STUDYING PROTEIN DYNAMICS WITH X-RAY FREE-ELECTRON LASERS:

OPPORTUNITIES & LIMITATIONS

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich Physik der Freien Universität Berlin

vorgelegt von
Marie Luise Grünbein

Berlin 2020

Erstgutachter:	Prof. Dr. Robert Bittl
Zweitgutachter:	Dr. Ilme Schlichting
Tag der Disputation:	19.02.2021

Summary

Protein structure and function are intimately connected. To deduce the mechanisms underlying specific functions, it is therefore of high interest to investigate structural changes during a reaction. Recently, the development of serial femtosecond crystallography (SFX) at X-ray free-electron lasers (XFELs) has attracted a great deal of attention by enabling time-resolved (TR) experiments at atomic spatial and femtosecond temporal resolution, thereby allowing unprecedented insight into protein dynamics.

The high intensity of the XFEL pulse destroys any sample that has been exposed to the focused beam. A new protein crystal thus needs to be supplied for each pulse. This is typically achieved using a continuously flowing jet. For light-triggered reactions, an optical pulse starts the reaction in crystals of photosensitive proteins and the X-ray pulse then interrogates the system after a given time interval. For such experiments there are two main issues: First, appropriate conditions have to be found for triggering the reaction of interest. This is particularly important given that to date all published ultrafast TR-SFX experiments used excessive (beyond biological relevance) excitation. Second, the measurement of weak signals is severely limited by the low data collection rate (≤ 120 Hz) at first-generation XFELs. Moreover high sample consumption is an issue at these X-ray sources.

The goals of this thesis were therefore twofold: In the first part, techniques were developed to enable studying the ultrafast isomerization following photon absorption by bacteriorhodopsin in a TR-SFX experiment. Extending these results, light-matter interactions changing the incident excitation intensity were quantified based on experiments and calculations. This allowed establishing guidelines how to generally determine appropriate excitation conditions in SFX employing light triggering. These findings are fundamental to avoid multiphoton artefacts arising from excessive excitation and are thus essential for studying biological reactions which take place almost exclusively in the single photon regime.

In the second part of this thesis, opportunities and challenges of SFX experiments at next-generation XFELs were explored. These new machines generate X-ray pulses at MHz peak repetition rate and promise significantly higher throughput and more efficient sample usage. However, the short spacing between pulses introduces new challenges: it needs to be ensured that fresh sample is supplied sufficiently fast for each X-ray pulse. Moreover, it has been shown that the XFEL pulse launches shock waves in the sample carrying jet. These may damage sample probed by subsequent pulses. Here, first experiments at MHz peak repetition rate were conducted to investigate both issues. It was demonstrated that data collection of undamaged sample is indeed possible at 1.1 MHz repetition rate. At shorter pulse intervals (corresponding to 4.5 and 9.2 MHz), shock wave induced damage may lead to a significant loss in diffraction resolution of the crystal and even to structural changes in the protein.

Together, the results of this thesis delineate the limitations of (TR-) SFX due to XFEL induced shock damage and pave the way towards exploiting the promising capabilities of MHz XFELs, in particular for studying biologically relevant light-triggered reactions in proteins.

Zusammenfassung

Proteinstruktur und –funktion sind eng miteinander verbunden. Um die zugrundeliegenden Mechanismen aufzuklären, ist es daher von hohem Interesse, strukturelle Änderungen während einer Reaktion zu verfolgen. Die Entwicklung serieller Femtosekunden-Kristallographie (SFX) an Freie-Elektronen-Lasern im Röntgenbereich (XFEL) hat folglich durch die einmalige Kombination von atomarer räumlicher und Femtosekunden zeitlicher Auflösung viel Aufmerksamkeit erregt, da sie beispiellose Einblicke in die Struktur und Dynamik von Proteinen erlaubt.

XFEL Pulse besitzen eine solch hohe Intensität, dass die Probe letztendlich zerstört und für jeden Puls ein neuer Proteinkristall benötigt wird. Ein Flüssigkeitsstrahl (Jet) liefert daher kontinuierlich frisches Material. Mit diesem Ansatz lassen sich auch lichtgesteuerte Reaktionen beobachten, indem ein optischer Puls die Reaktion in einem Kristall aus photosensitiven Proteinen startet, und der Röntgenpuls nach einer festgelegten Zeit das System abfragt. Bei dieser Herangehensweise gibt es zwei grundlegende Probleme: Erstens müssen geeignete Bedingungen zum Starten der Reaktion gefunden werden. Dies ist besonders wichtig, da bis heute alle publizierten ultraschnellen Experimente extreme Anregungsbedingungen (außerhalb biologischer Relevanz) verwendet haben. Zweitens ist an XFELs der ersten Generation die Messung schwacher Signale durch die geringe Repetitionsrate (≤ 120 Hz) limitiert, die zudem zu einem hohen Probenverbrauch führt.

Diese Arbeit hat daher zwei Ziele: Im ersten Teil wurden Methoden entwickelt, die die Grundlage für das Verfolgen der ultraschnellen lichtinduzierten Isomerisierung in Bacteriorhodopsin mittels SFX bildet. Anknüpfend daran wurden die Anregungsintensität ändernde Licht-Materie-Wechselwirkungen mithilfe von Experimenten und Berechnungen quantifiziert, sodass ein allgemeiner Leitfaden für die Bestimmung passender Anregungsbedingungen aufgestellt werden konnte. Dies ist ein entscheidender Schritt für das Vermeiden biologisch irrelevanter Multiphotonen-Effekte.

Im zweiten Teil der Arbeit wurden die Chancen und Herausforderungen von SFX an neuen XFELs untersucht, die Röntgenpulse mit bis zu MHz Wiederholrate produzieren können und dadurch versprechen, Durchsatz und Probeneffizienz zu erhöhen. Durch die kurzen Pulsabstände entstehen jedoch neue Probleme: einerseits muss die Zufuhr neuer Kristalle in den Strahl schnell genug geschehen. Andererseits wurde gezeigt, dass der XFEL Puls im Jet Schockwellen auslöst, die die Probe schädigen und so die Messung mit schnell aufeinanderfolgenden Pulsen beeinträchtigen könnte. In dieser Arbeit wurden erste Experimente bei MHz Wiederholrate durchgeführt und beide Problematiken untersucht. Messungen bei 1.1 MHz konnten erfolgreich ohne Beeinträchtigung durchgeführt werden. Es wurde aber auch gezeigt, dass bei kürzeren Pulsintervallen (entsprechend 4.5 und 9.2 MHz) die Schockwelle die Probe schädigen kann und dadurch zu einer reduzierten Auflösung der Kristalle, sowie zu Strukturänderungen im Protein führen können.

Die Ergebnisse dieser Arbeit sind wegweisend für das Ausschöpfen der vielversprechenden Möglichkeiten von MHz XFELs, insbesondere für das Beobachten biologisch relevanter, ultraschneller, lichtinduzierter Reaktionen in Proteinen.

Table of Contents

Publications

Major parts of this thesis have been or will be published in the following **research articles**:

1. **Grünbein, M.L.**, et al. "Illumination guidelines for ultrafast pump-probe experiments by serial femtosecond crystallography." Nature Methods 17(7): 681-684 (2020).
   Parts of this publication have been reproduced in chapters 1 – 3, 5, 7, 10 and 11.

2. **Grünbein, M.L.**, et al. "Effect of X-ray free-electron laser-induced shockwaves on haemoglobin microcrystals delivered in a liquid jet" Nature Communications 12(1): 1672 (2021).
   Parts of this publication have been reproduced in chapters 1 – 4 and 8 – 10.

3. **Grünbein, M.L.**, et al. "Observation of shock-induced protein crystal damage during megahertz serial femtosecond crystallography" Physical Review Research 3(1): 013046 (2021).
   Parts of this publication have been reproduced in chapters 1 – 4 and 8 – 10.

4. Gorel, A.*, **Grünbein, M.L.** *, et al. "Shock damage analysis in serial femtosecond crystallography data collected at MHz X-ray free-electron lasers." Crystals 10(12): 1145 (2020). *contributed equally
   Parts of this publication have been reproduced in chapters 1 – 4 and 8 – 10.

5. Nass Kovacs, G., Colletier, J.-P., **Grünbein, M.L.**, Yang, Y., Stensitzki, T., et al. "Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin." Nature Communications 10(1): 3177 (2019).
   Parts of this publication have been reproduced in chapters 3, 4, 6, 7 and 10.

6. **Grünbein, M.L.**, et al. "MHz data collection of a microcrystalline mixture of different jack bean proteins." Scientific Data 6(1): 18 (2019).
   Parts of this publication have been reproduced in chapters 1 – 4 and 8 – 10.

7. **Grünbein, M.L.**, et al. "Megahertz data collection from protein microcrystals at an X-ray free-electron laser." Nature Communications 9(1): 3487 (2018).
   Parts of this publication have been reproduced in chapters 1 – 4 and 8 – 10.

8. **Grünbein, M.L.**, et al. "Velocimetry of fast microscopic liquid jets by nanosecond dual-pulse laser illumination for megahertz X-ray free-electron lasers." Optics Express 26(6): 7190-7203 (2018).
   Parts of this publication have been reproduced in chapters 1 – 3 and 8 – 10.

As well as in the following **review article** and **book chapter**:

9. **Grünbein, M.L.**, G. Nass Kovacs, M. Kloos, A. Gorel, R.B. Doak, R.L. Shoeman, T.R.M. Barends and I. Schlichting. "Crystallographic studies of rhodopsins –structure and dynamics" chapter in "Rhodopsin: Methods and Protocols" edited by Valentin Gordeliy as part of the series "Methods in Molecular Biology" to be published by Humana Press, Springer Protocols
   Parts of this publication have been reproduced in chapters 1 – 2, 5 – 7 and 10.

10. **Grünbein, M.L.** and G. Nass Kovacs "Sample delivery for serial crystallography at free-electron lasers and synchrotrons." Acta Crystallographica Section D 75(2): 178-191 (2019).
    Parts of this publication have been reproduced in chapters 1 – 2 and 5 – 10.

A detailed description of my own contributions in the experiments presented in this thesis is delineated in detail in chapter 2.4.

Text and display items (Figures, Tables) taken from these manuscripts were produced by myself unless otherwise stated. In particular, any text reproduced literally from these publications has been written by myself. The direct supervisor of this work, Dr. Ilme Schlichting, confirms this.

1 Introduction

Three-dimensional structural knowledge plays a key role in understanding molecular functions, mechanisms and dynamics. This is important in a variety of fields ranging from fundamental research in all natural sciences to diverse applications such as improving the efficiency of catalysts, adapting material properties to a given task or studying drug targets for structure-based drug design. Of particular interest to all of these topics is the study of temporary structural configurations adopted along a reaction path which may help to elucidate the underlying mechanisms.

Light is an important trigger for many reactions. In biology proteins containing a photoactive chromophore enable light to serve as an energy resource that can be converted into chemical energy (e.g. in photosynthesis) or to act as a carrier of information initiating intra- or intercellular signalling (e.g. in vision). After absorption of a photon by the chromophore, specific interactions of the excited chromophore with the surrounding protein matrix lead to structural changes ultimately resulting in execution of the protein's specific function. Atomic rearrangement along the reaction pathway typically covers a long time scale of up to seconds, with first movements starting within femtoseconds (fs) after reaction initiation[14].

The possibility of triggering a light-driven reaction using femtosecond laser pulses enables performing ultrafast experiments to follow and understand the underlying mechanisms, allowing for example to learn how a reaction's efficiency can be optimized for a particular application. Previously this was only accessible via femtosecond spectroscopy but, while delivering a range of invaluable information on the energetic states occupied during the reaction, direct 3D structural information cannot be obtained. Structural information at atomic resolution are predominantly obtained via means of X-ray crystallography and used to be limited to a temporal resolution of ~100 ps at synchrotrons[15]. To improve the temporal resolution and allow observation of structural changes directly following bond breakage and bond formation, shorter X-ray pulses of sufficient intensity are required. With the advent of hard X-ray free-electron lasers (XFELs) the generation of ultra-short very intense X-ray pulses became possible, extending time resolution of structural meas-

urements down to femtoseconds[16], the time scale of molecular motion. However, the extraordinarily intense XFEL pulse also deposits large amounts of energy in the sample, leading to a very high energy density that ultimately destroys the sample within a single pulse[17]. Fresh sample must therefore be transported into the scattering region prior to arrival of each X-ray pulse, for example by means of a continuous sample stream or jet. Three-dimensional structural information can then be obtained by measuring the diffraction pattern of many identical crystals in different orientations, and merging these measurements to assemble a complete set of structure factors[18]. This serial approach to data collection combined with the femtosecond duration of the XFEL pulse is termed serial femtosecond crystallography (SFX)[19].

Given an electromagnetic trigger, time-resolved SFX experiments have thus opened up the possibility to observe the structure of transient states femtoseconds after reaction initiation. A range of initial time-resolved studies at XFELs exploited these advantages and elucidated ultrafast light-triggered reactions in a variety of proteins, including ligand dissociation[16] and trans/cis isomerization[20,21]. However, a difficulty in these experiments is that X-ray crystallography measures the superimposed signal of all species simultaneously. This is in contrast to spectroscopy which can distinguish small concentration differences of multiple states if these have unique spectral bands.

Any time-resolved crystallography experiment employing optical triggering therefore needs to ensure suitable means to excite the protein into the desired reaction pathway, and avoid forming other species via unwanted pathways. Additionally, for studying biologically relevant reactions optical excitation conditions need to be within the linear regime, in which one-photon processes are excited and the signal scales linearly with excitation intensity[22]. However, the attempt to maximise the fraction of excited molecules prompted previous experimenters[16,20,21] to work with very high pump laser intensities (of several hundred GW/cm^2). While these conditions may maximise light-induced differences in the deduced structures, it is unclear to what extent, if at all, these differences are due to biologically relevant structural changes and how many parallel pathways have been opened up[22].

The occupancies of intermediate states in the biologically relevant single-photon regime are generally low, requiring a lot of data to be collected to achieve a meaningful signal-to-noise ratio. While measurement time is extremely limited at first-generation XFELs due to

the limited number of facilities and a maximum pulse repetition rate of only 120 Hz[23], next-generation XFELs capable of running at MHz peak repetition rate promise to increase data collection rates significantly and thereby promise to facilitate observation of weak signals. However, data collection at MHz rate comes with new challenges: fresh sample needs to be supplied sufficiently fast for each X-ray pulse and it needs to be ensured that sample probed by subsequent pulses is not altered by the impact of previous pulses.

For rigorous experiments on protein dynamics using XFELs, the desired dynamics must be triggered reliably, and it is pivotal to know how experimental artefacts can be avoided. In the course of this thesis techniques were therefore developed to allow time-resolved experiments triggering the desired reaction, avoiding artificial XFEL-induced dynamics and thereby to explore the limitations of (time-resolved) SFX.

Specifically, techniques were developed to enable a time-resolved experiment on bacteriorhodopsin (bR), allowing to analyse the structural changes induced by photon absorption. bR is a light-driven proton pump that belongs to the family of rhodopsins containing retinal as their light-absorbing chromophore[24]. Rhodopsins are involved in a wide range of biological light-sensitivity, starting from vision, the regulation of the circadian rhythm to light-energy conversion in different microbial organisms[24]. The aim of the time-resolved SFX studies on bR was to understand the role of interaction between protein matrix and chromophore in directing the reaction. A better understanding of these processes would be of help for elucidating mechanisms of similar light receptors, as well as for a potential application as optogenetic tools.

Extending from the experiment on bR, a general characterization of appropriate (biologically relevant) photo excitation conditions in pump-probe experiments on protein microcrystals was performed, delineating specifically how to avoid multiphoton artefacts. For this purpose, a protocol was designed and characterized that allows determining photo excitation conditions within the sample, taking intensity changes along the optical path into account.

Finally, the possibilities and limitations of SFX measurements at MHz repetition rate were explored. To enable measurements at MHz repetition rate, fast jets capable of replenishing sample with sufficient speed need to be employed. Such jets require fast imaging methods

to characterize injection conditions. Suitable means to do so automatically were established in this thesis. Moreover, possible effects of the X-ray pulse on crystalline protein samples in liquid jets were studied to elucidate if and under what conditions experiments at MHz repetition rate may be conducted that indeed probe the protein's native state.

