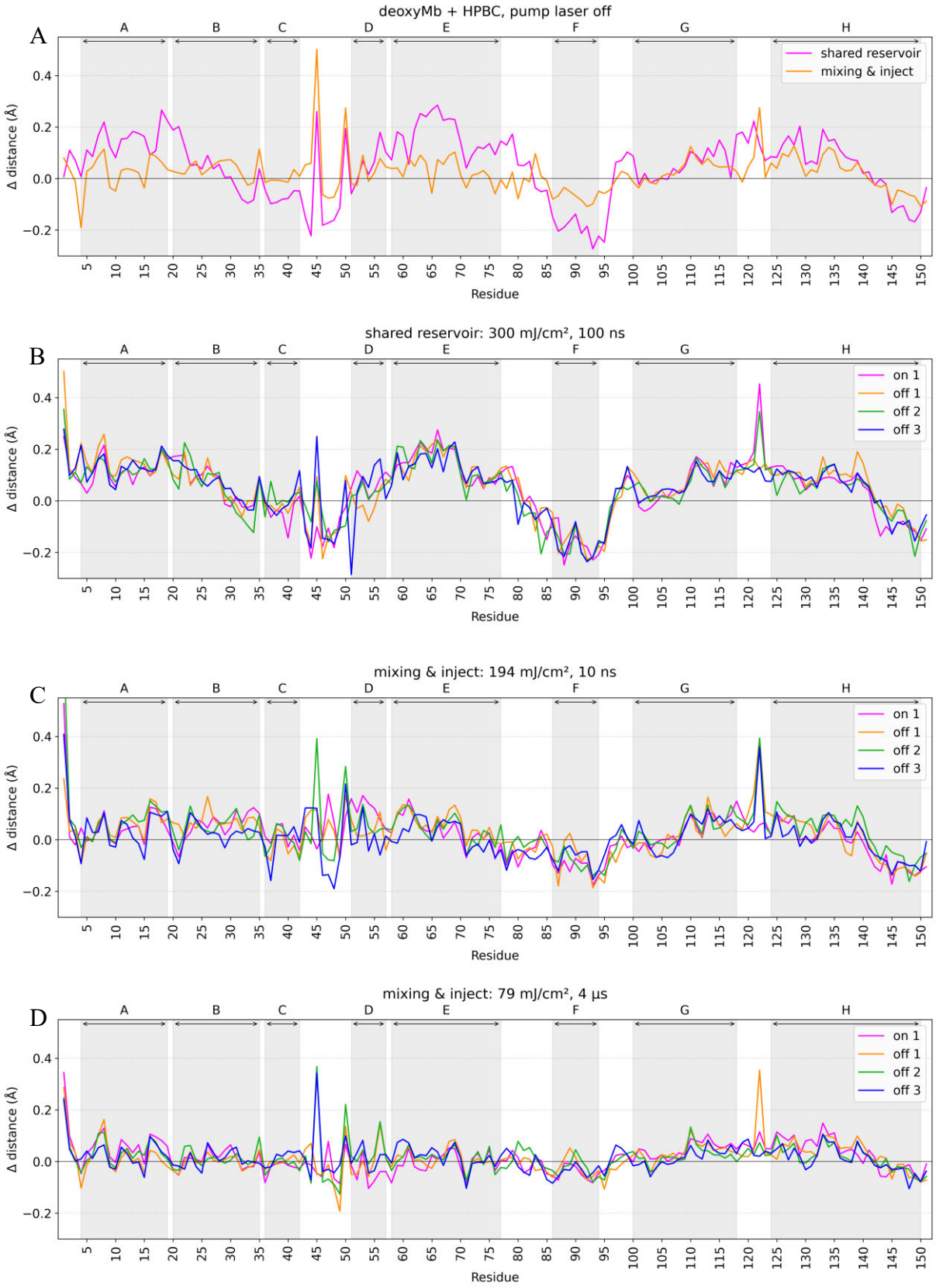


Figure 24: Distance difference graphs. Displayed are the distance differences between the refined deoxyMb + HBPC structures and deoxyMb for each amino acid. A: Distance differences of the shared-reservoir (pink) and mix-and-inject (orange) dark references. B: Distance differences of the shared-reservoir data triggered with 300 mJ/cm² and probed after 100 ns. C: Distance differences of the mix-and-inject data triggered with 194 mJ/cm² and probed after 10 ns. D: Distance differences of the mix-and-inject data triggered with 79 mJ/cm² and probed after 4 µs. The triggered data is displayed in pink: on 1, orange: off 1, green: off 2, blue: off3.

6. Discussion

6.1. DNA reporter system

The dimerisation of adjacent thymine bases deforms the DNA's linear structure and creates a kink. Due to this kink, crystals of DNA oligos with the same sequence, with and without dimerised thymine, exhibit different unit cell parameters [31]. Thymine dimers are induced by UV irradiation and form within picoseconds at sufficient irradiation dose, making the dimerisation of thymine in DNA crystals an excellent reporter system to test the instrumental set-up for pump-probed SFX experiments, without the need for structural phasing and refinement [29,45].

In this study, preliminary tests, including screening of crystallisation conditions, analysis of thymine dimerisation in liquid samples, and diffraction data collection at a synchrotron source, were performed. The tests aimed to gather insights into the system's applicability that are needed to establish a robust system before transferring it to an XFEL source, as experimental time at XFELs is highly limited and therefore valuable.

The analysis of the thymine dimerisation was performed using liquid 10-mer samples, as the number of suitable DNA crystals was too limited to establish workflows that yielded comparable results, and 20-mers showed poor performance in the crystallisation screening (chapter 5.1.2).

In the UV-absorption spectra, the samples irradiated at 360 mJ/cm² and higher doses showed the expected dose-dependent decrease at $A_{257.5}$ (chapter 5.1.2), which occurs due to the transformation of the C_5=C_6 double bond to single bonds in thymine photoproducts [30]. While the changes at 36 and 360 mJ/cm² remained within the variability of the unirradiated control, larger reductions were observed at 720 mJ/cm² and 1080 mJ/cm² (Figure 17). However, the 0.58% ± 0.74% decrease observed at 720 mJ/cm² cannot be clearly attributed to the radiation dose, given the high standard deviation. In contrast, although the standard deviation of the 1080 mJ/cm² samples with ± 1.38% was large compared to the mean change of -1.91%, the lower limit at 0.53% still exceeded the variability of the unirradiated samples. Overall, the data suggest a change in the UV absorption with irradiation dose, but a clear effect was observed only at the highest dose.

In Raman spectroscopy of DNA (chapter 5.1.2), the region from 1200 cm^{-1} to 1700 cm^{-1} is assigned to the stretching vibrations of the bases. In particular, the peak around 1250 cm^{-1} is

attributed to the ring stretching of purine bases, the peak at 1375 cm^{-1} to the methyl bending, and the peak around 1670 cm^{-1} to the carbonyl double bond. Each of those peaks is affected by the formation of T-T CPDs due to alternating electron distributions in the molecule, yet the contribution of thymine varies. The 1375 cm^{-1} peak additionally arises from adenine and guanine vibrations with only a weak contribution of the thymine methyl group [30,46,47]. Therefore, changes due to dimerisation can be small and difficult to detect.