PART 1: EXPERIMENTAL AND THEORETICAL BACKGROUND

2 Structural biology at XFELs

The first free-electron laser capable of producing hard X-ray radiation went into operation in 2009[23], opening up the possibility to perform protein crystallography at these new X-ray sources generating pulses of unprecedented short pulse duration and intensity. However, given the profound differences in mode of operation and experimental parameters compared to synchrotrons, crystallography at XFELs in fact requires a completely different approach to experiment and analysis before new types of fundamental questions can be tackled.

2.1 X-ray free-electron lasers

At XFELs a bunch of free electrons is accelerated to nearly the speed of light in a long linear accelerator[25]. These relativistic electrons are passed into an undulator, where alternating dipole magnets generate a sinusoidal magnetic field due to which the moving electrons experience a Lorentz force perpendicular to the magnetic field direction and the direction of movement. In consequence, the electrons follow a sinusoidal trajectory, emitting electromagnetic radiation with a wavelength down to the hard X-ray regime depending on the electron energy, the undulator period and the magnetic field strength[25].

Initially, the radiation emitted by the electrons in the bunch has a random relative phase distribution. Key to achieving the ultrabright intensity characteristic for XFELs is the resonant interaction between the electron beam and the emitted radiation as both propagate along the same axis, resulting in so-called “microbunching” of the electron density that leads to an exponential growth in the radiated intensity[26]. The interaction between electron beam and the emitted radiative field introduces a longitudinal force on the electrons, the strength of which depends on the electron energy[25-27]. Because of this energy-dependent velocity modulation, the electrons are forced into smaller bunches separated by a distance equal to the emitted wavelength[26-28]. Electrons within a wavelength distance of each other emit radiation in phase (coherently), producing a higher intensity of the emitted radiation due to constructive interference, which in turn enhances the microbunching by increased interaction between the electron beam and the stronger radiative field[25]. In the

extreme limit of perfect microbunching, the emitted intensity is proportional to N_e^2, the number of electrons in the bunch squared[27*]. In this fashion the intensity of the emitted radiation grows exponentially along the undulator distance until saturation is reached when bunching is close to its maximum[25]. When these electron microbunches start spontaneously from noise, the process is called self-amplified spontaneous emission (SASE)[29].

X-ray free-electron laser radiation has unique properties: the pulses can be as short as a few femtoseconds and the beam is transversally coherent with a peak brightness (number of photons per time, unit surface area, 0.1 % band width and unit solid angle) that is up to nine orders of magnitude higher than at synchrotrons[25]. A few femtosecond long X-ray pulse thus contains as many photons as emitted by a synchrotron in the X-ray regime over a whole second[30]. XFELs thus enable unprecedented time resolution, coherent imaging experiments as well as experiments on weakly scattering samples.

2.2 Breakthrough results enabled by crystallography at XFELs

The possibility to measure weakly scattering samples at XFELs not only holds promises for the future to measure diffraction from single particles[31] but importantly also allows to measure diffraction from small crystals[32]. At synchrotrons this is limited by radiation damage. For example, this possibility is of particular interest to the study of G-protein coupled receptors (GPCRs), a family of membrane proteins that is essential in many physiological signalling pathways[33]. Malfunction of GPCRs is associated with many pathophysiological conditions including some types of cancer, diabetes or cardiovascular diseases, making these proteins important drug targets[33]. Typically, it is extremely challenging if not impossible to grow GPCR crystals large enough for the study at synchrotrons[33]. SFX thus facilitated breakthrough advances in the study of GPCRs, allowing solving previously unknown or poorly understood structures of GPCR members[34] and complexes[35-37].

Furthermore, the short pulse duration of XFELs allows the elastic scattering signal to be recorded before significant structural rearrangements induced by radiation damage occur, enabling not only room temperature measurements[19] but also damage-free structure determination of radiation-sensitive proteins containing metallocentres[38,39]. Metal centres

[*] This is the reason for the orders of magnitude increase in peak brilliance compared to synchrotrons, where intensity only scales with the number of electrons.

are likely to undergo X-ray induced photoreduction, changing the structure in the vicinity of the metal centre even under cryogenic conditions[40-42]. Data collection at XFELs in contrast enabled radiation-damage free data collection of metalloenzymes like photosystem II[39,42,43], allowing to capture light-induced reaction intermediates[44,45], or coppernitrite reductase, the most important metalloenzyme in microbial denitrification[46].

The possibility to measure diffraction from small crystals, together with the short pulse duration, make XFELs moreover ideally suited for studying reactions in a time-resolved manner. The ability to measure diffraction from small crystals is key to this purpose, allowing reactions to be triggered homogeneously throughout the molecules constituting the crystal. In case of optically triggering a reaction, the typically low penetration depth of a protein crystal limits the useful crystal size to a few micrometres only. Similarly, diffusion times of chemical compounds into protein crystals increase with increasing crystal size, again limiting the useful crystal size to a few micrometres[47].

With sufficiently short optical triggers, SFX allows light-induced reactions to be followed with femtosecond time resolution, at atomic spatial resolution. Pioneering experiments elucidated ultrafast changes upon photo excitation to follow ligand dissociation[16] or the trans/cis isomerization in fluorescent[20,48] or otherwise photoactive proteins[21], paving the way towards a detailed understanding of the underlying mechanisms and potentially opening the door to design mutants of higher quantum yield that can be applied in fluorescent microscopy techniques[20].

Additionally, efforts have been made to establish means to chemically trigger a reaction by mixing ligands or buffers to induce pH changes with protein crystals prior to X-ray diffraction[49-51]. For example, such experiments captured multiple time points of binding of an expression regulating ligand to messenger RNA[52] or an antibiotic being cleaved by a bacterial enzyme causing antibiotic resistance[51,53].

2.3 The complexity of serial femtosecond crystallography

Since the pioneering work of Laue and Bragg more than 100 years ago, X-ray crystallography has rapidly developed into a widely used method for investigating the structure of molecules. Today X-ray crystallography at synchrotrons is well-established, with all steps

of data collection from mounting and rotating the crystal in the X-ray beam over data recording by highly developed detectors to subsequent data analysis being fully automated. As explained in the following, XFELs in contrast require a completely new experimental approach, rendering all established procedures at synchrotrons inapplicable.

The intense XFEL pulse deposits large amounts of energy in the sample, leading to a very high energy density that ultimately destroys the sample within a single pulse[17]. X-ray absorption, typically via the photoelectric effect, initiates ionization cascades which increase the electrostatic energy such as to ultimately result in a Coulomb explosion of the sample[54]. Crystallography at XFELs is therefore conducted in a mode that has been named "diffraction before destruction"[54], meaning that the femtosecond long XFEL pulse is short enough to collect a measurable diffraction signal before complete destruction of the material takes place[17].

Since the crystals are destroyed when exposed to the high intensity of a focused XFEL pulse, pristine sample must be transported into the scattering region prior to arrival of each X-ray pulse. Various techniques to transport fresh sample into the scattering region have been developed for SFX[55-60]. In the simplest approach, sample can be delivered with a fixed target, e.g. a chip, which is then translated through the interaction region synchronous to the arrival of X-ray pulses[59,61-63]. Alternatively, sample can be delivered into the X-ray interaction region by means of a continuously flowing jet that is oriented perpendicularly to the X-ray optical axis[64,65] (Figure 2.1). Due to several advantages, liquid jets[60,64] are currently the dominant methodology for sample delivery: suspending the crystals in jets keep the fragile protein crystals in their native growth solution, preventing drying out of the sample, and the micron-sized diameter of jets ensures minimal background signal from the surrounding liquid. Moreover, given sufficient jet speed, the continuous replenishment ensures that sample interrogated by subsequent pulses has not been exposed to stray radiation or other damaging effects of previous pulses.

The generation of stable, thin-diameter liquid jets suitable to reliably transport crystals into the X-ray interaction region is highly challenging. A jet diameter as small as possible minimizes solution scattering contributing to the background of the protein crystal diffraction pattern. Given crystal sizes of typically 1 – 10 µm, jets of similar size are desired. To

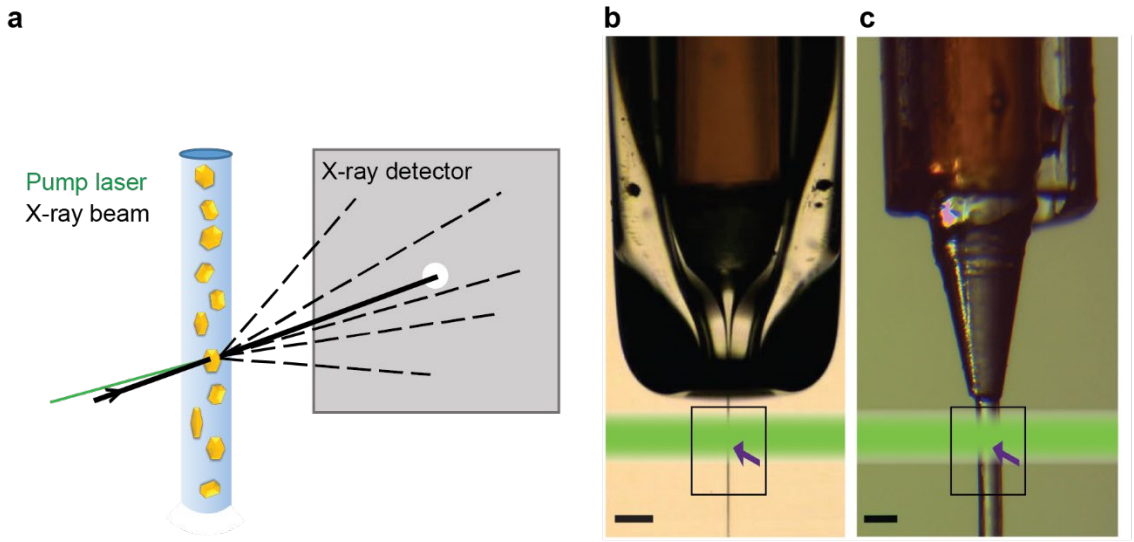


Figure 2.1 Experimental setup of SFX experiments. **a**. Microcrystals are delivered into the X-ray beam by means of a free-standing jet. The X-ray beam (black) intersects the jet, resulting in a diffraction pattern that is recorded by the detector. In case of pump-probe experiments, a pump laser (green) precedes the X-ray pulse and triggers a reaction in the crystallised molecules. Figure reproduced with formatting changes from Grünbein *et al*[4]. **b**. Liquid jet produced by a GDVN. Liquid sample is transported through the inner capillary (brown) while gas is delivered through the outer glass sheath. The gas stream focuses the liquid stream into a few micron sized jet exiting the nozzle's aperture. **c**. HVE injection producing a viscous jet. Viscous sample is pushed out of the sample capillary (brown) and forms a continuous stream that is stabilized by a coflowing gas stream. **b**, **c**. The scale bar is 100 µm. The arrow indicates the X-ray interaction region. For pump-probe experiments, a section of the sample jet is optically triggered (green shading) before X-ray probing. Both images adapted from Grünbein & Nass Kovacs[11] under the CC-BY 4.0[3].

generate these, however, it is impossible to employ equally small-diameter injection nozzles (Rayleigh nozzles) as these are easily clogged by either slightly larger crystals, families of crystals or simply contaminations in the liquid system[64]. To circumvent these issues associated with small-diameter sample lines, the so-called "gas-dynamic virtual nozzle" (GDVN) was developed[60,64]. GDVNs transport the crystal suspension in a larger diameter capillary, and, upon exit from the capillary, employ a coaxially flowing helium gas stream that focuses the liquid jet down to a diameter of typically 1 – 10 µm (Figure 2.1b).

The difficulty of working with GDVNs is due to the fabrication process, as well as the operation procedure, both of which have to be adapted to the specific sample properties and

experimental boundary conditions. A stable, continuous jet is the basal requirement to allow alignment of the jet with the X-ray beam, and thereby ensure that the fraction of X-ray pulses indeed interrogating the sample (the so-called "hit rate") is maximised. If the jet is not stable enough, either wiggling or dripping because nozzle geometry, flow rates and sample type have not been balanced correctly, a significant fraction of X-ray pulses misses the jet, dramatically decreasing the efficiency of the experiment both in terms of sample usage and experimental throughput which is highly unfavourable given the tight allowance of beam time and the difficulty of sample production. Depending on crystal size, crystal shape, nature of the surrounding liquid and boundary constraints of the experimental parameters (e.g. ambient pressure, X-ray repetition rate), different nozzle geometries need to be employed to achieve given jet diameters and jet speeds. For example, to generate a stable liquid stream exiting the nozzle into the X-ray interaction region requires different nozzle geometries for different types of protein buffers, it requires different flow rates and liquid line configurations for different crystal shapes and concentrations, and it requires different balances between liquid and gas flow rates for different target jet properties. Moreover, the generation of stable liquid jets is only one of many possible modes of ejecting sample from the nozzle: very often, if the right balance of injection parameters is not found, the ejected sample either drips or sprays[66,67]. These changes in stability cannot be detected by eye, but require high resolution imaging optics of sufficient contrast and temporal resolution.

Using jets as the most convenient and compatible method to deliver sample into the X-ray interaction region, sample flows continuously despite the pulsed nature of the data collection, resulting in many (>99 %) crystals flowing past the interaction region without ever being probed by an XFEL pulse when performing the experiment at first-generation XFELs running at up to 120 Hz. Since sample cannot be recovered afterwards, it is purely flowing to waste. In view of the often extremely valuable and difficult to obtain biological sample this situation is less than ideal. To increase efficiency of sample usage, sample volume passing the interaction region in between two X-ray pulses has to be reduced. This can be achieved either by decreasing sample flow rate or by increasing the X-ray pulse repetition rate.

Increasing efficiency of sample usage by decreasing sample flow rate in liquid jets can only be done in a marginal range, because a minimum flow rate is required for generating stable

jets from a low-viscosity liquid. In contrast, when extruding a viscous material, free-standing jets can be generated at much lower sample flow rates using similar nozzles[65] (Figure 2.1c). These jets have larger diameters (typically 50 – 100 μm), because the coflowing gas stream only stabilizes jet extrusion but cannot decrease the diameter of the extruded highly viscous stream. These types of jets are favourable for protein crystals that are grown in a viscous environment (e.g. crystals of membrane proteins[65]). Other protein systems require embedding the crystals in a viscous environment, which may work for some but not for all types of protein crystals without impairing crystal quality[68].

Increasing efficiency of sample usage by increasing the X-ray pulse repetition rate became possible with the advent of second generation XFELs capable of running at peak repetition rates of a few MHz. Compared to the previous maximum of 120 Hz at first generation XFELs[23], the tremendous increase in pulse repetition rate is enabled by employing superconducting linear accelerators. The European XFEL (EuXFEL) is the first superconducting XFEL, and is designed to deliver X-ray pulses at a maximum pulse repetition rate of up to 4.5 MHz, see[69].

Independent of the choice in sample delivery technology and X-ray repetition rate, crystallography at XFELs requires an unprecedented number of protein crystals to collect a single data set. Since only one diffraction pattern in a fixed orientation can be collected for each crystal which is then destroyed by exposure to the X-ray beam, many identical crystals in various orientations are required to obtain the complete set of structure factors[18]. Therefore, SFX experiments require often several tens of milligrams of protein or more. This introduces completely new challenges not only in obtaining sufficient amounts of protein, but also in subsequently crystallising the purified protein at scale. While screening conditions to achieve large protein crystals for structure determination at synchrotrons are well established, obtaining large quantities of micron-sized crystals proves extremely difficult and requires new approaches[15,70].

Moreover, the stochastic nature of the X-ray generation process at FELs leads to a large fluctuation in pulse properties from shot to shot: each X-ray pulse is characterized by a unique photon energy spectrum, pulse energy, fluence and pointing[25]. On top of that, a different crystal of different (unknown) orientation, different size and at a slightly different position in space covered with different amounts of surrounding liquid due to inherent changes in jet behaviour introduces additional fluctuations in diffraction intensities. All of

these properties need to be recorded and stored for each single X-ray pulse alongside the corresponding diffraction pattern. Some experiments even require the history of previous pulse parameters to be known additionally, because these may affect data collection of upcoming pulses.

As a consequence of the serial nature of data collection and the intrinsic shot-to-shot variations in experimental parameters, data analysis also reaches a new level of complexity. Not only does one need to figure out the a priori unknown orientation of each single diffraction pattern, but moreover all single-shot changes have to be taken into account when merging the thousands of single diffraction patterns in order to obtain a meaningful three-dimensional set of structure factors which can then be used for structural interpretation[18].

Alongside the mentioned challenges of crystallisation, sample delivery, single-shot parameter characterization and data analysis, it requires large efforts and manpower to run the accelerator in a stable mode, to ensure reliable X-ray detector operation and synchronisation of all single-shot recordings of pulse properties and beam line parameters, to guarantee error-free storage of the collected data in real time, to align the X-ray beam with the delivered sample jet or to steer and align additional optical lasers if these are required. Experiments at XFELs are therefore a highly complex enterprise that cannot be performed by a single person but in order to succeed require a team of experts working hand in hand.

2.4 Collaborative effort

All XFEL experiments described in this thesis were conducted in a team, with tasks subdivided between the different experts. Samples were prepared and crystallised by Elisabeth Hartmann, Gabriela Nass Kovacas and Ilme Schlichting. Bruce Doak, Marco Kloos, Robert Shoeman and myself represented the team responsible for the general experimental setup and appropriate sample delivery. During and after data collection, data storage and synchronisation was ensured by Lutz Foucar, Mario Hilpert and Chris Roome with the help of the beam line scientists on site who also operated and aligned the X-ray and optical pump laser beam. Crystallographic analyses and structural interpretation were performed by Thomas Barends, Alexander Gorel and Ilme Schlichting.

In particular, my personal responsibilities included for all experiments the individual fabrication, testing and operation of sample delivery instrumentation appropriate for each

specific experiment. I designed and implemented a suitable method to perform light-adaptation of bacteriorhodopsin microcyrstals, as well as ensured sample delivery matching the photoexcitation goals of the experiment (chapter 6). I characterized photoexcitation conditions employed in the bacteriorhodopsin experiment and generalized these findings by designing and conducting appropriate experiments and theoretical calculations to establish a protocol how appropriate photo excitation conditions can be determined for optical pump X-ray probe SFX experiments (chapter 7). To explore the possibilities and limitations of SFX experiments at MHz repetition rate at second-generation XFEL facilities, I coordinated and conducted the experiments targeting the effects of XFEL-induced shockwaves on protein crystals and analysed and interpreted its diffraction data (chapters 8 and 9).

3 Materials and Methods

3.1 Samples and sample preparation

Samples were prepared by Elisabeth Hartmann, Gabriela Nass Kovacs and Ilme Schlichting unless otherwise stated. Parts of the descriptions of samples and methods of this section have been published in references[1,2,4,5,7,9,71].

3.1.1 Viscous media

A variety of viscous media commonly employed in SFX experiments conducted in viscous jets were prepared as described in previous publications:

Lipidic cubic phase (LCP) and LCP mixed 4:1 with 40 % (w/v) Pluronic F-127 (from Sigma-Aldrich, P2443) were prepared as described in reference[68].

Sodium carboxymethylcellulose (NaCMC, obtained from Sigma-Aldrich, 21904) in lysozyme crystal storage solution was prepared by mixing 8 % (w/v) NaCMC in ultra-pure water with an equal volume of 2x concentrated lysozyme storage solution (20 % NaCl, 0.2 M Na acetate pH 4.0) as described in reference[68].

A 5.6 % (w/v) agarose gel (prepared from low gelling agarose obtained from Sigma-Aldrich, A9414)[68] in a 30 % aqueous glycerol solution was prepared as described in reference[72].

The synthetic grease Super Lube (# 21030, Synco Chemical)[56] is a commercial product and was obtained from Amazon.

3.1.2 Refractive-index matching reference media

For characterizing the optical properties of the viscous media described above, the following refractive-index matching reference media were prepared: Aqueous glycerol solutions (65 % w/w and 40 % w/w) were prepared by mixing glycerol with ultra-pure water in the appropriate weight per weight ratio. Ultra-pure water was taken from an ELGA-Purelab ultra system (ELGA LabWater, Celle, Germany).

3.1.3 Crystallisation

Hen egg-white lysozyme (HEWL) crystals were prepared as described in references[4,1,2,5].
Purified bacteriorhodopsin (bR) was obtained from Ramona Schlesinger (FU Berlin) and crystallised in LCP as described in references[7,68].
Crystals of the three jack bean proteins concanavalin A, concanavalin B and urease were prepared by Elisabeth Hartmann as described in reference[1].
Myoglobin (Mb) crystals were prepared as described in reference[16].
Haemoglobin (Hb) crystals were prepared as described in reference[9].

3.1.4 Embedding of crystals into viscous media

bR was crystallised in LCP, making embedding unnecessary. To adjust the crystal concentration and improve injection properties, the crystal-loaded mesophase was mixed with an equal volume of monoolein (from Nu-Chek Prep) and diluted (4+1) with Pluronic F-127 40 % (w/v) prior to all experiments as described in references[4,7,68].

As described in reference[4] 10 µl (1.25 µl) crystal pellet of 12 µm (1 µm) sized lysozyme microcrystals was mixed with 90 µl (98.75 µl) NaCMC equilibrated with lysozyme storage buffer using two coupled Hamilton syringes, obtaining a v/v concentration of 10 % (1.25 %). Care was taken to not introduce any air into the mixture. To obtain suspensions of half the v/v concentration, the suspension was mixed in equal ratios with additional NaCMC equilibrated with lysozyme storage buffer using two coupled Hamilton syringes.

3.2 Methods for characterizing light propagation through different jet media

This section contains descriptions of methods published in references[4].

3.2.1 Refractive index measurement

Refractive indices were determined in the visible spectrum using an Abbe refractometer (Zeiss 133968) for various wavelength ranges λ within 440 nm $\leq \lambda \leq$ 645 nm by using a white light lamp (Schott KL 2500 LED) in combination with coloured filters (Andover Cor-

poration 440FS10-25 (440±5nm), Andover Corporation 460FS10-25 (460±5 nm), Edmund Optics #65-700 (532±5 nm), Edmund Optics #65-703 (546±5 nm), Edmund Optics #65-711 (632±5 nm), Andover Corporation 645FG07-25 (long pass filter, 645 nm)).

3.2.2 UV-vis transmission measurements

A Jasco V-760 Spectrophotometer was used to measure the UV-vis transmission spectrum in a spectral range from 200 – 750 nm in steps of 2 nm with a bandwidth of 0.5 nm and a scan speed of 400 nm/min. The acceptance angle of the spectrophotometer was determined to be ≤ 12°. The baseline (no cuvette in either the measurement path or the reference path) was recorded before each set of measurements. All samples were measured against air (i.e. no cuvette in the reference beam path) such that referencing was done after data recording by subtracting the extinction spectrum of the corresponding reference medium computationally using Python 3.7. Flat demountable rectangular cells with 100 µm path length (106-0.10-40, Hellma Analytics) were used as cuvettes for all optical measurements. When loading the cuvette with 10-40 µl sample, care was taken to not introduce any air bubbles and to fill the whole volume illuminated during extinction measurements. A Hamilton syringe or a positive displacement pipette was used for loading highly viscous samples such as LCP. Two or more separate measurements per sample were made and averaged.

3.3 XFEL data analysis

This section partly contains descriptions of methods published in references [1,2,5,9].

Detector calibration and image correction

Data from the AGIPD detector at EuXFEL was calibrated using the calibration pipeline established at the EuXFEL[73,74] with constants provided by the facility and the AGIPD consortium; data from the CSPAD at LCLS was calibrated and corrected in CASS[75,76]. The positions and orientations of individual detector modules were refined as described[16].

Hit finding and indexing

CASS[75,76] was used for online data analysis, hit identification and data preprocessing. A hit is defined as an image where more than ten peaks were identified by CASS. Indexing and integration were performed with CrystFEL[77] version 0.6.3 for the first experiment at EuXFEL[1,71] and version 0.8.0 for experiments at LCLS[2,9] and the follow-up at EuXFEL[5].

Analysing data quality

For all experiments the resolution of a diffraction image at a given signal-to-noise ratio (SNR) x is determined as the best resolution of a diffraction peak with SNR ≥ x. Here, the signal-to-noise ratio of each reflection in every indexed hit was calculated as $I/\sigma(I)$ extracted from the stream file output from CrystFEL after indexing.

Forming data subsets with equalized probe pulse energies

For two existing data sets A and B having given distributions in probe pulse energy with mean μ_A and μ_B and standard deviation σ_A and σ_B the target probe pulse energy distribution was determined as a Gaussian distribution centred on $\mu=0.5\cdot(\mu_A+\mu_B)$ with standard deviation $\sigma=0.5\cdot(\sigma_A+\sigma_B)$. From this distribution N random probe pulse energies e were drawn, with N corresponding to ~57 % of the total number of data points in the smaller of the two original data sets.

To form two subsets A^* and B^* with equal probe pulse energy distributions, all drawn probe pulse energies e were compared to the probe pulse energies in the original data sets in A and B, and the hit with the closest probe pulse energies in each of A and B was added to the subsets A^* and B^*, respectively.

The effect of probe pulse energy equalization is exemplarily shown in Figure 3.1 for the Lysozyme data sets collected as part of the X-ray pump X-ray probe experiment with 122.5 ns time delay at LCLS.

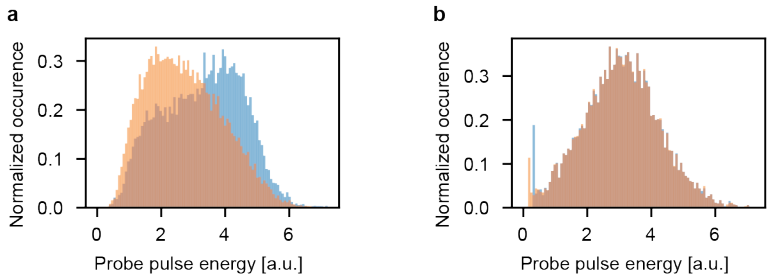


Figure 3.1 Probe pulse energy equalization. Probe pulse energies were measured by X-ray sensitive diodes. **a**. Histogram of probe pulse energies of the pump-probe (blue, average probe pulse energy 3.4 ± 1.3) and the single-pulse reference data set (orange, average probe pulse energy 2.6 ± 1.2). Due to fluctuations in the X-ray generation process, the probe pulse energy distribution differs slightly between the two data sets. **b**. Histogram of probe pulse energies of the subsets of the pump-probe (orange) and single-pulse (blue) data subsets sampled to represent the same probe pulse energy distribution centred on 3.0 ± 1.2. **a**,**b**. Lysozyme data from the X-ray pump X-ray probe experiment with 122.5 ns time delay at LCLS.

3.4 Imaging and velocimetry of fast microscopic jets

This section contains descriptions of methods published in reference[6].

3.4.1 Imaging setup and optical components

Different light sources were tested for imaging the microscopic, fast jets produced by GDVNS. For this purpose the setup shown in Figure 3.2 was employed. The light source (either a laser (Table 1) or a white-light LED (Schott KL 2500)) was positioned on the optical axis of a high-speed camera (Photron Fastcam SA-Z type 2100) facing directly into the camera to backlight the liquid jet, which flowed vertically downward through the field of view of the camera. The camera was equipped with a motorized Navitar 12X zoom lens. Various optical elements (diffusers, lenses, collimators and multimode fibres) could be inserted into the laser beam between the laser and the jet, as indicated by the dashed box in Figure 3.2. The pulse energy, repetition rate and pulse width of the two fibre lasers were controlled via their computer interface from a laptop. Input from a function generator (Tektronix AFG3102) monitored by a fast oscilloscope (Teledyne LeCroy HDO6054) set

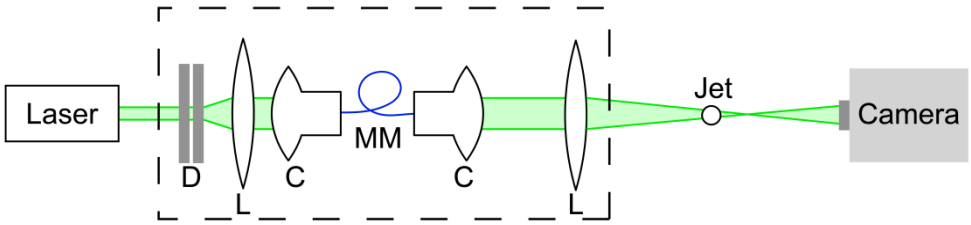


Figure 3.2 Experimental setup for testing jet imaging and velocimetry. Back-illumination of the jet requires the least laser power and so is the preferred mode of operation. The optical elements within the dashed box were used in varying combinations (or removed completely). D: Diffusers; L: Lens; C: Collimator; MM: Multimode fibre. Figure without changes from Grünbein *et al*[6].

the repetition rate of the diode laser, with pulse energy and duration set by its dedicated electronic circuit to the values listed in Table 1.

To scramble the angular and temporal coherence of the laser beam and thereby mitigate speckle if necessary, the following devices were used in varying combinations: ∅ 1" ground glass diffusers with 120, 220, 600 and 1500 grit polishes (Thorlabs), multimode fibres (∅ 300 µm core, 49 m length, Thorlabs FT300UMT) with fibre collimators (F220SMA-532 and F810SMA-543, Thorlabs) for coupling in and out of the fibre, and standard lenses (∅ 1" N-BK7, Thorlabs) to expand/reduce beam size, to collimate the beam after the diffusers and to introduce intentional beam divergences.

The degree of speckle formation in the image was characterized by the speckle contrast C defined as the intensity standard deviation σ relative to the average intensity value I_{avg} of the image:

$$C = \sigma / I_{avg} \qquad (1)$$

Type	Fiber laser, VGEN-G-HE-20	AlGaInP Laser Diode, HL6545MG	Tailored pulse fiber laser, Pyroflex-25-GR
Manufacturer	Spectra-Physics, Newport Cooperation	Ushio Opto Semiconductors, Inc.	Eolite Lasers, Electro Scientific Industries, Inc.
Wavelength [nm]	532	652-664; 660 typ.	532.2
Line width [nm]	≤1	NA	<1.5
Pulse width [ns]	3 – 20	3[(c)]	1 – 600
Max. repetition rate [MHz]	1	50	0.5
Peak Power [kW]	22[(a)]	0.0003[(c)]	≤10
Max. Pulse Energy [µJ]	204[(b)]	NA	45[(e)]
Output Beam Dia. [mm]	2.0 ± 0.3	0.3 (FWHM)[(d)]	NA
Full Beam Divergence [mrad]	0.4	≤ 0.8[(d)]	<2.5
M^2	<1.1	NA	<1.3

Table 1. Specifications of pulsed light sources tested for imaging fast GDVN jets and measuring their speed. Values taken from manufacturer manual, datasheet and test results. Table from Grünbein *et al*[6].
[(a)] Max. peak power with 7 ns pulse width and 150 kHz pulse rate
[(b)] Max. pulse energy with 10 ns pulse width and 100 kHz pulse rate
[(c)] Measured value
[(d)] Measured value after collimation with an aspherical lens as used for the described experiments
[(e)] With 10 ns pulse width and 200 kHz pulse rate

3.4.2 Velocimetry of fast microscopic jets

GDVN jets were imaged as described in section 3.4.1. Using a fast camera (Photron Fastcam SA-Z type 2100) capable of running at several hundred kilohertz repetition rate, single exposure images of the jet were recorded. Alternatively, double-exposure single frame images were recorded which were obtained by illuminating the jet twice within the exposure

time of a single frame of the camera. For speed measurement, the distance travelled by droplets beyond the Rayleigh break-up point of the jet were tracked over time. For this purpose custom-written Python scripts (Python 3) were employed as described in chapter 8.2. As a comparison to the automatic analysis of these scripts, speed was measured manually by measuring the position of the emitted droplets manually over time.

4 XFEL data collection parameters and setup

4.1 Bacteriorhodopsin

Descriptions from this section are in part published in reference[7].