The results from Raman spectroscopy showed the same behaviour as the UV-absorption (chapter 5.1.2). While the intensity of the 36 mJ/cm² irradiated samples increased, the intensity of the samples irradiated with higher doses decreased (Figure 18). In the collected spectra, changes in intensities in the thymine-associated peaks were strongly visible at 1250 cm^{-1} and 1670 cm^{-1}, whereas changes at 1375 cm^{-1} were only visible in the highest irradiation dose. The greatest changes in those peaks were observed at an irradiation dose of 360 mJ/cm². The increase in intensities with higher irradiation doses is consistent with the results from Nagpal *et al.*, who analysed changes in UV-absorption and Raman spectra in thymine solution and calf thymus DNA upon UV irradiation [30]. Nagpal *et al.* reported that after a dose of approximately 400 mJ/cm², the reversible formation of CPDs reaches equilibrium, while the ratio of the T-T (6-4) photoproducts increased [29,30]. In parallel, their recorded Raman intensities decreased up to an irradiation dose of approximately 400 mJ/cm² and increased at higher doses, while the UV absorption continued to decrease. The similar trends in the UV-absorption and Raman spectra between the data collected in this study and those of Nagpal *et al.* indicate similar behaviour of the induced photodamage. More detailed differentiation of those photoproducts would require additional methods like high-performance liquid chromatography, liquid chromatography-mass spectrometry or the elucidation of the structures by X-ray crystallography [29]. However, in a final application of the DNA as a reporter system, the kind of induced photoproducts would be of lower importance, as all mentioned thymine photoproducts kink the DNA, which should all be visible in the unit cell parameters, as (6-4) photoproducts bring higher structural perturbation [29,48].

After the initial crystallisation condition screening, new crystals were grown in the 10.1 buffer at 35% (w/v) PEG and 2.0 mM DNA for diffraction data collection at the EMBL P13 beamline at DESY (chapter 5.1.3). Unirradiated and irradiated crystals of that condition, as well as residual crystals from the condition screening, were tested for diffraction. The tested crystals rarely produced distinct Bragg peaks (Figure 19), and usable data could not be collected.

Instead, the features visible in the single diffraction patterns were similar to features of DNA fibre diffraction. The X-shaped shadows are a typical feature resulting from the helical turns of double-stranded DNA, whereas circular-arranged spots at low resolution arise from the 10 base pairs of one helical turn [49,50]. The rings around 4 Å are less distinct and are not a characteristic feature of DNA fibres. Since they occur at the same resolution as the X-features, they may be smears of those features arising from less-oriented molecules. The less distinct features of DNA fibre diffraction arise from the higher molecular disorder in those fibres compared to the order in crystals. In crystals, each unit cell is repeated by translation while maintaining a fixed orientation, forming a three-dimensional periodic lattice [51]. In contrast, DNA fibres maintain periodic translational order along the fibre axis, but the molecules' orientations along the axis vary. This disorder broadens the diffraction into features such as layer lines, arcs, or rings at higher disorder, instead of discrete Bragg peaks [49,52].

The features in the collected diffraction images suggest that the internal structure of the obtained crystals showed low order and not a true crystal lattice. Instead, the features can be explained by a fibre-like arrangement of the molecules or high mosaicity in the crystals. Additionally, the crystals examined at the synchrotron measured between 10 and 50 µm, while the P13 beamline suggests a minimum size of 20 µm [40]. The small size may have contributed to the low diffraction, together with the disordered structure, as fewer molecules are present that could contribute to crystal diffraction [15].

To overcome both limitations, further optimisation of the crystallisation conditions is needed. As shown in chapter 5.1.1, lower PEG and DNA concentrations resulted in visually more ordered crystals. Therefore, further reduction may also increase the internal order. Based on the phase diagram, a good crystallisation condition would enter the nucleation zone close to the supersaturation line, so that after a few crystallisation nuclei are formed, the metastable zone is reached. Then crystals can grow from those nucleation points without new nuclei interfering, leading to a more ordered structure and larger crystals. Adjusting the pH and growth temperature, in addition to changing the DNA and PEG concentrations, could also positively affect the internal order by altering the phase diagram. Alternatively, seeding can be used to avoid entering the nucleation zone. By placing microcrystals in a metastable solution, crystals can grow at mild conditions, avoiding high supersaturation levels that could lead to irregularly shaped crystals grown from multiple nucleation points [14,15,31,51].

6.2. HPBC oxygen cage

The experiments performed at SPB/SFX aimed to characterise the HPBC oxygen cage for future implementation as a light-sensitive component in pump-probe TR-SFX experiments. In the first approach, HPBC and deoxyMb were injected together from a shared reservoir, but the dark states showed high oxygen occupancies in the myoglobin and therefore severe leakage of the cage (chapter 5.2.2, Table 10). To counteract the leakage, deoxyMb and HPBC were supplied from two separate reservoirs in a new approach, reducing the contact time between the myoglobin and leaked oxygen to 15 ms of mixing time (chapter 5.2).

With the reduced contact time in the mix-and-inject approach, it was possible to reduce the oxygen occupancies in the "off" states from 60-77% to 15-35%. However, in the mix-and-inject approach, the calculated occupancies did not overall correlate with the displayed electron densities (Table 13) or geometry measurements in the heme pocket and the helix movements (chapter 5.2.2, Table 10, e.g. 194 mJ/cm² "on 1" compared to 194 mJ/cm² "off 1"). Although in the shared-reservoir approach the distinct electron densities were visible that could be accommodated by single O_2 molecules without additional water in the heme pocket, the calculated occupancies also did not correlate with the measured heme geometries and helix movements (e.g. "off 1" compared to "off 2") (Table 10).

The occupancy refinement of *phenix* is based on the refinement of model-derived structure factors to fit experimental structure factors [16]. Low occupancies and overlapping densities from the deoxy water, bound oxygen and *aquomet*-H_2O complicate the differentiation and may lead to the wrong estimation of a single state with structure factors calculated from multiple states, as the *phenix* occupancy refinement is not specified for that task [53]. The lack of differentiation was visible in the refinement of the supposed oxyMb reference structure. When a bimolecular oxygen was used for the refinement, the oxygen atom closer to the heme could be modelled into the free density without issues, visible in the surrounding blue electron density, whereas the second oxygen atom was surrounded by negative, red density (Figure 21 D). In that case, the problem was obvious due to the negative density. Data from *aquomet*-Mb, deoxyMb and oxyMb can show larger densities arising from multiple overlapping densities that can lead to wrongly estimated occupancies. Therefore, the calculated occupancies from the mix-and-inject approach should be treated with greater caution. Nevertheless, the measured distances in the heme pocket and the movements of the E and F helices were consistent in indicating that the oxygen occupancy in the mix-and-inject data was reduced compared to the shared-reservoir approach (chapter 5.2.2).