The experiment was performed at the Linac Coherent Light Source (LCLS) of the Stanford Accelerator Center (SLAC) in July 2017 (proposal LP55). Data was collected at the Coherent X-ray Imaging (CXI) instrument using a photon energy of 9.8 keV and ~30 fs long X-ray pulses of ~0.2 mJ at the sample position. X-rays were focused to a spot of ~1.3 μm diameter in the microfocus chamber at CXI. Reference data of the light-adapted bR ground state were collected at 120 Hz X-ray repetition rate, while time-resolved pump-probe data were collected at 10 Hz X-ray repetition rate. Diffraction of the photoexcited bR crystals was recorded for time delays of 0.5, 1, 3 and 10 ps as well as 33 ms time delay between the optical pump laser and the X-ray probe pulse. For the shortest time delay, the true time-delay between pump and probe was measured for each pulse pair using the timing tool[78]. X-ray diffraction was recorded with a CSPAD detector[79] operated in dual-gain mode with a low and high gain in the low and high resolution region of the detector, respectively.

Femtosecond pump laser excitation of bR crystals

Circularly polarized laser pulses of ~5.9 μJ at 532 nm wavelength were produced by an optical parametric amplifier pumped by a Ti-Sapphire laser regenerative amplifier system and stretched to pulse durations of 145 ± 5 fs. The laser was focused to a spot of 99 μm $1/e^2$ diameter having a Gaussian intensity profile, which was positioned 25 ± 5 μm offcentre of the X-ray focus for all ps and sub-ps time points. For the 33 ms time delay, the laser was positioned slightly above the X-ray focus and data was collected at 30 Hz repetition rate.

Sample injection

bR crystals embedded in a mixture of LCP and Pluronic F-127 as described in section 3.1.4 were injected into the X-ray interaction region using a High Viscosity Extrusion (HVE) injector[55] equipped with a 100 µm inner diameter UV-vis transparent sample capillary at 1.9 µl min^{-1} sample flow rate.

Jet imaging

The viscous jet was imaged by the off-axis camera installed at the CXI instrument. Movies of the jet were recorded at 120 Hz repetition rate.

Retinal extraction and quantifying bR light adaptation

Light adaptation of bR was analysed by determining the ratio of $bR_{all\text{-}trans}$ to $bR_{13\text{-}cis}$. For this purpose, retinal was extracted chemically from the microcrystals and separated via high performance chromatography as described in reference[7].

4.2 MHz data collection at European XFEL

Descriptions in this section are partly published in references[1,5,71].

Two experiments were performed at the European XFEL (EuXFEL): the first one in June 2018 (proposal 2038) and a follow-up experiment in March 2019 (proposal 2156). Both experiments were performed at the Single Particles, Clusters and Biomolecules and Serial Femtosecond Crystallography (SPB/SFX) instrument[80] of the EuXFEL and X-ray pulses were delivered in 10 pulse trains per second, with pulses within each train being delivered at a maximum repetition rate of 1.1 MHz. The pulse length was ~50 fs (FWHM) based on electron beam diagnostics. X-ray diffraction was recorded using a 1 megapixel Adaptive Gain Integrating Pixel Detector (AGIPD)[81]. For each individual X-ray pulse, the pulse energy was recorded by two X-ray gas monitors (XGM)[82] upstream of the experimental hutch: one in the tunnel after the SASE1 undulator and one at the end of the tunnel upstream of the SPB/SFX instrument. For data presented in this thesis, the XGM closer to the SPB/SFX instrument was used.

June 2018

Photon energies of 7.47 keV and 9.22 keV were used for collection of two control lysozyme data sets and a photon energy of 7.48 keV was used for collection of diffraction data on jack bean protein crystals. The X-ray focal size was ~15 μm diameter at 7.47 keV and 7.48 keV and ~28 μm diameter at 9.22 keV, and each pulse had ~0.9-1.5 mJ pulse energy. Each X-ray pulse train consisted of 50 pulses separated by 889 ns (1.125 MHz). The maximum number of 60 pulses per train the machine was capable of at the time of the experiment was used. The first 10 pulses were used for electron orbit feedback and then sent to the pre-undulator dump, without lasing. This was done to preclude the possibility of the feedback loop potentially lowering the pulse intensity.

March 2019

A photon energy of 9.3 keV was employed, and X-rays were focused to a spot of ~3 μm diameter. At the time of the experiment, a maximum number of 176 X-ray pulses at 1.1 MHz repetition rate could be delivered within each train, and a new bunch patterning capability[83] was exploited to subdivide each X-ray pulse train into smaller subtrains (called "wagons") consisting of 5-6 pulses each.

Producing GDVNs capable of generating high-speed jets

GDVN nozzles were produced as described[60]. To achieve high-speed jets, the geometry of the gas sheath is very important. High-speed jets can be obtained by flame polishing the gas sheath such as to have a small orifice, and increase the distance of the sample capillary tip relative to the orifice[84]. In this configuration, the focusing gas exerts larger forces onto the jet, leading to larger accelerations and thus higher jet speeds. For GDVNs with sample capillaries of 75 μm (100 μm) inner diameter, the distance of the sample capillary to the orifice was in the range of ~150 – 200 μm (200 – 300 μm), and gas apertures were smaller than ~100 μm (120 μm).

Sample injection

A suspension of microcrystals in their mother liquor was injected into the X-ray interaction region via a liquid microjet produced by a gas dynamic virtual nozzle (GDVN)[60] with 75

and 100 µm inner diameter using helium as the focusing gas. The sample flow rate was 30-40 µl min^{-1}, and gas pressure 400-500 psi at the inlet of the GDVN's gas supply line, corresponding to a flow rate of 140-250 ml min^{-1}. All samples were 20 µm filtered prior to injection, and the suspension was adjusted to contain 10-15 % (v/v) settled crystalline material. During injection the sample was kept in a rotating temperature-controlled reservoir (20 °C for lysozyme microcrystals, 4 °C for jack bean protein microcrystals) to prevent crystal settling[85].

Jet imaging

The liquid jet was imaged from an off-axis perspective (orthogonal to both X-rays and jet flow direction) using a 10× infinity-corrected objective in combination with a 200 mm tube lens and a camera (Basler pilot pIA2400-17gm, Basler AG, Germany). The optical resolution of the imaging system, determined with a resolution target (Edmund Optics), was 1.6 µm. During data collection the camera pixels were 2×2-binned, resulting in recorded images with a scale of 0.68 µm $pixel^{-1}$. For illumination the fs SASE1 optical pump laser[86] was employed for jet illumination as described in[8]. The fs laser pulse and the camera were triggered by the EuXFEL global trigger (10 Hz) that indicates the arrival of an X-ray pulse train, thus the images were recorded at a set delay relative to the arrival of the pulse train and one image per pulse train was recorded. Jet imaging was primarily set up by Claudiu Stan.

Extracting a proxy for the solvent scattering intensity in MHz repetition rate experiments

The solvent scattering intensity on the detector indicates whether a given X-ray pulse interacted with the sample jet. For fast and efficient analysis of the scattered intensity of the solvent ring, data from only one detector module (module 02) that overlapped well with the signal of the solvent ring was used (Figure 4.1). To quantify the solvent scattering intensity, the standard deviation σ of all raw pixel intensities was used. It was found that σ was a more robust measure than other variables such as the median pixel intensity. The distribution of pixel intensities in the module becomes bimodal when solvent scattering is present. Thus, the higher the solvent scattering intensity the larger the standard deviation

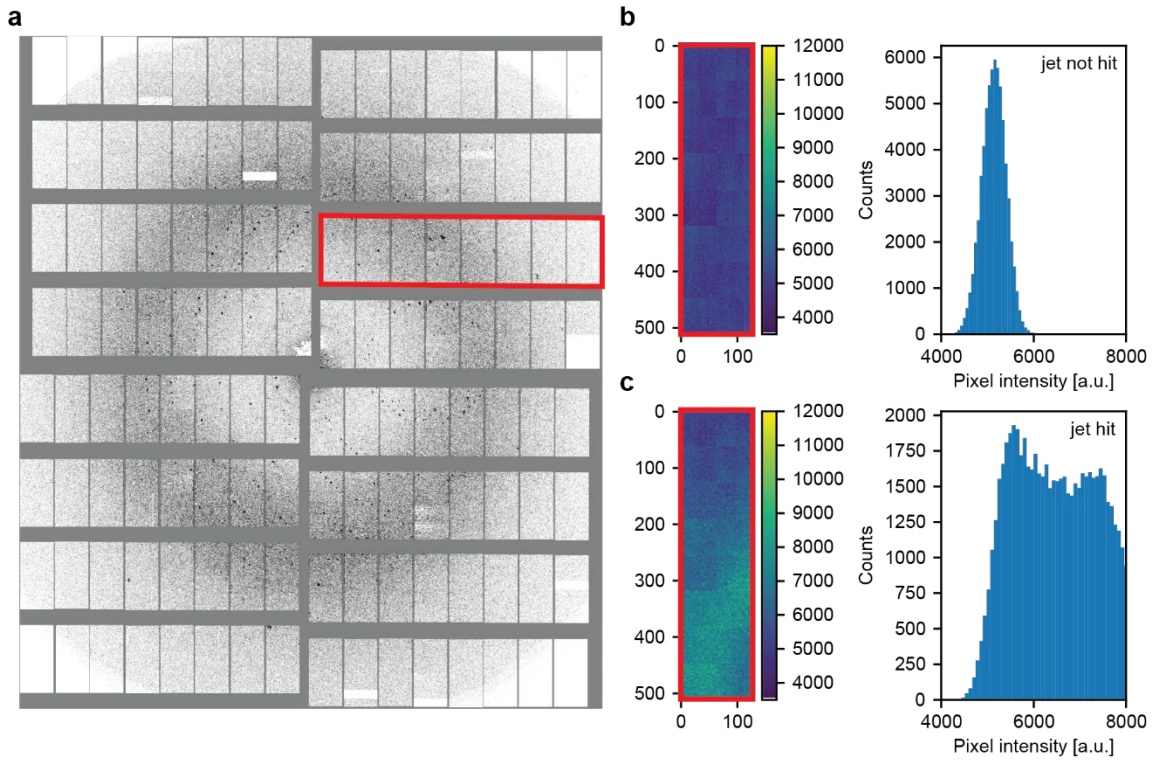


Figure 4.1 Quantifying solvent scattering intensity detected on the AGIP detector (AGIPD) at European XFEL. **a**. Image of the AGIPD geometry. Module 02 (red rectangle) overlaps partly with the solvent scattering ring. Panel adapted from Grünbein *et al.*[1] under the CC-BY 4.0[3]. **b**. If the jet is not hit, no solvent scattering is detected on the detector. The histogram of pixel intensities of module 02 is approximately gaussian. **c**. If the jet is hit and solvent scattering detected, the histogram of pixel intensities of module 02 becomes bimodal.

in pixel intensities. σ was therefore used as a proxy for the measured intensity of the solvent scattering. It was found empirically that solvent scattering is visible on the detector for $\sigma \gtrsim 400$, which was subsequently used as a threshold to identify whether the jet was hit.

Extracting the number of previous shocks for a given pulse position in a MHz repetition rate pattern at EuXFEL

To identify the number of shocks N launched by previous pulses, the solvent scattering strength σ was stored together with the trainID t and cellID c of each shot that uniquely identify it and its relative position to other shots. For each shot i hitting the jet (i.e. $\sigma_i \geq 400$)

with trainID t_i and cellID c_i, the standard deviation of pixel intensities σ_{i-1} of the previous shot (trainID t_i, cellID c_{i-1}) was interrogated. If the jet was hit ($\sigma_{i-1} \geq 400$), the number of pre-shocks N_i for shot i was increased by 1 and shot i-2 was analysed. If this again hit the jet, N_i was again increased by 1, and shot i-3 was analysed. This iterative procedure was terminated either when the beginning of an X-ray wagon was reached or at the first occurrence of not hitting the jet.

4.3 X-ray pump X-ray probe experiments at LCLS

Descriptions in these sections are in part published in references[2,9].

Two X-ray pump X-ray probe experiments (proposal LM18, July 2016 and proposal LR76, February 2018) were performed at LCLS to investigate XFEL-induced shock wave effects on protein crystals. Both experiment were performed in the microfocus chamber of the CXI instrument. Two 30 fs X-ray pulses of were produced using the two-bunch mode[87]. The photon energies of the two pulses were separated by ~70 eV, centred on the iron K-edge at 7.11 keV. The temporal separation of the two pulses was either 8.4 ns (July 2016) or 122.5 ns (February 2018) with negligible temporal jitter on the fs time scale. Additionally, the two pulses were vertically offset such that the interaction regions of the two pulses with the sample jet was separated by 5 µm. Tuning of the laser intensity on the gun cathode allowed some control over the relative pulse energies of the two pulses. Due to the unusual operation mode with a transverse offset, source size and divergence may have been larger than in standard single-pulse experiments.

As reference to the X-ray pump X-ray probe data, single-pulse diffraction data was collected in both experiments by suppressing the first of the two pulses. In this single-pulse mode beam feedbacks were turned off to prevent automatic adjustment of the single-pulse photon energy to the mean photon energy of the two pulses and to prevent shifting of the spatial location of the single pulse. The machine was run only for short time periods (<5 min) in this mode such that drifting of the energy was small. Because of the missing wakefields from the first bunch, the second bunch has a slightly different orbit, leading to a lower lasing intensity in the single pulse (probe-only) runs.

X-ray diffraction was recorded with a CSPAD detector[79].

Measuring pulse energies

In both experiments, X-ray sensitive diodes were installed to measure X-ray scattering from materials placed downstream of the interaction chamber. This was setup by Claudiu Stan. Diode signals were recorded synchronously with the CSPAD data. Pulse energies were evaluated as the background-corrected integrated signal of each pulse. For LM18, a diode (AXUVHS11, Opto Diode Corp) measured X-rays backscattered off a silica lens placed in the optical path at the end of the experimental arrangement. For LR76, two X-ray sensitive diodes (G4176-03, Hamamatsu Photonics K. K.) measured X-rays scattered off a Kapton foil placed in the optical path at the end of the experimental arrangement. One of the diodes was covered with a 25 μm Fe foil like the CSPAD detector. The non-masked diode measured the relative pulse energy of both pump and probe pulse, while the masked diode only measured the fraction of both pulses having photon energies below the Fe K absorption edge. The masked diode was used to evaluated probe pulse energies, while the unmasked diode was used to evaluate pump pulse energies.

Sample injection

A suspension of lysozyme or haemoglobin microcrystals in their mother liquor was injected into the X-ray interaction region via a liquid microjet produced by a GDVN[60] with 75 and 100 μm inner diameter using helium as the focusing gas. The sample flow rate was 30-50 μl min^{-1}, and gas pressure 300–600 psi at the inlet of the 2 m long GDVN gas supply line (with an inner diameter of 100 μm), producing ~50 m s^{-1} jets with 4-5 μm diameter. All samples were filtered prior to injection through a 20 μm filter, and the suspension was adjusted to contain 10–20 % (v/v) settled crystalline material. During injection the sample was kept in a rotating temperature-controlled reservoir at 20 °C to prevent crystal settling[85].

Jet imaging

The jet was imaged from the off-axis perspective (orthogonal to both X-rays and jet flow direction) a 50× infinity corrected objective (Plan Apo SL, Mitutoyo) in combination with a 200 mm tube lens and two different cameras camera (Opal 1000, Adimec and Vison Research, Miro R341) that could be used interchangeably for different purposes during the

experiment. The Opal camera was used to image the overlap between jet and X-rays, and for each pump-probe-pulse pair a jet image was recorded synchronously with the corresponding diffraction pattern. A fs laser was used for illumination as described in reference[8]. All of this was set up by Claudiu Stan. The fs laser pulse and the camera were triggered by electronic signals synchronized with the X-ray pulses, such that the images were recorded at a set delay relative to the arrival of the X-ray pulses. In the 8.4 ns experiment the jet was imaged at the arrival time of the probe. In the 122.5 ns experiment the jet was imaged a few nanoseconds after the probe pulse, to observe the formation of two distinct gaps in the jet impact of both pulses on the jet.

5 Characterizing intensity losses for photo excitation in jets

Descriptions of the methods in this section have been published in reference[4]. Sections have also been published in reference[11,88].