Besides the mix-and-inject data triggered with 79 mJ/cm² and probed after 4 µs, all structures with HPBC, triggered with 194 mJ/cm² and probed after 10 ns and triggered with 300 mJ/cm² and probed after 100 ns (chapter 5.2), including the references without any laser pulse, showed clear signs of oxygen as densities in the heme pocket, in their measured heme geometries and helix movements. In addition, the isomorphous difference maps from all structures showed residual densities in the heme's vicinity (chapter 5.2.2, Figure 23, Table 13). Although all structures with HPBC were refined with an alternating conformation of the deoxy- and oxyMb state, the ligands in the 79 mJ/cm² triggered data set cannot be determined with confidence, as the differences in the measured movements remained close to the deoxyMb state and densities close to the heme could also arise from *aquomet*-Mb. However, the distance changes in structures of the mix-and-inject data triggered with 194 mJ/cm² and the reference without laser triggering are proof that premature oxygenation still occurred. The deoxyMb reference structure showed no sign of oxygen in the electron density (chapter 5.2.2, Figure 21 A, B). Accordingly, the oxygenation of the tested structures can be ascribed solely to leakage of the HPBC compound, as the deoxyMb structure shows that external oxygen could be excluded from the experimental setup. Therefore, the contact time between the myoglobin and leaked oxygen could not be sufficiently reduced to enable the use of the HPBC oxygen cage as the intended light-sensitive component in pump-probed TR-SFX.

Photocages are meant to release their caged molecule with high temporal accuracy to enable the analysis of short-lived, transient reaction intermediates. To gather data from those intermediates, the initiation of the reaction under investigation needs to be timed homogeneously, as intermediates from varying time points would blur the data. Hence, good stability, a high quantum yield and a fast decaging rate are desired. However, good stability often comes with compromised release and vice versa [5,32]. In this study, the stability over time in the experimental environment was the major limitation, which has been reported before. In a study from Sandelin *et al.* the observed instability that occurred within a few seconds could be traced back to reducing conditions in their buffer. Yet, additional temperature dependent degradation of the cage was observed over the course of several hours at room temperature in aqueous solution [32]. The HPBC used for the data analysed in this study was stored as a powder at 4°C and brought into solution shortly before the acquisition of the data. The observed leakage therefore might be due to the buffer condition, which could be improved for better stability.

Alternatively, HPBCs with additional side groups have been reported with higher resistance against degradation but also reduced quantum yield [32]. Lower quantum yields require higher laser fluences for a suitable decaging of the reactant, but high fluences can also lead to multiphoton effects in other reaction components if they absorb at a similar wavelength [5,11,33]. Multiphoton absorption in myoglobin and other proteins can induce non-physiological excited states that potentially alter the reaction under investigation and lower the data quality due to inhomogeneous excitation [39,54]. In this study, the highest laser fluence used was 300 mJ/cm². Comparing the indexing statistics of those structures with the data triggered at 79 mJ/cm² showed higher overall R_{split} and lower overall $CC_{1/2}$ values, demonstrating lower data quality (Table 8) [28,55]. While the lower quality could arise from the higher laser fluence, the data sets were also collected at different time points. As the lower laser fluence data was collected after 4 µs and the higher fluence after 100 ns, a longer relaxation time after the triggering could also positively influence the data quality. A clear effect of the pump laser was visible in the oxygen occupancies of the "on" and "off" data. The calculated occupancy, iron-to-N-plane distance and the movement of the E and F helices (Table 10) indicate that the reference data without laser triggering showed the highest oxygen occupancy. Meanwhile, based on occupancy refinement, the Fe-to-His distance and helix movement, the "on 1" data showed the lowest occupancy, which increased again in the "off" structures, indicating photolysis of the oxygen upon the 300 mJ/cm² laser pulse. Although comparison to lower laser fluences used in this study is not possible, as in the mix-and-inject data, the analysed parameters were more variable within a single structure and did not show a trend, therefore, the photodissociation in the shared-reservoir emphasises the undesired side effects of high fluences.

Although not suitable as a controlled starting point, the HPBC could instead be used for experiments that temporarily require higher oxygen concentrations than achievable in liquids saturated with air or pure oxygen. Under optimised conditions, Sandelin *et al.* temporarily increased the oxygen concentration to 13 times higher than possible in air-saturated and three times higher than in oxygen-saturated aqueous solution [32]. However, those concentrations were achieved with a 5 ns pulse at 1 mJ/cm. Literature and the results of this experiment suggest that the laser fluence should be kept as low as possible to avoid multiphoton and contradictory reactions [39,54]. Therefore, the use of HPBC for higher oxygen concentrations could be limited.

7. Conclusion and Outlook

The irradiation of the liquid DNA samples showed promising results in the UV and Raman spectra, which indicated a change in the DNA's structure depending on the irradiation dose. The structural change in the liquid sample is a first indication that the DNA can be used as a validation system for the instrumental set-up for pump-probed TR-SFX at the SPB/SFX instrument. However, a major shortcoming was the crystallisation of the DNA into well-ordered crystals, visible in the crystal morphologies as well as the collected synchrotron data. For further investigation, the crystallisation conditions need to be optimised to lower the nucleation points and increase the internal order. Besides adjusting the buffer conditions and DNA concentration, seeding with microcrystals could potentially enhance the order as well as the time needed until the crystals grow to a suitable size. In addition, validation of structural changes within the crystals upon irradiation will also need to be further investigated. The crystallisation and diffraction screening results highlight the importance of careful preparation and preliminary experiments to identify issues early and increase the likelihood of successful data collection at XFELs.

In the analysis of the HPBC diffraction data, the oxygen occupancies could not be determined with confidence, particularly at low levels. However, the dark states clearly showed signs of oxygenation, showing that leakage of the photocage could not be fully compensated by a reduced contact time. Additionally, the reduced occupancies in the respective "on 1" states compared to the "off" states showed how high pump laser fluence not only affects the photocage but also bonds within the protein. One major finding was that no oxygen was present in the deoxyMb structure, proving that at the SPB/SFX instrument, experiments can be performed under the full exclusion of oxygen.

In conclusion, this work showed how preliminary experiments can enhance the success rate for XFEL beamtime by investigating and excluding shortcomings beforehand. Initial tests on the DNA reporter system paved the way for further investigations needed to build a reliable validation tool that increases confidence in collected diffraction data. In addition, the analysis of the HPBC photocage demonstrated the complexity of post-experimental data processing in time-resolved experiments, particularly when structural changes are small. The HPBC did not reliably retain the oxygen and is therefore not suitable for reaction initiation in pump–probe experiments. However, the presented results demonstrated that experiments at the SPB/SFX

instrument at the European XFEL can be performed under anaerobic conditions, thereby establishing a foundation for future studies of oxygen-sensitive systems.

8. References

1. Chrencik JE, Su H-P, Gomez Llorente Y, Palte RL, Klein DJ, Hayes RP, et al. The evolving role of structural biology in pharma: integration of X-ray crystallography, cryo-electron microscopy and beyond. Acta Crystallogr Sect Struct Biol. 2026;82:587–602. https://doi.org/10.1107/S2059798326004390

2. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. Nature. 1958;181:662–6. https://doi.org/10.1038/181662a0

3. De Wijn R, Melo DVM, Koua FHM, Mancuso AP. Potential of Time-Resolved Serial Femtosecond Crystallography Using High Repetition Rate XFEL Sources. Appl Sci. 2022;12:2551. https://doi.org/10.3390/app12052551

4. Pandey S, Bean R, Sato T, Poudyal I, Bielecki J, Cruz Villarreal J, et al. Time-resolved serial femtosecond crystallography at the European XFEL. Nat Methods. 2020;17:73–8. https://doi.org/10.1038/s41592-019-0628-z

5. Monteiro DCF, Amoah E, Rogers C, Pearson AR. Using photocaging for fast time-resolved structural biology studies. Acta Crystallogr Sect Struct Biol. 2021;77:1218–32. https://doi.org/10.1107/S2059798321008809

6. Mancuso AP, Aquila A, Batchelor L, Bean RJ, Bielecki J, Borchers G, et al. The Single Particles, Clusters and Biomolecules and Serial Femtosecond Crystallography instrument of the European XFEL: initial installation. J Synchrotron Radiat. 2019;26:660–76. https://doi.org/10.1107/S1600577519003308

7. Oeser T. Kristallstrukturanalyse durch Röntgenbeugung: Spektroskopiekurs kompakt. Wiesbaden: Springer Spektrum; 2019. https://doi.org/10.1007/978-3-658-25439-1

8. Chapman HN. X-Ray Free-Electron Lasers for the Structure and Dynamics of Macromolecules. Annu Rev Biochem. 2019;88:35–58. https://doi.org/10.1146/annurev-biochem-013118-110744

9. Botha S, Fromme P. Review of serial femtosecond crystallography including the COVID-19 pandemic impact and future outlook. Structure. 2023;31:1306–19. https://doi.org/10.1016/j.str.2023.10.005

10. Calvey GD, Katz AM, Schaffer CB, Pollack L. Mixing injector enables time-resolved crystallography with high hit rate at X-ray free electron lasers. Struct Dyn. 2016;3:054301. https://doi.org/10.1063/1.4961971

11. Schulz EC, Yorke BA, Pearson AR, Mehrabi P. Best practices for time-resolved serial synchrotron crystallography. Acta Crystallogr Sect Struct Biol. 2022;78:14–29. https://doi.org/10.1107/S2059798321011621

12. Darmanin C, Strachan J, Adda CG, Ve T, Kobe B, Abbey B. Protein crystal screening and characterization for serial femtosecond nanocrystallography. Sci Rep. 2016;6:25345. https://doi.org/10.1038/srep25345

13. Friedmann D, Messick T, Marmorstein R. Crystallization of Macromolecules. Curr Protoc Protein Sci [Internet]. 2011 [cited 2026 July 18];66. https://doi.org/10.1002/0471140864.ps1704s66

14. Krauss IR, Merlino A, Vergara A, Sica F. An Overview of Biological Macromolecule Crystallization. Int J Mol Sci. 2013;14:11643–91. https://doi.org/10.3390/ijms140611643

15. McPherson A, Gavira JA. Introduction to protein crystallization. Acta Crystallogr Sect F Struct Biol Commun. 2013;70:2–20. https://doi.org/10.1107/S2053230X13033141

16. Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. Acta Crystallogr Sect Struct Biol. 2019;75:861–77. https://doi.org/10.1107/S2059798319011471

17. Turkot O, Dall'Antonia F, Bean RJ, E J, Fangohr H, Ferreira De Lima DE, et al. EXtra-Xwiz: A Tool to Streamline Serial Femtosecond Crystallography Workflows at European XFEL. Crystals. 2023;13:1533. https://doi.org/10.3390/cryst13111533

18. Yefanov O, Mariani V, Gati C, White TA, Chapman HN, Barty A. Accurate determination of segmented X-ray detector geometry. Opt Express. 2015;23:28459. https://doi.org/10.1364/OE.23.028459

19. Mittemeijer EJ, Welzel U, editors. Modern Diffraction Methods [Internet]. 1st ed. Wiley; 2012 [cited 2026 July 30]. https://doi.org/10.1002/9783527649884

20. Evans P, McCoy A. An introduction to molecular replacement. Acta Crystallogr D Biol Crystallogr. 2008;64:1–10. https://doi.org/10.1107/S0907444907051554

21. Taylor GL. Introduction to phasing. Acta Crystallogr D Biol Crystallogr. 2010;66:325–38. https://doi.org/10.1107/S0907444910006694

22. European XFEL in international comparison [Internet]. [cited 2026 July 25]. https://www.xfel.eu/facility/comparison/index_eng.html. Accessed 25 July 2026

23. Altarelli M, Brinkmann R, Chergui M. The European X-ray Free-Electron Laser. Hamburg: DESY XFEL Project Group European XFEL Project Team Deutsches Elektronen-Synchrotron Member of the Helmholtz Association; 2007 July.

24. How to make great even better – European XFEL receives new electron source "GUN5" · DESY [Internet]. Hamburg: Deutsches Elektronen-Synchrotron DESY; 2025 [cited 2026 July 25]. https://desy.de/desy_latest_news/2025/gun5/index_eng.html. Accessed 25 July 2026

25. Altarelli M. The European X-ray Free-Electron Laser: toward an ultra-bright, high repetition-rate x-ray source. High Power Laser Sci Eng. 2015;3:e18. https://doi.org/10.1017/hpl.2015.17

26. Beamlines [Internet]. [cited 2026 July 31]. https://www.xfel.eu/facility/beamlines/index_eng.html. Accessed 31 July 2026

27. Koliyadu JCP, Letrun R, Kirkwood HJ, Liu J, Jiang M, Emons M, et al. Pump–probe capabilities at the SPB/SFX instrument of the European XFEL. J Synchrotron Radiat. 2022;29:1273–83. https://doi.org/10.1107/S1600577522006701

28. White TA, Kirian RA, Martin AV, Aquila A, Nass K, Barty A, et al. *CrystFEL* : a software suite for snapshot serial crystallography. J Appl Crystallogr. 2012;45:335–41. https://doi.org/10.1107/S0021889812002312

29. Kciuk M, Marciniak B, Mojzych M, Kontek R. Focus on UV-Induced DNA Damage and Repair—Disease Relevance and Protective Strategies. Int J Mol Sci. 2020;21:7264. https://doi.org/10.3390/ijms21197264

30. Nagpal A, Dhankhar D, Cesario TC, Li R, Chen J, Rentzepis PM. Thymine dissociation and dimer formation: A Raman and synchronous fluorescence spectroscopic study. Proc Natl Acad Sci. 2021;118:e2025263118. https://doi.org/10.1073/pnas.2025263118

31. Park H, Zhang K, Ren Y, Nadji S, Sinha N, Taylor J-S, et al. Crystal structure of a DNA decamer containing a *cis-syn* thymine dimer. Proc Natl Acad Sci. 2002;99:15965–70. https://doi.org/10.1073/pnas.242422699

32. Sandelin E, Johannesson J, Wendt O, Brändén G, Neutze R, Wallentin C-J. Characterization and evaluation of photolabile (μ-peroxo)(μ-hydroxo)bis[bis(bipyridyl)cobalt caged oxygen compounds to facilitate time-resolved crystallographic studies of cytochrome c oxidase. Photochem Photobiol Sci. 2024;23:839–51. https://doi.org/10.1007/s43630-024-00558-x

33. Mikšovská J, Larsen RW. Photothermal studies of the photodegradation of (μ-peroxo)(μ-hydroxo)bis[bis(bipyridyl)Co(III)] and (μ-peroxo)(μ-hydroxo)bis[bis(phenanthroline)Co(III)] complexes in water. Inorganica Chim Acta. 2003;355:116–20. https://doi.org/10.1016/S0020-1693(03)00272-X

34. MacArthur R, Sucheta A, Chong FF, Einarsdóttir O. Photodissociation of a (mu-peroxo)(mu-hydroxo)bis[bis(bipyridyl)-cobalt(III)] complex: a tool to study fast biological reactions involving O2. Proc Natl Acad Sci. 1995;92:8105–9. https://doi.org/10.1073/pnas.92.18.8105