5.1 Characterizing transmission losses within jet media not containing crystals

The extinction A (relating to the transmission T via $T = 10^{-A}$) of various media typically used for viscous jet injection was recorded as described in section 3.2. Each medium was referenced to a non-absorbing medium of the same refractive index n to compensate for transmission losses due to reflection at the interfaces (Table 2).

Medium	Refractive index *n*	Reference medium
Synthetic grease Super Lube	1.47	100 % glycerol
LCP	1.42	65 % aqueous glycerol (w/w)
LCP + Pluronic F-127	1.43	65 % aqueous glycerol (w/w)
NaCMC	1.33	Ultra-pure water
Agarose gel	1.38	40 % aqueous glycerol (w/w)

Table 2. Refractive indices of common media for viscous jet injection in SFX experiments and suitable reference media of appropriate refractive index. Refractive indices were measured in the visible range (440 – 645 nm) in which dispersion was found to be small ($\Delta n < 0.02$). LCP: Lipidic Cubic Phase. LCP + Pluronic F-127: LCP and Pluronic F-127 (40 % (w/v)) mixed 4+1. NaCMC: Sodium carboxymethylcellulose (4% w/v in water). Agarose gel: 5.6 % (w/v) in 30 % aqueous glycerol. The preparation of all media is described in section 3.1. Table from Grünbein *et al*[4].

The measurements show that for all media dispersion is small ($\Delta n < 0.02$), demonstrating that referencing to a single reference medium is valid over the whole wavelength range. The measured extinction of the reference (where $A>0$ due to reflection at the interfaces to the cuvette and the medium) was subtracted from the measured extinction of the medium (where $A>0$ due to reflection at the interfaces, scattering and absorption within the medium) to obtain the extinction of the medium solely due to absorption by and scattering within the medium.

5.2 Orientation-averaged reflectance of a protein crystal

Reflectance R at the crystal surface was calculated based on the Fresnel equations describing the transmitted and reflected fraction when light travels from a region of refractive index $n=n_1$ to a region where $n = n_2 \neq n_1$. R generally depends on the two refractive indices n_1, n_2, the angle of incidence θ onto the interface and the polarisation of light. The reflectance at normal incidence for a given n_1, n_2 and polarisation of light was obtained by applying the Fresnel equations once for $\theta = 0°$. To calculate the orientation-averaged (mean or median) reflectance R, the reflectance $R(\theta)$ for a given pair of n_1 and n_2 was calculated for a set of 1000 angles within $0° \leq \theta < 90°$. R was obtained by averaging (applying the mean or median) on all $R(\theta)$ for each pair of n_1 and n_2. Separate calculations were performed for orthogonally (s) and parallel (p) polarised light; the behaviour of natural (unpolarised) light was obtained by averaging the results for s and p polarisation. The described operations were performed in Python 3.7.

5.3 Raytracing calculations on light propagation through large diameter jet

Light propagation through viscous jets was approximated using ray tracing techniques based on geometrical optics since the diameter $2R$ of viscous jets is much larger than the employed pump wavelength λ. Reflectance and refraction at interfaces between media of different refractive indices was calculated using the vector form of the laws of reflection and refraction. Time-resolved optical pump X-ray probe SFX experiments are usually performed with a focused Gaussian pump laser beam. In the focal region the wavefront can be approximated to be planar. For the ray tracing simulations, the incident beam is therefore

approximated by a parallel bundle of rays being incident on the jet with the optical axis perpendicular to the jet axis. The FEL probe pulse is typically a few μm in diameter, thus only a small slice of the incident pump beam is relevant for exciting sample inside the region where the FEL probes. This makes it sufficient to calculate the propagation for a fan of rays only inside this region instead for the whole beam. Geometrical ray tracing was performed for each ray in an iterative manner, calculating its propagation through free-space as a linear equation until intersection with an interface, where the laws of reflection and refraction were applied. Changes in local intensity I can be estimated by the change of the cross-sectional area of the ray tube formed by closely lying rays assuming a flat intensity profile of the incident beam[89]. Since light is only focused in the plane perpendicular to the jet axis the change in intensity scales linearly with the distance d to the neighbouring rays in the xy-plane: $I \sim d^{-1}$. Intensity changes due to reflection at the jet interface were not taken into account. The mean increase in intensity was calculated as the average increase in intensity along the optical axis of the X-ray beam.

PART II: RESULTS

6 Time-resolved SFX to observe ultrafast dynamics in bacteriorhodopsin

Parts of this chapter have been published in reference[7,11,88].

6.1 Introduction: bR, a light-driven proton pump

Bacteriorhodopsin (bR) is a light-driven proton pump found in Archea which exploit the proton gradient generated by bR for ATP synthesis[24]. bR is the best characterized member of the family of rhodopsin membrane proteins which contain retinal as a light-absorbing chromophore. Rhodopsins are involved in many different light-activated processes such as vision or light-energy conversion[24]. Despite this wide range in functionality, the general structure of the members of the family is conserved, as is, one believes, the underlying mechanism of how the excited retinal interacts with the surrounding protein matrix[24].

At the core of all light-driven rhodopsin dynamics is the trans↔cis isomerization of the retinal chromophore upon light absorption[24] (Figure 6.1). In bR the all-*trans* to13-*cis* isomerization takes place within 3 ps after photon absorption when the K intermediate is formed[7,12] (Figure 6.1). The rearrangement of the surrounding protein matrix then leads to the unique functionality, in case of bR to unidirectional proton pumping[24]. This ultrafast rearrangement was hitherto only accessible via ultrafast spectroscopy. While these experiments were successful in determining time constants and energetic states, a 3D picture of the structural rearrangement, the interaction of the chromophore with the surrounding protein matrix guiding the reaction towards its functional path, was missing. It thus remained unknown how photo-excitation of the chromophore induces functional motions of the protein. For this task optical pump X-ray probe experiments at XFELs are the method of choice: a short light pulse emitted by an optical laser initiates the reaction inside a protein microcrystal, the diffraction of which is interrogated after a set time delay by an X-ray pulse. This enables to follow bR's structural rearrangement from the ultrafast timescale after photon absorption onwards.

At the time when the experiment presented in this thesis was conducted, no ultrafast time-resolved SFX experiment on membrane proteins had been performed. Time-resolved ex-

periments in the ultrafast regime had been performed before on soluble proteins[16,20,21] using liquid jets produced by GDVNs for sample delivery, ensuring rapid sample exchange and thus fresh sample for each pump-probe pair. The membrane protein bR, however, is crystallised in lipidic cubic phase (LCP), a viscous matrix[68,90]. Sample therefore needs to be delivered into the interaction region by means of viscous streams. This imposes a number of additional technical challenges, often including low stability of the jet. Moreover, jet velocity is low and can show large variations over time[68]. As a result, excited sample may not have cleared the X-ray interaction zone in time before arrival of the next X-ray pulse such that the starting conditions of the subsequent pump-probe pair are ill-defined. Matching and controlling sample delivery speed, X-ray and excitation repetition rate thus become central issues to the experiment, as the interplay of those three parameters determines the state captured in the pump-probe experiment (see section 6.4)[11].

A further complication of the experiment is that the retinal in bR occurs in two isomeric forms, of which only one can undergo a functional photocycle[91]. These two forms coexist

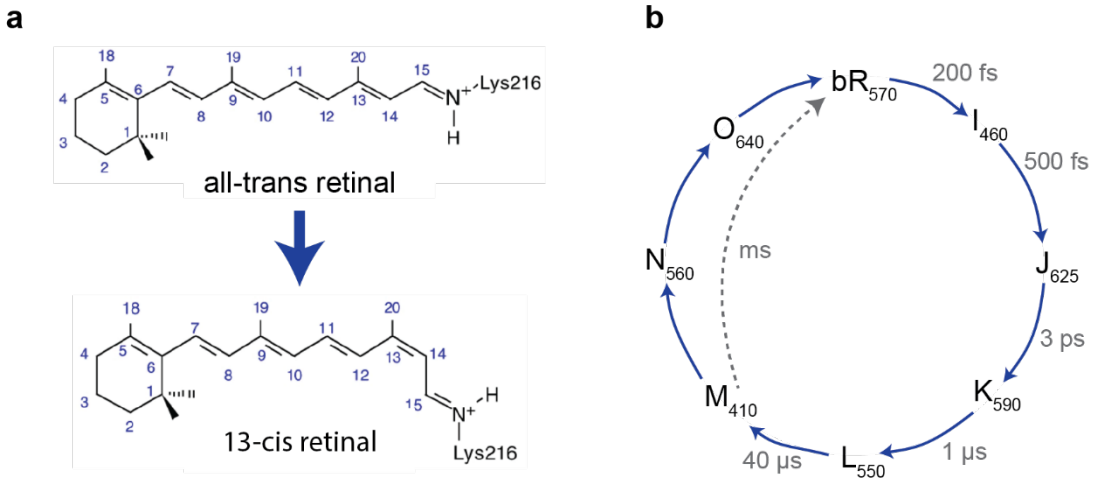


Figure 6.1. Light-induced dynamics in retinal-containing bacteriorhodopsin. **a**. Chemical structure of the light-absorbing chromophore, the retinal, in the active groundstate (all-trans) and the isomerized 13-cis retinal after light absorption. Figure adapted from Nass Kovacs *et al.*[7] under the CC-BY 4.0[3]. **b**. Photocylce of bR showing the different intermediate states after light absorption of the ground state (bR_{570}) and the time scale on which these intermediates may be observed. The superscript denotes the wavelength of the absorption maximum of all states. Absorption maxima and time constants taken from[12,13].

in thermal equilibrium in a ratio depending on the light conditions. Therefore, to increase photoexcitation yield of the desired process, illumination conditions have to be found that increase the fraction of the functional isomeric form.

To enable time-resolved experiments on bR crystals, techniques had to be developed to ensure that the sample delivery matches the pump-probe repetition rate and to ensure that the fraction of bR in the configuration capable of entering the functional photocyle is maximised.

6.2 Experimental design

To capture the ultrafast light-triggered dynamics in bR including isomerization of the bound retinal, a time-resolved SFX experiment on bR microcrystals was conducted, using a femtosecond optical pump laser for excitation and the XFEL pulse to probe the protein structure at given time delays after reaction initiation. To deliver a fresh bR crystal for each pump-probe pulse pair the crystals embedded in LCP were injected as a viscous stream into the interaction region using a high viscosity extrusion (HVE) injector[55]. Viscous jet flow is achieved by pushing the sample at constant flow rate out of a reservoir through a coned, 100 μm inner diameter (ID) capillary using a hydraulically driven piston and employing a co-flowing helium gas stream to improve homogeneous extrusion as a free jet as shown in Figure 6.3[11].

To trigger the reaction, a 532 nm circularly polarized laser pulse of 145 fs duration was employed. The experimental arrangement of jet, X-ray and pump laser optical axes is depicted in Figure 2.1a. The axis of the optical pump laser was aligned to be nearly parallel to the X-ray beam axis to allow for accurate timing also for sub-picosecond time delays which cannot be achieved in case of perpendicular alignment.

To minimize background scattering impairing the signal to noise ratio of the recorded diffraction signal of the crystals, the experiment was conducted under vacuum conditions. Thus, all additions to the experimental setup that are described in the following sections needed to be compatible with vacuum operation conditions and moreover needed to fit the space constraints within the vacuum chamber.

6.3 Preillumination

In the dark-adapted resting state bR's retinal adopts one of two configurations, 13-*cis*,15-*syn* and all-*trans* ($bR_{13\text{-}cis}$ or $bR_{all\text{-}trans}$), with the equilibrium depending on the light condition[91,92]. In a fully light adapted ensemble of non-crystallised bR, 100 % of the chromophores are found in the all-*trans* configuration, while dark-adapted bR consists of similar parts of $bR_{all\text{-}trans}$ and $bR_{13\text{-}cis}$, respectively[91]. Importantly, only $bR_{all\text{-}trans}$ can enter the functional photocycle including the functional retinal isomerization and subsequent proton transfer[91]. Photon absorption by the non-functional $bR_{13\text{-}cis}$ leads to a state with millisecond life time, that thermally decays to either the 13-*cis* or, to a lesser degree, the all-*trans* ground state[91]. Thus, for maximising the population of $bR_{all\text{-}trans}$ that may undergo the desired photoreaction, light adaptation by continuous illumination needs to be performed. By repeatedly driving the $bR_{13\text{-}cis}$ photocycle the amount of $bR_{all\text{-}trans}$ can be enriched because the accumulated $bR_{all\text{-}trans}$ state is stable for several minutes. The half-life of light-adaptation in bR microcrystals was measured to be 17 min using absorption spectroscopy (data not shown).

Light adaptation of bR is typically achieved by continuous illumination with visible light[91-93]. The measured half-life of light adaptation is very short compared to the time required for different steps of the SFX pump-probe experiment: after loading the HVE injector with LCP-containing microcrystals, mounting the injector inside the experimental vacuum chamber takes typically 15 - 30 min after which data collection of the loaded sample takes 2 – 3 h. Therefore, light adaptation has to be performed *in situ* during the experiment by means of "online" preillumination, and not prior to injector loading as done by other groups for simplicity[94,95]. However, online preillumination is not straight forward and has to fulfil the following constraints: (i) it must allow continuous light-adaptation of sample prior to reaching the X-ray interaction point to maximise the fraction of light-adapted molecules while (ii) ensuring at the same time that all molecules have returned to the ground state when reaching the interaction point; (iii) it has to be compatible with vacuum conditions and (iv) fit the spatial limitations inside the chamber

Since bR crystals are optically very dense (1/e penetration depth of 3.5 µm), only thin sample layers of sufficiently low crystal concentration can be efficiently light-adapted, otherwise the crystals shield each other. It was thus decided to preilluminate through the sample capillary, instead of preilluminating the much thicker sample reservoir inside the HVE

injector. Since the standard sample capillary of an HVE injector has a brown polymer coating to increase mechanical stability, a UV-vis transparent capillary was used instead, allowing sample preillumination through the side wall. To maximise transmission towards the bR sample, the orientation of the square gas sheath surrounding the sample capillary has to be aligned to be perpendicular to the incident light source (see Figure 2.1 for the geometry of HVE nozzles). Tests were performed to establish light adaptation conditions that are compatible with the experiment, testing different light sources, the required preillumination time and intensity were tested.