35. Olson JS. Kinetic mechanisms for O2 binding to myoglobins and hemoglobins. Mol Aspects Med. 2022;84:101024. https://doi.org/10.1016/j.mam.2021.101024

36. Vojtěchovský J, Chu K, Berendzen J, Sweet RM, Schlichting I. Crystal Structures of Myoglobin-Ligand Complexes at Near-Atomic Resolution. Biophys J. 1999;77:2153–74. https://doi.org/10.1016/S0006-3495(99)77056-6

37. Park J, Park S, Nam KH. Preliminary Serial Femtosecond Crystallography Studies of Myoglobin from Equine Skeletal Muscle. Crystals. 2025;15:905. https://doi.org/10.3390/cryst15100905

38. Barends TRM, Foucar L, Ardevol A, Nass K, Aquila A, Botha S, et al. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. Science. New York, N.Y.; 2015;350:445–50. https://doi.org/10.1126/science.aac5492

39. Barends TRM, Gorel A, Bhattacharyya S, Schirò G, Bacellar C, Cirelli C, et al. Influence of pump laser fluence on ultrafast myoglobin structural dynamics. Nature. Nature Publishing Group; 2024;626:905–11. https://doi.org/10.1038/s41586-024-07032-9

40. Cianci M, Bourenkov G, Pompidor G, Karpics I, Kallio J, Bento I, et al. P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. J Synchrotron Radiat. 2017;24:323–32. https://doi.org/10.1107/S1600577516016465

41. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of *Coot*. Acta Crystallogr D Biol Crystallogr. 2010;66:486–501. https://doi.org/10.1107/S0907444910007493

42. Kabsch W. *XDS*. Acta Crystallogr D Biol Crystallogr. 2010;66:125–32. https://doi.org/10.1107/S0907444909047337

43. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. *Phaser* crystallographic software. J Appl Crystallogr. 2007;40:658–74. https://doi.org/10.1107/S0021889807021206

44. Vojtěchovský J, Chu K, Berendzen J, Sweet RM, Schlichting I. Crystal Structures of Myoglobin-Ligand Complexes at Near-Atomic Resolution. Biophys J. 1999;77:2153–74. https://doi.org/10.1016/S0006-3495(99)77056-6

45. Marguet S, Markovitsi D. Time-Resolved Study of Thymine Dimer Formation. J Am Chem Soc. American Chemical Society; 2005;127:5780–1. https://doi.org/10.1021/ja050648h

46. Singh JS. FTIR and Raman spectra and fundamental frequencies of biomolecule: 5-Methyluracil (thymine). J Mol Struct. 2008;876:127–33. https://doi.org/10.1016/j.molstruc.2007.06.014

47. Ruiz-Chica AJ, Medina MA, Sánchez-Jiménez F, Ramírez FJ. On the interpretation of Raman spectra of 1-aminooxy-spermine/DNA complexes. Nucleic Acids Res. 2004;32:579–89. https://doi.org/10.1093/nar/gkh232

48. Park H, Zhang K, Ren Y, Nadji S, Sinha N, Taylor J-S, et al. Crystal structure of a DNA decamer containing a *cis-syn* thymine dimer. Proc Natl Acad Sci. 2002;99:15965–70. https://doi.org/10.1073/pnas.242422699

49. Egli M. Diffraction Techniques in Structural Biology. Curr Protoc Nucleic Acid Chem [Internet]. 2010 [cited 2026 July 16];41. https://doi.org/10.1002/0471142700.nc0713s41

50. Latychevskaia T, Fink H-W. Three-dimensional double helical DNA structure directly revealed from its X-ray fiber diffraction pattern by iterative phase retrieval. Opt Express. Optica Publishing Group; 2018;26:30991–1017. https://doi.org/10.1364/OE.26.030991

51. McPherson A, Kuznetsov YG. Mechanisms, kinetics, impurities and defects: consequences in macromolecular crystallization. Acta Crystallogr Sect F Struct Biol Commun. 2014;70:384–403. https://doi.org/10.1107/S2053230X14004816

52. Marvin DA. Fibre diffraction studies of biological macromolecules. Prog Biophys Mol Biol. 2017;127:43–87. https://doi.org/10.1016/j.pbiomolbio.2017.04.005

53. De Zitter E, Coquelle N, Oeser P, Barends TRM, Colletier J-P. Xtrapol8 enables automatic elucidation of low-occupancy intermediate-states in crystallographic studies. Commun Biol. 2022;5:640. https://doi.org/10.1038/s42003-022-03575-7

54. Grünbein ML, Stricker M, Nass Kovacs G, Kloos M, Doak RB, Shoeman RL, et al. Illumination guidelines for ultrafast pump–probe experiments by serial femtosecond crystallography. Nat Methods. 2020;17:681–4. https://doi.org/10.1038/s41592-020-0847-3

55. Karplus PA, Diederichs K. Assessing and maximizing data quality in macromolecular crystallography. Curr Opin Struct Biol. 2015;34:60–8. https://doi.org/10.1016/j.sbi.2015.07.003

9. Appendix

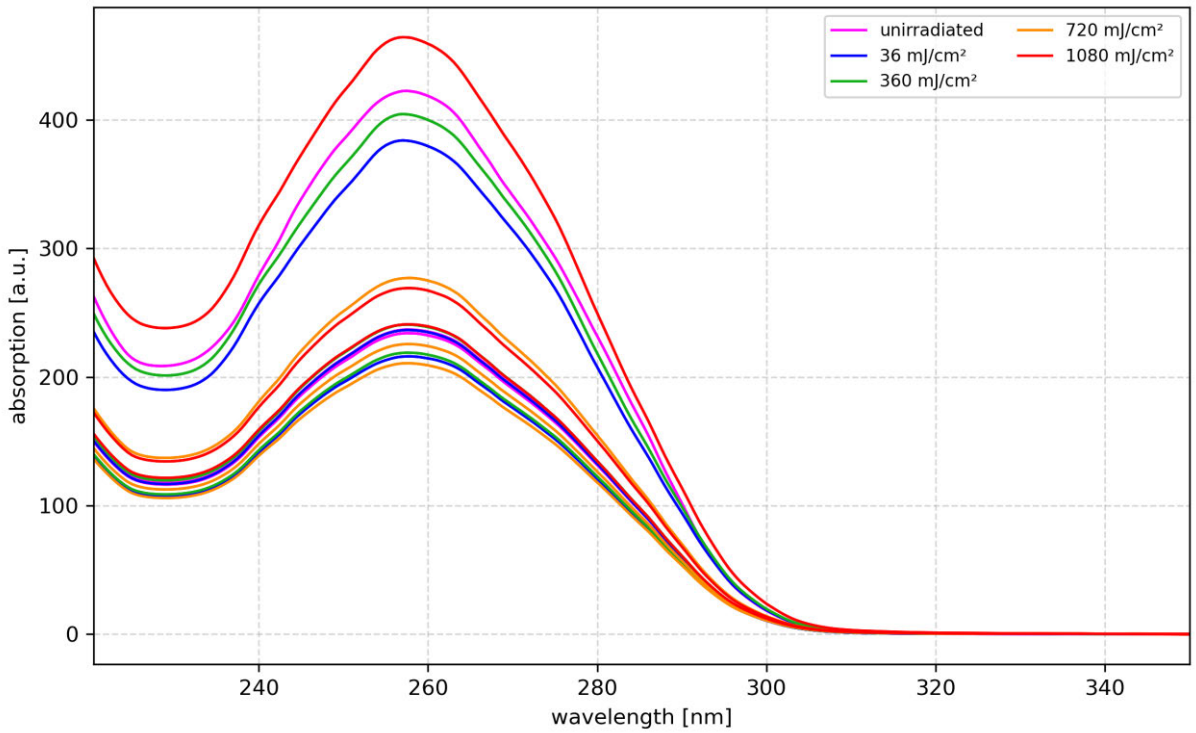


Figure 25: Collected UV-absorption spectra . Displayed are all UV-absorption spectra collected at 220-350 nm. Samples were irradiated at 36 mJ/cm² (blue), 360 mJ/cm² (green), 720 mJ/cm² (orange) or 1080 mJ/cm² (red). Unirradiated DNA (magenta) was measured as reference.