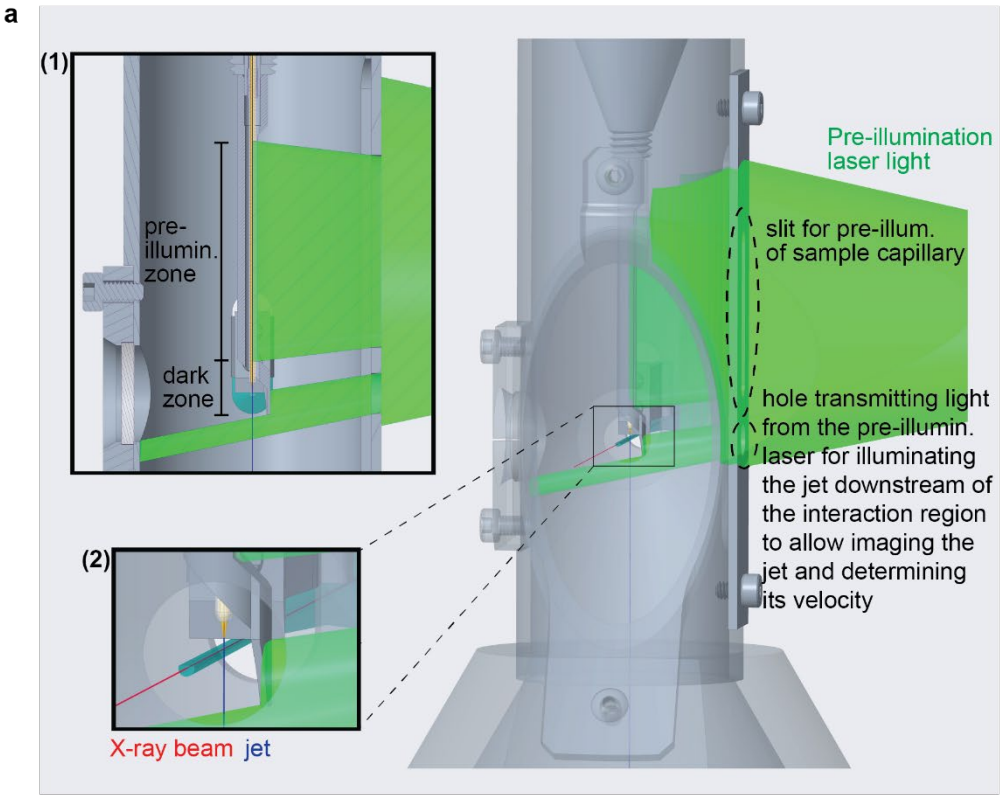
The light adaptation efficiency of a white light LED (Schott KL2500), a Xenon lamp, a green LED (the green channel of a SOLA 3-channel light engine, Lumencor), a 532 nm CW laser (RGBLaser 200 mW Fiber Coupled Raman Laser) and a 519 nm CW laser (Oxxius LBX-525-800-HPE-PP) were assessed. The non-coherent sources (the white light and the green LED source) were tested first as these would be easier to implement into the experimental arrangement at the beamline given the strict safety controls concerning application and implementation of lasers. Due to the finite size of the experimental vacuum chamber and a lot of equipment already being in place, additional light sources need to be either vacuum compatible and small enough to fit inside the chamber, or be able to use the free viewing port at ~50 cm distance to the interaction point to guide light towards the interaction region. At this distance, intensity has decreased substantially even when focusing and/or collimating the tested incoherent sources, leading to insufficient light-adaptation of bR. It was thus tried to couple the light output into a large-diameter (1.5 mm core) multimode fibre to guide the light inside the chamber an allow placing the fibre's end closer to the interaction region. However, due to the low transfer efficiency this arrangement was inadequate for significant light adaptation. In contrast, light emitted by the coherent light sources can be well collimated an could thus be efficiently transferred through the optical fibre. Initial tests showed that both lasers, i.e. both light at 532 nm as well as light at 519 nm, can be used for light adaptation as both increased the amount of $bR_{\text{all-}trans}$. Given the higher output power leading to more efficient light adaptation in bR, the Oxxius 519 nm CW laser was the preferred choice. The free viewing port is located on the side of the vacuum chamber that is regularly opened up during the experiment for checking and changing equipment inside. It was therefore decided to use multimode fibres to couple light into the chamber from a (non-viewing) side port, thus avoiding having to realign the optics each time the chamber is opened.

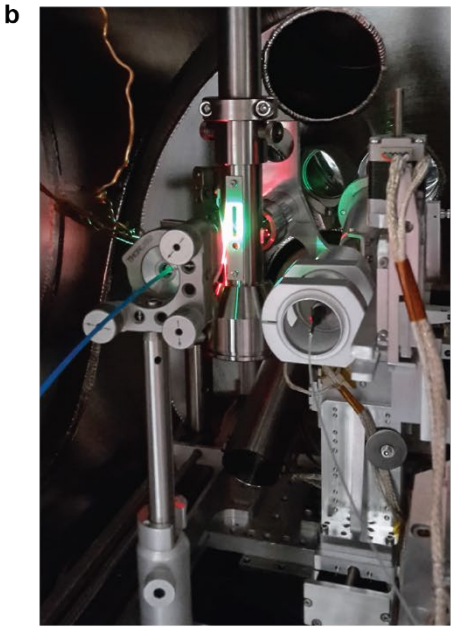
The laser's output was therefore coupled into a multimode fibre of 1.5 mm core diameter. A custom-made flange was prepared to serve as an O-ring sealed feedthrough into the vacuum chamber, such that a continuous light fibre (FT1500UMT, TECS clad, 10 m long) could be used for transport into the vacuum chamber, increasing transmission to the interaction point compared to the standard setup using separate fibres outside and inside the vacuum chamber. Compression of the O-ring onto the fibre for sealing purposes did not lead to significant changes in the transmitted intensity. Both of the fibre's ends were polished after passing through the feedthrough. One end was equipped with an SMA connector fitting the laser head, the other polished end was used as the light source for preillumination of bR inside the vacuum chamber.

The efficiency of light adaptation was measured by the increase in the fraction of $bR_{all\text{-}\textit{trans}}$ upon extracting the retinal chemically and separating it via high performance liquid chromatography (HPLC)[7,92,96]. It was found to be optimal for few-second, medium-intense illumination: for ~5 s long illumination light adaptation increased with light intensity up to ~200 mW output of the multimode fibre contained in a 4 cm spot. At higher intensities, heating damaged the protein, evidenced by a change in colour and a change in the consistency of the extruded sample, preventing smooth injection.

To allow ~5 s long preillumination inside the sample capillary, the UV-vis capillary was extended from the usual 5 to 26 mm emerging from the HVE injector into free space. Likewise, also the length of the gas sheath was increased such that the co-flowing gas stream stabilizing injection could be guided to the point of extrusion. In order for all bR molecules to be back in the ground state before pumping and probing the system, sample should not be illuminated a short (~1 s) time prior to reaching the interaction point. This was achieved by sliding a metal mask onto the tip of the HVE injector: the custom-made metal mask contained a 22 mm long window through which light could illuminate the capillary. The last 4-5 mm of the mask were solid, shielding sample from further illumination to allow ground state recovery (Figure 6.2). The inside shape of the mask was concave, reflecting any transmitted light back into the sample for increasing illumination from all directions. Due to the high viscosity of the embedding matrix, the velocity profile inside the sample capillary is constant (plug flow). Given the inner diameter of 100 µm of the sample capillary, sample flow rate was set to 1.9 µl such that sample translates at 4 mm/s, spending 5.5 s in the preilluminated and 1 s in the masked region, respectively.

At the beam line, preillumination efficiency was tested *in situ* by injecting preilluminated sample into a light-tight dark Eppendorf tube, the content of which was subjected to HPLC analysis. The all-*trans* retinal content was 65-80 %, improving the originally dark-adapted state in bR microcrystals containing only 40-45 % $bR_{all\text{-}trans}$ by a factor of two. Moreover, successful measurement of a high-quality ground state structure of bR obtained without femtosecond pump laser excitation showed that the mask blocked the preillumination light efficiently and that the extent of the masked region provided sufficient time for all molecules to return to the ground state. If either of these two constraints would not have been met, the measured ground state structure would have resembled the structurally very distinct M state which accumulates under continuous illumination and decays with a time constant of approximately 100 ms in crystals[94] (Figure 6.1). Indeed, the measured dark state structure did not show any such light-induced changes demonstrating the successful preillumination and retinal in its all-*trans* configuration.


a
(1)
pre-
illumin.
zone
dark
zone
(2)
X-ray beam
jet
Pre-illumination
laser light
slit for pre-illum.
of sample capillary
hole transmitting light
from the pre-illumin.
laser for illuminating
the jet downstream of
the interaction region
to allow imaging the
jet and determining
its velocity


b

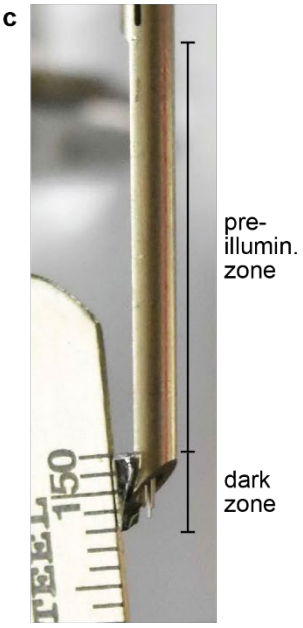

c
pre-
illumin.
zone
dark
zone

Figure 6.2 Preillumination geometry. **a**. Schematic drawing of the preillumination geometry inside the SFX chamber. The sample capillary extending from the injector body (not shown) is located in a differential pumping shroud to shield the vacuum chamber from the gas load of gas-focused injection[60]. A slit in the shroud's side wall allows preillumination of sample flowing through the capillary. A custom-made metal mask allows illumination along the upper 22 mm of the sample capillary, but blocks light along the 4-5 mm above the interaction point, providing a dark zone in which all molecules have time to return to the ground state (inset 1). Its diagonal finish cut allows the XFEL probe pulse to interrogate the jet at a position shielded from the preillumination laser (inset 2). A second hole in the shroud allows illumination of the extruded jet right below the interaction point, allowing jet imaging and speed measurements using a camera on the opposite side of the experimental chamber. Drawing and shroud design by Bruce Doak, figure adapted from Nass Kovacs *et al*.[7] under the CC-BY 4.0[3]. **b**. Picture of the experimental arrangement inside the vacuum chamber, showing the mounting of the multimode fibre guiding light towards the sample capillary behind the slit inside the shroud. **c**. Close up view of the metal mask on the sample capillary.

6.4 Matching jet speed, pump laser spot size and pump-probe repetition rate

When performing pump probe measurements on a moving target such as a free-flowing jet, jet speed, the spatial extent of the pump excitation region and the repetition rate of the experiment have to be matched. After initiating a reaction in a given segment of the jet, the very same segment must be located in the X-ray interaction region when the X-ray pulse arrives, otherwise the ground state is measured instead of the evolving reaction (Figure 6.3). Moreover, the excited sample segment must have left the interaction region before the next pump-probe event such that each pump-probe event starts with the same, clean

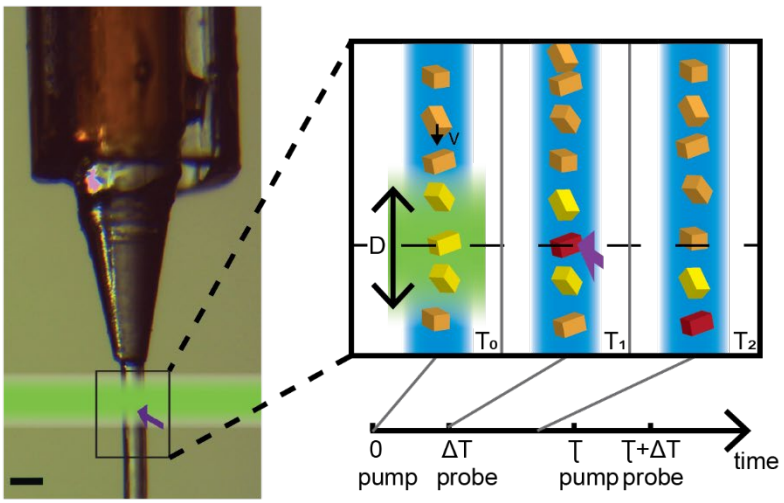


Figure 6.3 Pump-probe experiments on samples embedded in a moving jet. Left: HVE jet of vaseline. The purple arrow indicates the X-ray interaction. For pump-probe experiments, a section of the sample jet is optically triggered (green shading) before X-ray probing. The black scale bar is 100 µm. Right: Constraints on time-delays for time-resolved experiments valid for any injection system. The reaction is triggered in crystals (orange) within the segment hit by the pump pulse (optical axis indicated in green) at time T_0, after which the reaction proceeds in the excited crystals (yellow). After a time delay ΔT the X-ray pulse (purple arrow) probes one of the excited crystals (red) at time T_1. Jet speed v must be sufficiently slow so that not all crystals excited within the region upstream of the X-ray optical axis have passed the interaction region (dashed line), i.e. $v<D/\Delta T$. All crystals triggered at T_0 must clear the interaction region before arrival of the subsequent probe pulse at $T=\Delta T+\tau$, $1/\tau$ being the X-ray repetition rate, requiring jet speed to be $v>D/(\tau+\Delta T)$. Figure adapted from Grünbein & Nass Kovacs[11] under the CC-BY 4.0[3].

conditions in the ground state (Figure 6.3). This constrains jet velocity to a defined regime depending on the pump-probe time delay and the pump laser beam size. i) If the jet is too fast, the excited segment has already travelled beyond the interaction point before the X-ray probe pulse arrives and ii) if the jet is too slow, the excited segment has not yet cleared the X-ray interaction region before the subsequent pump-probe exposure. Moreover, if the point of reaction initiation is fully upstream of the X-ray interaction, the excited segment has not yet arrived in the X-ray interaction region if the jet is too slow.

For successful pump probe measurements, jet speed must therefore be known and matched to the repetition rate of the experiment and the extent of the sample segment illuminated by the pump laser (Figure 6.3). This is particularly important for pump-probe measurements on viscous jets, where jet speed is often not a constant value, but may vary significantly due to sample inhomogeneity affecting jet flow. Both unavoidable inhomogeneity after mixing different substances for the final matrix and the crystals themselves disturb jet flow, leading to changes in jet shape due to changing surface tension and changing propensity to adhere to the sample capillary tip. Moreover, the pressure required to transport sample through the sample capillary depends on sample viscosity as well as crystal size and density. Changes in these parameters that naturally occur in any sample of microcrystals embedded in a viscous matrix lead to variations in sample flow rate and thus variations in jet speed due to the lag time required for the hydraulic system driving sample flow to re-adjust the driving pressure. It is thus not sufficient to assume a constant jet velocity derived from the inner diameter of the sample capillary as done by other groups[95].

In the described experiment, jet speed and the extent of the pump laser excitation zone were constrained by the preillumination requirements and properties of the optical system at the beam line, respectively. The pump laser beam had a 99 µm $1/e^2$ diameter at the interaction point and its centre was aligned to 25 ± 5 µm below the X-ray probe point. This reduced the jet extent above the interaction point in which sample is excited and which needs to be displaced for clearing the interaction region before the next pump-probe pair. Based on the Gaussian intensity distribution of the beam profile, the pump laser intensity decreased to 0.1 % of the peak intensity at 91 µm radial distance from the focus. Given the vertical offset between pump and probe laser centre, the jet thus has to translate by at least 66 µm between shots to clear the excited region in time for the next pump-probe pair (assuming that illumination by less than 0.1 % of the peak intensity is insufficient for reaction

triggering). While faster jet speeds are favourable to ensure such clearing, efficient preillumination puts an upper bound of 4 mm/s on jet speed. At higher speeds light adaptation of the microcrystals in the described preillumination setup would have been less efficient and bR molecules in the X-ray interaction region may not yet have returned to the ground state after light adaptation (see section 6.3).

Given the natural variation in jet speed, jet speed was measured during the experiment and the repetition rate of pump and probe were adjusted to allow clearance of the excited sample segment. During the experiment jet velocity and its distribution were determined both periodically and after each change in flow conditions such as changing sample batch. Movies of the flowing jet were recorded using an off-axis camera positioned outside the experimental chamber on the opposite side of the preillumination light source that was also used for jet illumination (Figure 6.2). The pixel size was calibrated by imaging a capillary of 360 µm outer diameter, resulting in 0.54 µm/pix and thus a field of view of 553 x 553 μm^2 (1024 x 1024 pix^2). To allow tracking features over multiple frames at this field of view, 2 - 4 min long movies were recorded at the maximum repetition rate of the camera of 120 Hz. For a given movie, jet speed was measured in regularly spaced time points (every 250-1000 frames depending on the length of the recorded movie) for 11-40 time points per movie. Histograms of jet speed were continuously evaluated (Figure 6.4), monitoring average jet velocity and its distribution. To allow clearing the excited sample segment prior to each pump probe shot, pump probe data of the ultrafast time points was collected at 10 Hz such that for > 90 % of pump-probe pairs all sample having been exposed to at least 0.1 % of the peak pump intensity had cleared the interaction region.