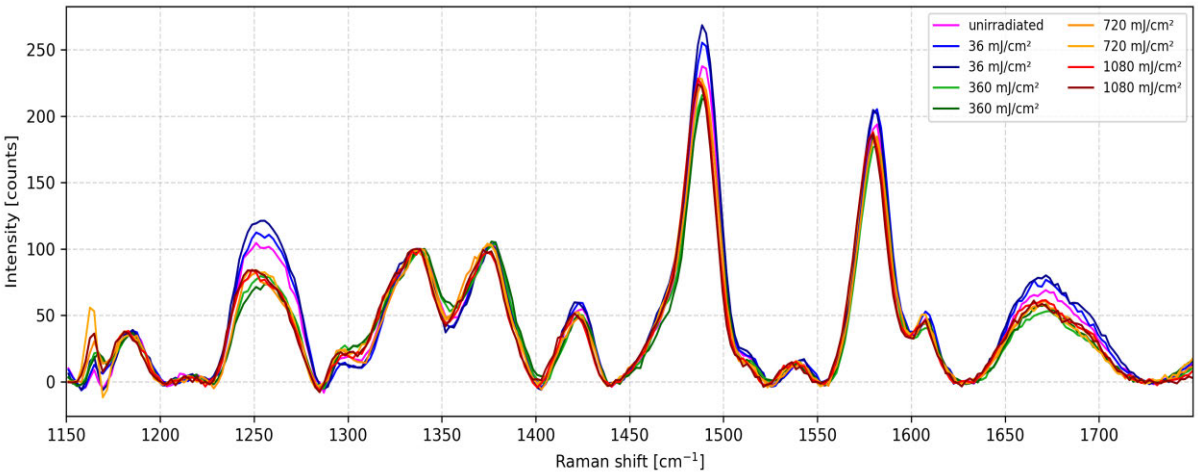


Figure 26: Collected Raman spectra . Displayed are the single collected Raman spectra. Raman shifts were measured from 1150-1750 cm^{-1} from unirradiated (magenta) and irradiated at 36 mJ/cm² (blue), 360 mJ/cm² (green), 720 mJ/cm² (orange) and 1080 mJ/cm² (red) 10-mer DNA samples.

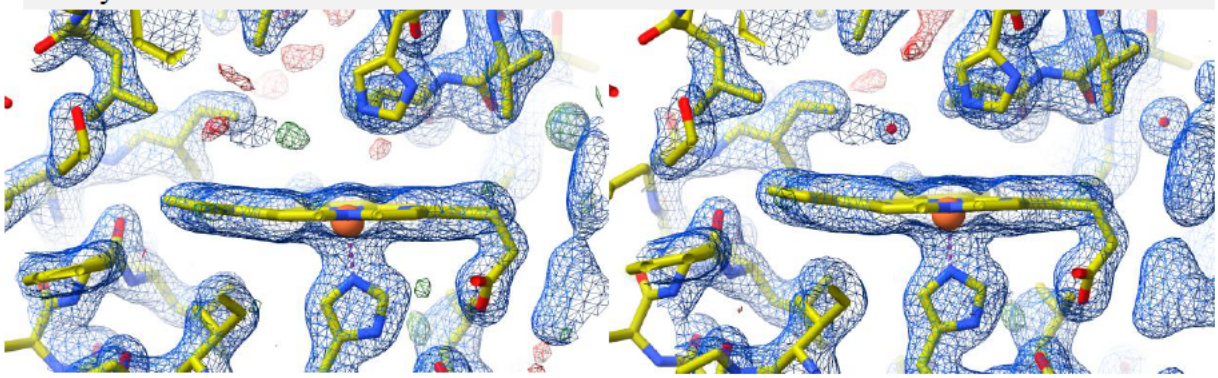
Table 11: Screened and collected unirradiated crystals. Listed below are the unirradiated crystals screened for diffraction. Crystals that showed potential diffraction signals in the screening were used for full data set collection. *Given are the number of crystals used for a full data set collection in parentheses, the number of acquired diffraction patterns per set and the maximum resolution of acquired data sets at the detector's edges.

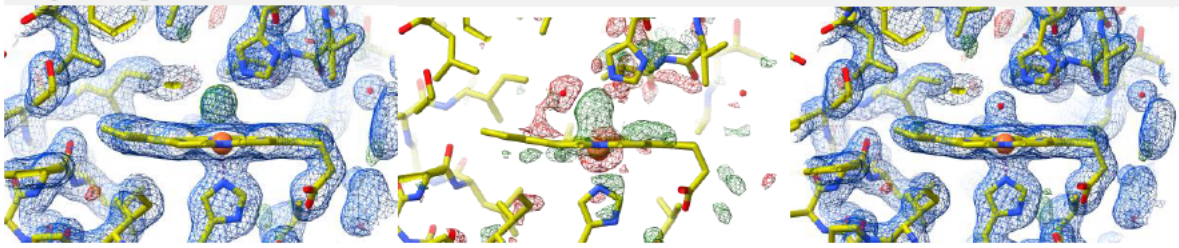
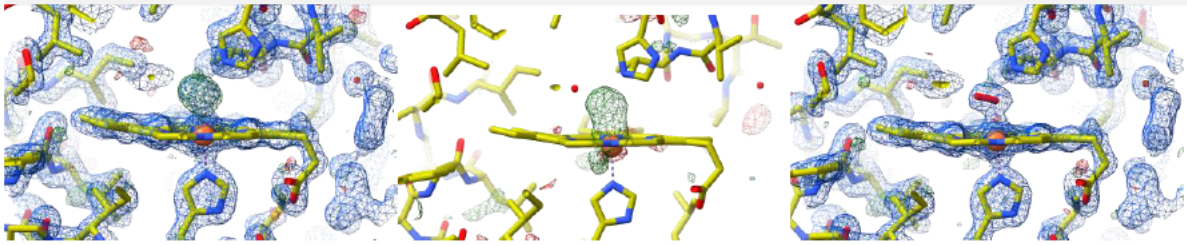
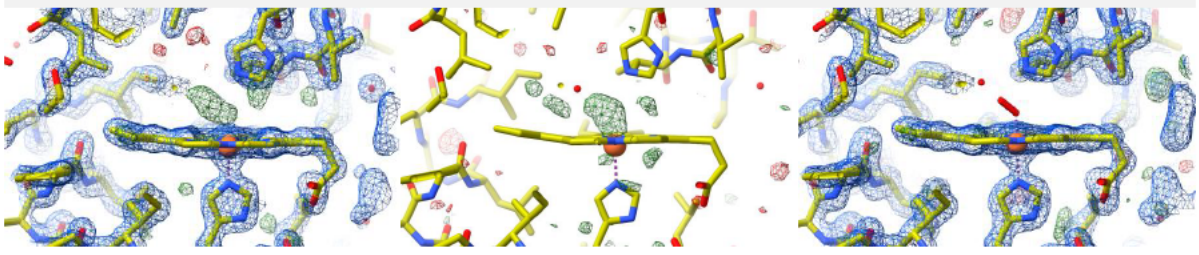
Oligo	Crystallisation condition	Additional cryoprotectant	screened crystals	Data collection*
10-mer	10.1, 25% (w/v) PEG, 2.0 mM DNA	paraffin oil	4	-
	10.1, 35% (w/v) PEG, 2.0 mM DNA	-	5	(2) 360: 1.5 Å, 2.58 Å
	10.2, 35% (w/v) PEG, 2.0 mM DNA	-	6	(1) 360: 1.2 Å
	10.3, 25% (w/v) PEG, 2.0 mM DNA	-	1	-
20-mer	20.1, 25% (w/v) PEG, 2.0 mM DNA	-	1	-
		paraffin oil	1	-
	20.1, 25% (w/v) PEG, 1.5 mM DNA	paraffin oil	2	(2) 3600: 2.58 Å
	20.2, 25% (w/v) PEG, 2.0 mM DNA	paraffin oil	1	-
	20.2, 25% (w/v) PEG, 2.5 mM DNA	paraffin oil	2	(2) 3600: 2.58 Å
	20.3, 35% (w/v) PEG, 2.0 mM DNA	-	6	(3) 360: 1.5 Å, 1.2 Å, 2.58 Å

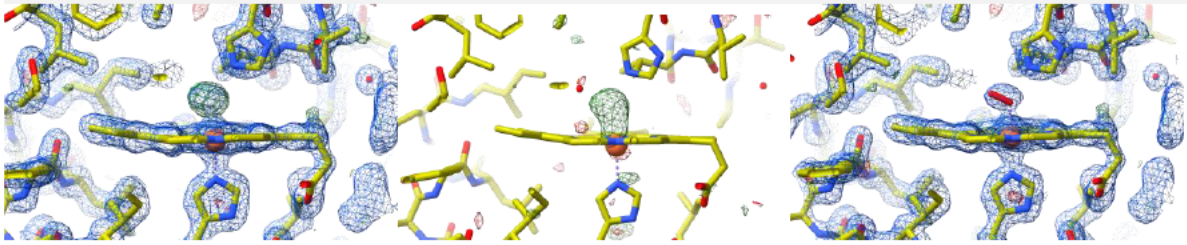
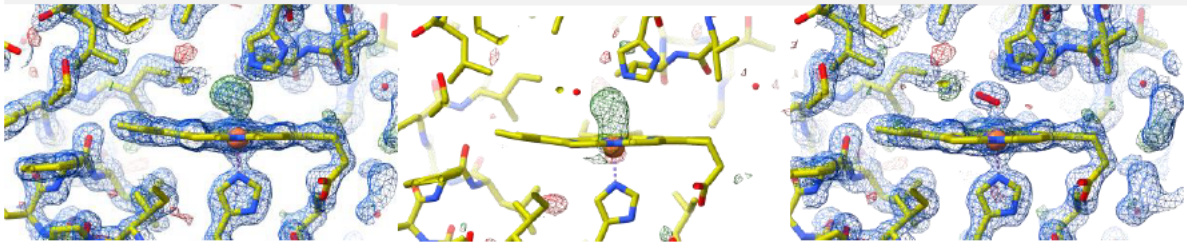
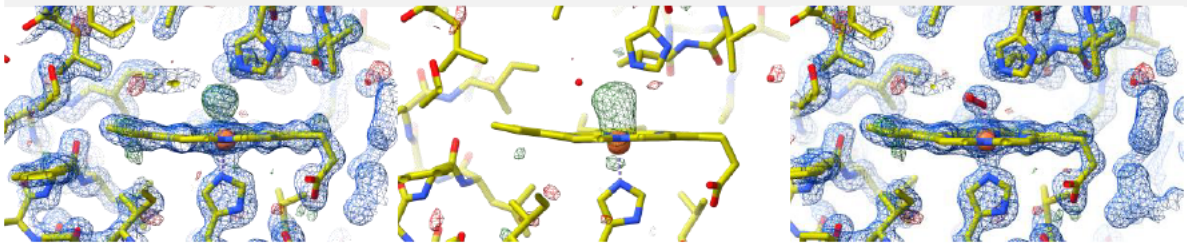
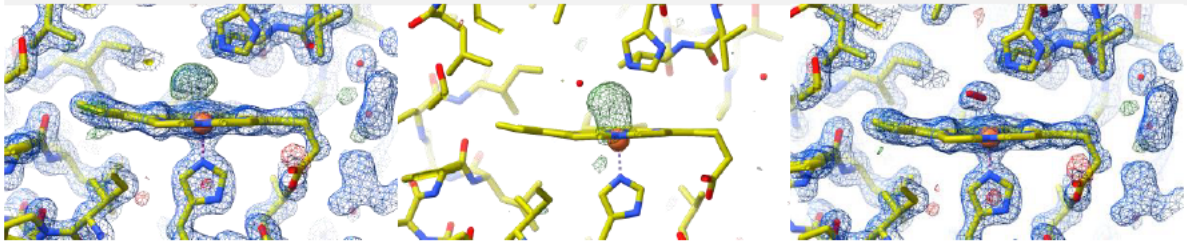
Table 12: Screened and collected irradiated crystals. Listed below are the irradiated crystals screened for diffraction. Crystals that showed potential diffraction signals in the screening were used for full data set collection. *Given are the number of crystals used for a full data set collection in parentheses, the number of acquired diffraction patterns per set and the maximum resolution of acquired data sets at the detector's edges.