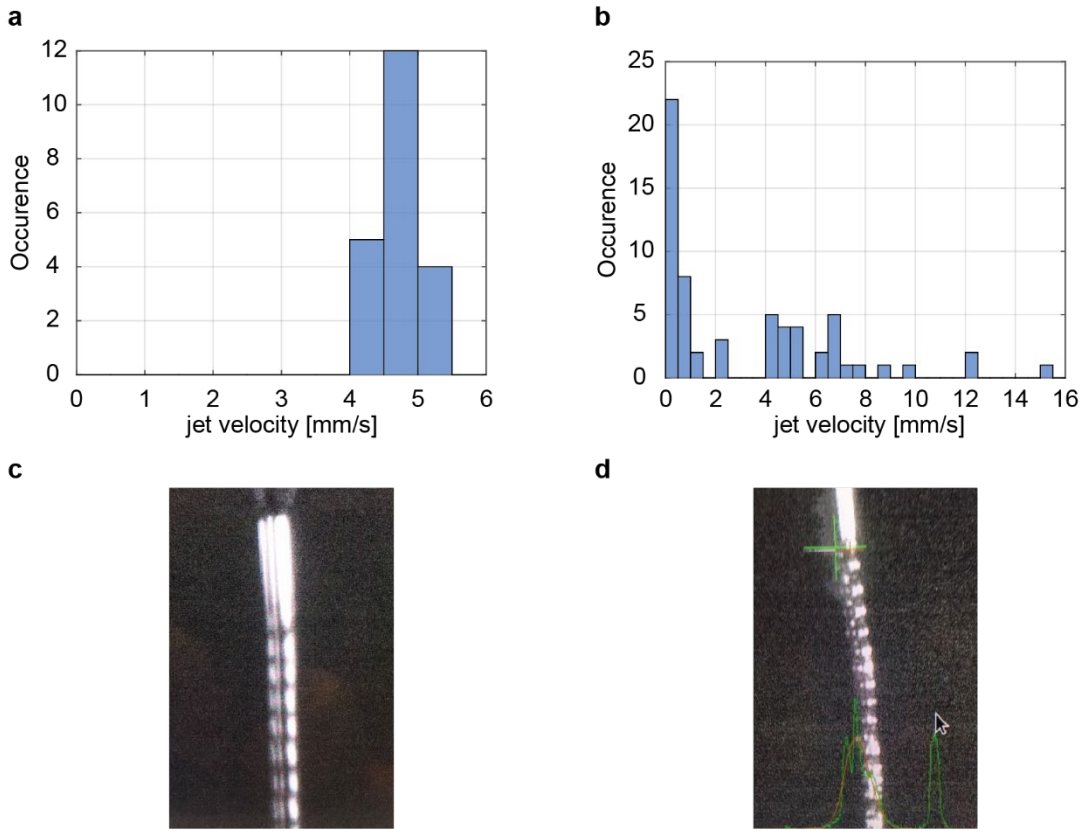


Figure 6.4 Monitoring jet velocity during the experiment. **a**,**b**. Histogram of the jet velocity measured at regularly spaced time points, showing the number of occurrences a given jet velocity range was measured. Perturbations in the LCP matrix, e.g. as represented by the microcrystals, change the flow behaviour of the extruded stream and lead to variations in jet velocity. These were monitored by periodic speed measurements of the jet by tracking features over multiple frames at evenly sampled time intervals. The figure shows histograms of jet velocity measured over a time period of ~4 min. **a**. Jet speed histogram of a well-behaved sample flowing with the desired constant velocity. **b**. Jet speed histogram of a sample that cannot be injected well. The jet velocity varies significantly, nearly coming to a halt at times and then running very fast at others. Data collected with low jet speed are problematic, as the probed jet segment may have been illuminated twice by the pump laser such that the time-delay probed by the X-ray pulse is ill-defined. Any data collected with sample jets showing such behaviour was therefore discarded. **c**,**d**. Pictures of jets in which the impact of the X-ray beam on the jet is visible as dark stripes. The spacing between two segments affected by X-ray pulses depends on jet speed, leading to "ladders" of larger or smaller spacing for faster (**c**) or slower (**d**) jets. Jet imaging can thus already be used as a quick online monitor to detect problematic behaviour. Pictures displayed here were obtained from a different beam time at SwissFEL, but are conceptually similar to those taken during the bR beam time.

6.5 Ultrafast structural changes in bR upon photoexcitation

bR diffraction data was collected at the CXI instrument at the LCLS. In addition to the light-adapted unpumped ground state, time-resolved data was recorded for time delays of 0.5, 1, 3 and 10 ps as well as 33 ms between the optical pump laser and the X-ray probe pulse. For the shortest time delay, the exact time-delay between pump and probe was measured for each pulse pair using a timing tool[78]. Based on the measured time delay, the sub-ps data was sorted into 12 distinct time points ranging from 240 fs to 740 fs. Due to the high quality of the diffraction data with bR microcrystals diffracting to 1.8 Å resolution, small structural changes occurring on the sub-ps time scale could be resolved (Figure 6.5). Starting from the very first time point at 240 fs post triggering twisting of the retinal starts to become apparent (Figure 6.5). Moreover, changes start occurring in the surrounding protein structure and water network at sub-ps time scale and an oscillatory behaviour in the retinal and its surrounding residues and waters suggest a strong vibrational coupling between retinal and its environment[7].

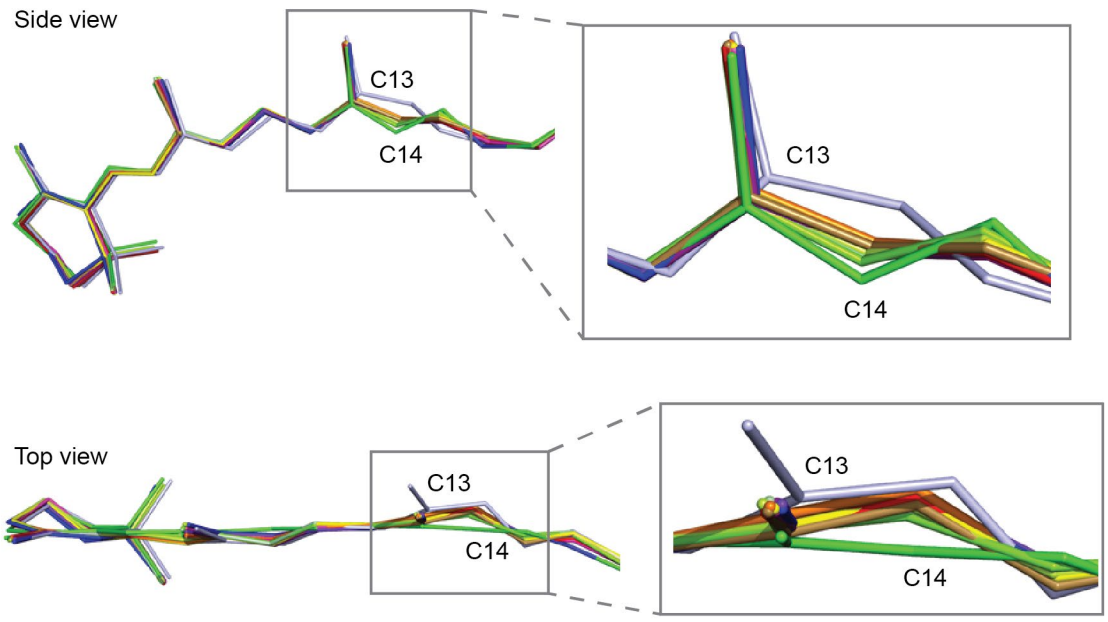


Figure 6.5 Isomerization of the retinal C13-C14 bond on the fs time scale. Green: ground state (all-*trans*), light purple: after 10 ps (13-*cis*), rainbow: sub-ps time delays ranging from 240 fs to 740 fs. The detailed structural analysis was performed by Thomas Barends and Gabriela Nass Kovacs and is published in reference[7].

7 Controlled photo-excitation in light-sensitive protein crystals

The time-resolved experiment yielded detailed insight into the ultrafast reaction of bR, resolving retinal isomerization and a vibrational coupling between retinal and the surrounding protein matrix[7]. However, following a reviewer's comment during the revision process of the manuscript describing these experiments and their findings, further analysis of the photoexcitation conditions was performed, including a spectroscopic power titration.

Upon closer investigation it became clear that the excitation conditions used for the time-resolved SFX experiment were far beyond the single photon regime: When focusing 5.9 μJ into a 99 μm $1/e^2$ diameter spot of Gaussian intensity distribution, the power density for the 145 fs long pulse at the X-ray probe position (offset by 25 ± 5 μm) is ~630 GW/cm^2 assuming a flat temporal intensity profile. From spectroscopy it is known that the ultrafast response of bR changes at high excitation power[97-100], with the onset of nonlinear behaviour - indicating multiphoton events - starting at ~30 GW/cm^2 (see reference[7]). While photoproduct formation is observed in the determined structures of bR, it does not mean that the structural changes and dynamics leading there resemble the biologically relevant reaction[22].

Extensive parts of this chapter have been published in reference[4]. Sections have also been published in references[7,11,88].

7.1 Problematic photo-excitation common practice in previous experiments

The in-depth analysis of the time-resolved SFX experiment on bR initiated also a re-analysis of the excitation conditions of previously published time-resolved SFX experiments using light excitation for reaction triggering. To resemble naturally relevant processes, pump-probe experiments on light-triggered biological systems need to be carried out in the linear regime in which the observed signal scales linearly with excitation intensity, translating to one or less absorbed photons per absorbing molecule. All ultrafast pump-

probe experiments[16,20,21,95] have used power densities of 360 – 500 GW/cm² (Table 3), corresponding to nominally tens[20,95] to hundreds[101] of photons absorbed per molecule on average*. Such multiphoton excitation of the chromophore is a radical departure from the biologically relevant single-photon processes studied by spectroscopy employing only tens of GW/cm² (see references[22,99]).

A common justification for the high pump laser intensity is that a large fraction of the incident pump beam scatters off the jet employed to deliver the crystals and is therefore lost before reaching the interaction region[95,101]. However, the magnitude of these claimed losses was not characterised.

Experiment	Time delay	Power density	Fluence
CO myoglobin[16]	500 fs – 150 ps	380 GW/cm²	566 µJ/mm²
Photoactive yellow protein (PYP)[21]	100 fs – 3 ps	320 GW/cm²	450 µJ/mm²
rsEGFP[20]	1 ps – 3 ps	360 GW/cm²	830 µJ/mm²
Bacteriorhodopsin[95] (*)	~100 fs – 10 ps	520 GW/cm²	520 µJ/mm²
Bacterial phytochrome[102]	1 ps – 10 ps	6300 GW/cm²	4400 µJ/mm²
Krokinobacterei-kastus rhodopsin 2 (KR2)[103]	800 fs – 20 ms	800 GW/cm²	1200 µJ/mm²

Table 3 Photo excitation conditions in previous ultrafast tr SFX experiments.
(*) First time points included data in a range of time delays from 0 fs to 141 fs.

* assuming equal intensity reduction and penetration depths for single and multiphoton absorption

Therefore, to allow mechanistically meaningful data collection in the biologically relevant regime, the actual excitation intensity reaching the sample needs to be quantified. To determine the number of photons available for triggering a reaction several changes to the incident intensity have to be taken into account (Figure 7.1): (i) Reflection of the incident beam at the interfaces surrounding-to-jet and (ii) jet-to-crystal, as well as (iii) potential absorption by the jet medium decreases the incident intensity. Moreover, (iv) refraction at those interfaces changes the intensity distribution. Last but not least, (v) within the protein crystal absorption by the constituting molecules reduces intensity for deeper layers, potentially leading to different excitation regimes within the crystal. The analysis of these contributions is described separately in the following sections.

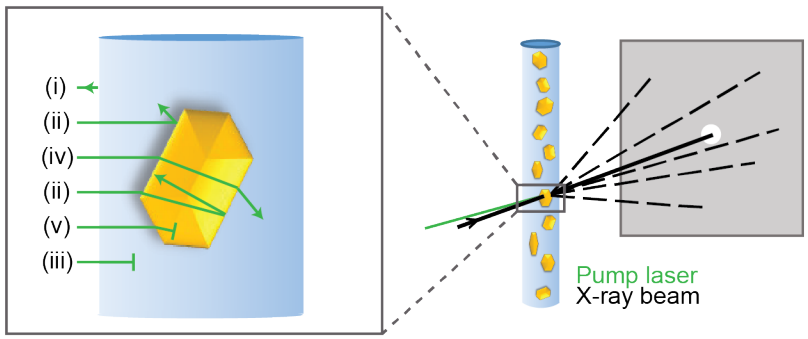


Figure 7.1 Intensity change upon passage through the liquid jet to the protein crystal. Reflection of the incident beam at the interfaces surrounding-to-jet (i) and jet-to-crystal (ii) as well as potential absorption by the jet medium (iii) decreases the incident intensity. Moreover, refraction at those interfaces changes the intensity distribution (iv). Last but not least, within the protein crystal absorption by the constituting molecules reduces intensity for deeper layers (v). All these processes change the intensity I reaching a chromophore inside the jet at depth x. Figure based on a figure published in Grünbein *et al*[4].

7.2 Excitation regimes within protein crystals

The intensity I_0 of a pump laser beam travelling through a crystal containing a concentration c of absorbing chromophores decreases with propagation length d. In general, assessing the exact change in intensity is highly complex, since all possible molecular states and transitions would need to be taken into account, requiring knowledge of the corresponding linear and non-linear absorption cross sections. Since these are unique to each system, generally unknown and difficult to obtain[104], an easy-to-apply approach to judging the absorption regime in pump-probe experiments on protein crystals is required. As a strong oversimplification[105] but nonetheless straight-forward model for analysing excitation conditions, the Beer-Lambert law can be used†.

To assess whether excitation conditions target a biologically relevant or multophoton regime, it is generally not sufficient to calculate the average number of photons absorbed per molecule. Molecules at different depths in the crystal are subjected to different intensities (Figure 7.2) and since the X-ray diffraction signal averages over all molecules in the X-beam, the effects of different excitation regimes cannot be separated. Moreover, structural changes due to higher order processes may dominate over those due to the single photon absorption process.

To obtain information on the different photon absorption regimes at different depths within a crystal, Beer-Lambert's law has to be applied differentially. At depth x in a sample with particle density ρ_{mol} the number of photons absorbed per molecule $N_{abs}(x)$ given an incident number of photons per area N_{phot} becomes[4]

$$N_{abs}(x) = \frac{c \cdot \epsilon}{\rho_{mol}} \cdot \ln(10) \cdot N_{phot} \cdot 10^{-\epsilon \cdot c \cdot x} \qquad (2)$$

Using equation (2), pump-probe experiments maximising the fraction of molecules in a desired regime may be designed. However, for a clean experiment on light-sensitive proteins properly delineating the biological response, the possibility of higher-order processes

† For a crystal containing a concentration c of the absorbing molecule of molar absorption coefficient ϵ, the intensity at penetration depth d is $I = I_0 \cdot 10^{-\epsilon \cdot c \cdot d}$.
It assumes that the chromophore concentration c remains essentially unchanged during absorption and that the extinction coefficient ϵ is independent of laser power. As such it is not a universally valid physical law (e.g., it breaks down when interfaces reflect a fraction of the intensity backwards, when fluorescence or stimulated emission occur or when the molar extinction coefficient depends on the concentration of the molecule), see reference[105]. However, assuming equal intensity reduction and penetration depths for single and multiphoton absorption as a strong oversimplification, Beer-Lambert's law can be used to calculate the number of photons absorbed per molecule N_{avg} in a straight-forward manner.

should be avoided at all cost by setting the incident intensity such that the linear regime applies to all molecules in the crystal.