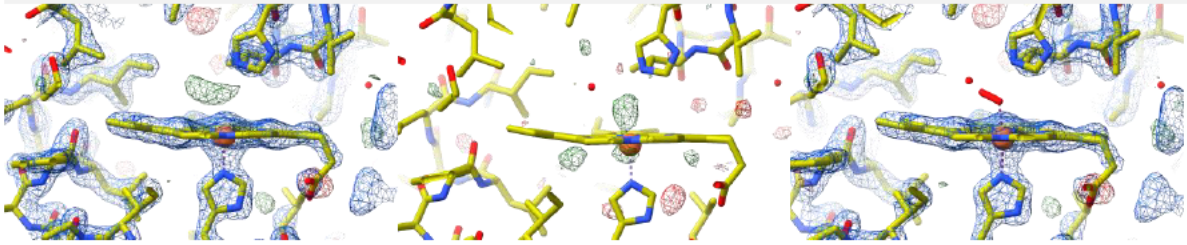
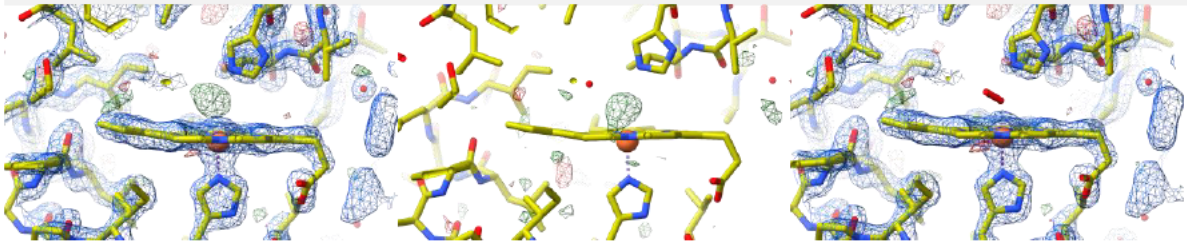
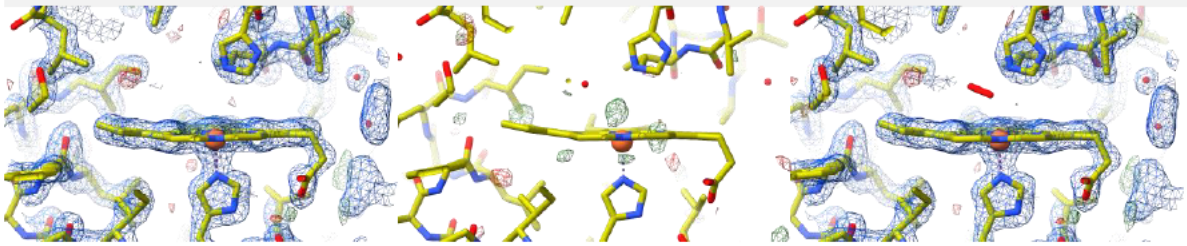
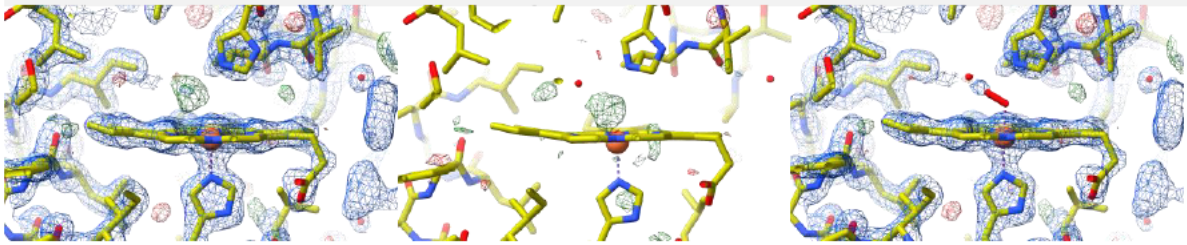
Oligo	Crystallisation condition	Irradiation time [min]	Screened crystals	Data collection*
10-mer	10.1, 35% (w/v) PEG, 2.0 mM DNA	3	5	(2) 3600:1.50 Å, 2.0 Å
		30	4	(1) 3600: 1.50 Å
		60	5	(3) 3600: 1.50 Å, 2.58 Å
		90	5	(2) 3600: 2.58 Å

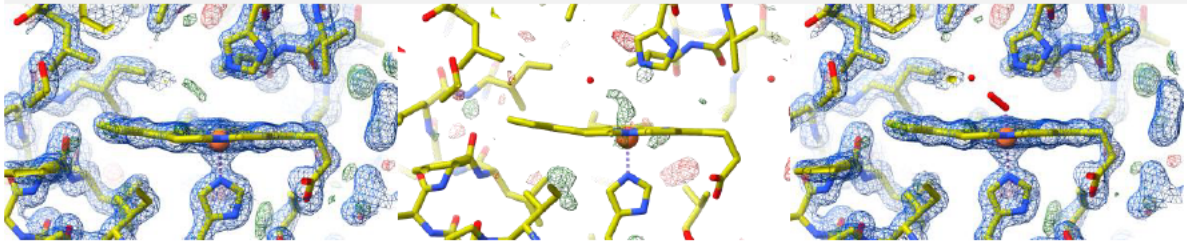
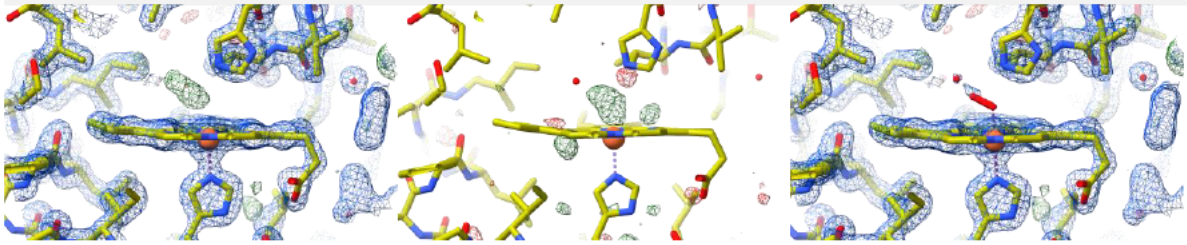
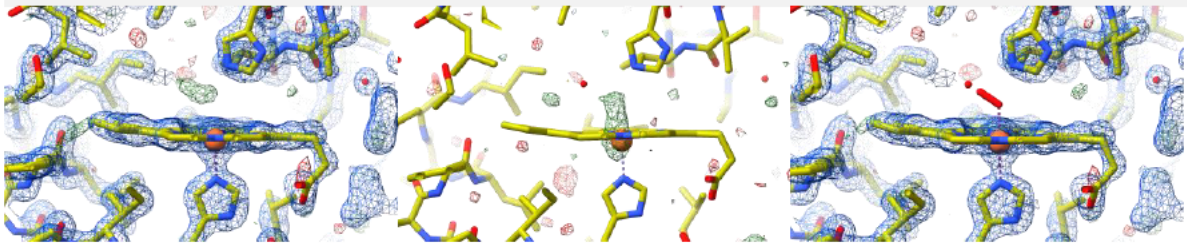
Table 13: List of all analysed myoglobin structures. Displayed are the structures before molecules were added to the heme pocket, the isomorphous difference maps and the final refined structure. Blue: 2mFo-DFc map at an r.m.s.d of 1.50, green: positive mFo-DFc densities at +3σ, red: negative mFo-DFc densities at -3σ.

Structure without ligand	Final refined structure
reference structures	
deoxyMb	

Structure without ligand	Isomorphous difference map	Final refined structure
oxyMb/*aquomet*-Mb		
shared-reservoir deoxyMb + HPBC, no laser		
mix-and-inject deoxyMb + HPBC, no laser		

Structure without ligand	Isomorphous difference map	Final refined structure
shared-reservoir deoxyMb + HPBC, 300 mJ/cm2, 100 ns		
on 1		
off 1		
off 2		
off 3		

Structure without ligand	Isomorphous difference map	Final refined structure
mix-and-inject deoxyMb + HPBC, 194 mJ/cm2, 10 ns		
on 1		
		
off 1		
		
off 2		
		
off 3		
		

Structure without ligand	Isomorphous difference map	Final refined structure
mix-and-inject deoxyMb + HPBC, 79 mJ/cm2, 4 µs		
on 1		
off 1		
off 2		
off 3	