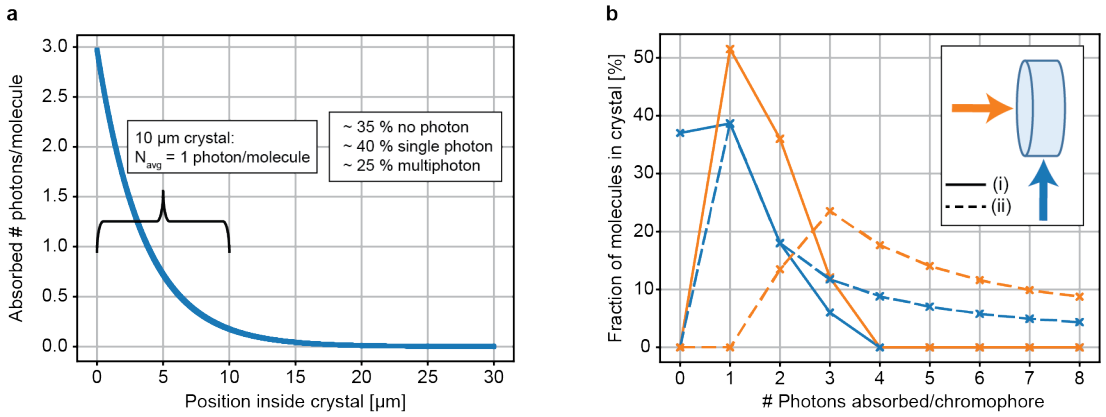


Figure 7.2 Depth-dependent photon absorption regimes inside bR crystals. The average number of absorbed photons per molecule was calculated at varying depth *x* inside the crystal using equation (2). $n > 1$ absorbed photons per molecule in the above plots imply increasing probabilities of higher-order processes, and not necessarily a specific n^{th} order process. **a**. Depth-dependent photon absorption regimes in bR crystals (c = 27 mM, ϵ_{532nm} = 45600 $M^{-1}\cdot cm^{-1}$) with incident pump laser illumination (pulse energy E = 0.25 µJ, $1/e^2$ diameter 100 µm, wavelength 532 nm, fluence F = 64 $J\cdot m^{-2}$) resulting in $N_{abs,avg}$ = 1 for 10 µm thick crystals. **b**. Effect of crystal orientation on the relative contribution of different photon absorption regimes inside bR crystals for two photoexictation scenarios (wavelength 532 nm, $1/e^2$ diameter 100 µm in both cases). All molecules absorbing on average n-0.5 ≤ n < n+0.5 are counted towards the regime absorbing n photons. Blue (orange) lines describe the situation for crystals of 10 µm (5 µm) thickness. (**i**) Incident laser pulses (E = 0.25 µJ, F = 64 $J\cdot m^{-2}$) result in $N_{abs,avg}$ = 1 for 10 µm thick crystals. For 10 µm (5 µm) thick crystals, 35 % (0 %) of molecules do not absorb photons, 40 % (50 %) absorb 1 photon and 25 % (50 %) are in higher-order regimes. (**ii**) Incident pump laser pulse (E = 0.715 µJ, F = 182 $J\cdot m^{-2}$) set such that N_{abs}≥1 everywhere in a 10 µm crystal. The situation shifts significantly towards higher-order processes. For 10 µm (5 µm) thick crystals, 0 % (0 %) of molecules do not absorb photons, 40 % (0 %) absorb 1 photon and 60 % (100 %) are in higher-order absorption regimes. **a**, **b** with formatting changes taken from Grünbein *et al*[4].

7.3 Intensity loss along the beam propagation path

The intensity of the incident pump pulse available for triggering a reaction is reduced by reflection on the jet and the crystal surface, as well as by potential absorption by and scattering within the jet medium (Figure 7.1).

Intensity loss due to absorption by jet media

While traversing the viscous medium constituting the jet, a fraction of the incident light is absorbed by the medium itself. If these losses are large, the number of photons impinging onto the protein crystal at a given depth in the carrier medium can be significantly reduced. To quantify the fraction of photons available for starting a reaction within a jet of a given thickness, the transmission of light through the jet must be characterized. Since absorption scales with path length, absorption by the jet medium is of relevance for large (50 – 100 µm) diameter viscous jets, but not for the micron-sized GDVN jets, which have diameters of typically only 1 – 10 µm. Refraction on the cylindrical jet surface leads to a large divergence of a beam after passage through the jet (see section 7.4), making transmission measurements through the jet itself highly errorprone. Therefore, absorption by jet media was measured in spectroscopic cells of appropriate thickness (see chapter 3.2.2).

Extinction after passage through a 100 µm thick layer was measured for typical high viscosity jet matrices: synthetic grade Super Lube grease[56], LCP[55,65,68], LCP with the additive Pluronic F-127[68], agarose in a water-glycerol mixture[72] and sodium carboxymethylcellulose (CMC)[68]. In summary, typical homogeneous jet carrier, including all tested water-based viscous matrices (CMC, Agarose) as well as LCP show negligible absorption and scattering losses (Table 4). The only exception is Super Lube which contains PTFE micro powder that scatters light strongly.

Medium	λ = 450 nm		λ = 532 nm		λ = 670 nm	
	A	***T* [%]**	***A***	***T* [%]**	***A***	***T* [%]**
Synthetic grease Super Lube[#]	0.416±0.007	38	0.378±0.007	42	0.314±0.006	49
LCP	0.009±0.001	98	0.007±0.001	98	0.006±0.001	99
LCP + Pluronic F-127	0.006±0.002	99	0.005±0.002	99	0.004±0.001	99
NaCMC*[§]	0.000±0.001	100	0.000±0.001	100	0.000±0.001	100
Agarose*[¶]	0.000±0.002	100	0.000±0.002	100	0.000±0.002	100

Table 4. Transmission losses due to absorption by and scattering within jet media not containing crystals as measured by a downstream detector. *A*: measured fractional extinction after referencing to the reference medium. Corresponding transmission computed from the formula $T = 10^{-A}$.

[#] Super Lube contains PTFE micro powder that is visible under the microscope as densely packed particles in the medium. These scatter light strongly.

[*] absorbance was 0 within the experimental error of the measurement.

[§] 4 % NaCMC ultra in water

[¶] Results at 60 °C with warm, liquid, homogenous medium. Cold medium has the same absorbance if it is completely homogenous, i.e. if it is cooled down within the cuvette before the measurement. If cooled down before filling the cuvette, the medium becomes very brittle and filling the cuvette without breaking the gel into many pieces / introducing many air bubbles is impossible. In that case, the absorbance is increased by ~0.01, giving a transmission of T ~ 98 %. Table from Grünbein *et al*[4].

Changes in intensity distribution due to scattering centres in the jet medium

Excitation conditions need to be well-determined and thus reproducible. Irreproducible and numerous scattering centres in the jet, as for example intrinsically contained nanoparticles in the commonly employed Grease[56] matrix or inhomogeneities introduced by mixing (e.g., gas or liquid bubbles), change the incident intensity in an undefined manner by redirecting light from the original beam path. These changes cannot be quantified by transmission measurements as done previously[102] since a downstream detector generally does not capture all scattering angles, such that the intensity loss measured by the detector is

much larger and not representative of the intensity change within the medium itself. Intensity redistribution within the scattering medium depends strongly on the number density and scattering properties (shape, size) of the scattering particles. For example, particles much smaller than the light wavelength scatter light at all angles, while forward scattering dominates for larger particles[106]. Introduction of scattering centres in undefined quantities and size ranges therefore must be avoided.

Intensity loss due to reflection on the jet surface

The magnitude of reflection on the jet surface is described by the Fresnel equations and depends on the refractive indices of the jet medium and the surroundings (generally vacuum, helium or air with refractive index $n = 1$), the light polarisation and the angle of incidence. The refractive indices of typical jet media are within a range of $1.33 \leq n \leq 1.47$, important examples being water(-based substances) with $n = 1.33$, LCP with $n = 1.42$ and synthetic grease Super Lube with $n = 1.47$ (Table 2).

Using the Fresnel equation on a set of parallel rays incident on the jet as described in section 5.3 shows that ~2-9 % of the incident intensity is reflected on the surface of the jet, independent of jet diameter (Table 5). The same relative magnitude (2-9 %) is internally reflected on the rear surface of the jet, leading to some intensity modulation inside the jet. Internally reflected light may moreover intersect crystals embedded in the jet with a temporal offset of ~100 fs relative to the primary pulse (see reference[4] for details). The effect of both this "after-pulse" re-excitation and the intensity modulation is however small due to the low intensity of the internally back reflected light compared to the initial intensity.

Intensity loss due to reflection on the crystal surface

The magnitude of reflection on the crystal surface is again described by the Fresnel equations and depends on the refractive indices of the crystal and the surrounding medium, the light polarisation and the angle of incidence. In SFX experiments, a large number of similar microcrystals is probed at random orientations. Therefore, the exact fraction of pump light reflected from the crystal surface varies from shot to shot depending on the orientation of each individual crystal. To estimate the average reflectance over all shots one can either

Refractive index *n*	**Mean reflectance *R* [%]**			***R* at 0° incidence [%]**
	s pol.	p pol.	Circ pol. / nat. light	
1.33	6.5	2.2	4.3	2.0
1.42	8.4	2.7	5.6	3.0
1.47	9.4	3.0	6.2	3.6

Table 5. Reflectance of an incident beam on the jet surface. The environment of the jet is assumed to have a refractive index of $n = 1.0$ (vacuum). The mean reflectance is obtained by averaging the reflectance of a set of equally spaced rays perpendicularly incident onto the jet. If the X-ray probe arrives perpendicular to the pump, it will intersect at near-orthogonal angle essentially all of the internal light rays of the pump beam (see Figure 7.4), making the mean reflectance a meaningful and useful characterization of the loss due to surface reflection. If, on the other hand, the probe beam arrives parallel to the pump axis, it will intersect at near-zero angle only a very limited subset of the internal light rays of the pump beam. For a small diameter probe beam passing directly through the centre of the jet, the polarization-independent reflectance of the central light ray having a 0º angle of incidence is then a more useful characterization.
n: refractive index of the jet. s (p) polarisation: Electric field vector of the incident electromagnetic wave is perpendicular (parallel) to the plane-of-incidence.
Table from Grünbein *et al*[4].

use the reflectance at normal incidence as a rough approximation, or calculate the average reflectance over all orientations.

Determining the refractive index of protein crystals is far from trivial given the sensitivity and generally small size of macromolecular crystals. For large lysozyme crystals the refractive index was determined to be $n \sim 1.55 \pm 0.03$ depending on crystal lattice, face and wavelength[107]. The refractive index for other protein crystals will likely lie in the same range given generally similar density and composition. Assuming an absolute refractive index change of 0.2 ± 0.1 between a water-based medium ($n = 1.33$) and the crystal, typically ≤ 10 % of the incoming intensity is reflected at the crystal surface when averaging over all possible orientations assuming equal probabilities for all angles of incidence $0° \leq \theta < 90°$

(Figure 7.3). This number presents an upper limit on the orientation-averaged reflection since the assumption of equal distribution of all angles of incidence overestimates the contribution of large angles of incidence. Moreover, reflectance is further reduced for common jet carrier media like LCP which have higher refractive indices than water.

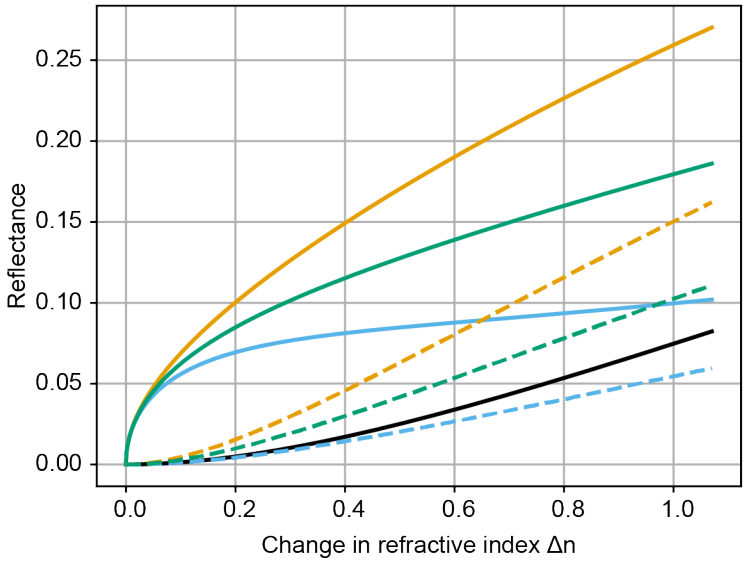


Figure 7.3 Reflectance at a protein crystal's surface. Reflectance on the crystal surface was obtained by calculating the orientation-averaged reflectance of an object of refractive index $n_{xtal} = 1.33 + \Delta n$ in water ($n = 1.33$) assuming a flat probability distribution for all angles of incidence $0° \leq \theta < 90°$. Δn corresponds to the difference in refractive index between water and the protein crystal. The plot shows the reflectance at normal incidence (black solid curve) as well as the mean (solid curves) and median (dashed curves) orientation-averaged reflectance for orthogonally (s, orange) and parallel (p, blue) polarised and natural (unpolarised, green) light. Figure with formatting changes from Grünbein *et al*[4].

7.4 Changes in intensity distribution due to refraction at the jet interface

Due to its cylindrical shape, the jet acts like a cylindrical lens with effective focal length EFL = $\frac{n\,R}{2(n-1)}$, where n is the refractive index of the jet and R its radius[108]. Any light beam impinging orthogonally onto the jet will therefore be focused into a light sheet.

Light propagation in large jets

For large (50 – 100 μm) diameter viscous jets, propagation of the incident beam through the jet can be well approximated using optical ray tracing (Figure 7.4). Due to the short focal length (e.g., for a 2R = 100 μm LCP jet with n=1.42 the effective focal length is 85 μm) the rays diverge strongly beyond this focal line (Figure 7.4a) and focusing by the jet interface significantly changes the light intensity within the jet.

For pump-probe experiments in large-diameter jets, two extreme scenarios are relevant, namely orthogonal pump illumination and probing of the jet with the X-ray beam either perpendicular (pumping along the x-axis in Figure 7.4a) or parallel to the pump beam (pumping along the y-axis in Figure 7.4a). In the perpendicular case the power density inside the jet is on average calculated to be ~1.2-1.3 times higher than that of the incident beam (Figure 7.4b). The power density in the jet centre is relatively uniform but there are regions of very low or even zero power density near the jet edges (Figure 7.4b). In the parallel case, the average power density within the jet is calculated to be ~1.4-1.6 times higher than in the incident beam (Figure 7.4c). There are no regions of zero power density, but the power density gradually increases along the probe axis by as much as a factor of $\alpha \geq 2$ (Figure 7.4c). For both calculations, no back reflections inside the jet were considered.

To achieve a relatively uniform excitation intensity inside the jet a perpendicular pump-probe geometry or two counter-propagating pump beams[109] should be utilized. However, a perpendicular pump-probe geometry is more dependent on having a very stable jet position since jet movements affect the distance that the X-ray and pump light travel differently, and thereby affect the pump-probe time delay. For example, a 50 μm displacement along the probe axis induces a ~165 fs time difference in vacuum. Likewise, the temporal resolution in jets with a diameter of tens of microns is smeared on the femtosecond timescale.

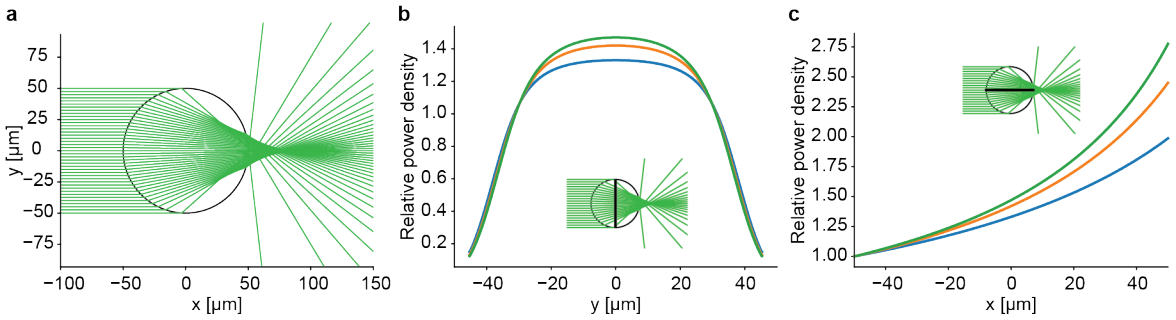


Figure 7.4 Propagation of light through a cylindrical jet of $2R$ = 100 µm diameter. **a**. Geometrical ray tracing of a parallel beam of light impinging onto a 100 µm LCP jet (refractive index n = 1.42) shows that the pump beam is focused by the sample delivery jet. **b**,**c**. Increase in power density inside a 100 µm jet of refractive index 1.33 (blue line), 1.42 (orange line) and 1.47 (green line) in case of perpendicular (**b**) and parallel (**c**) pump probe geometry. The incident beam has a uniform intensity profile, is collimated and intersects the jet in a plane orthogonal to the jet axis. The inset shows the axis along which the power density was evaluated: The black line shows the direction of the X-ray beam, the green fan corresponds to the geometrical ray tracing of the pump laser shown in **a**. Both reflectance at the jet interface and the increase in power density within the jet are independent of the jet diameter as long as the incident intensity profile is uniform and geometrical optics are a valid approximation ($2R \gg \lambda$). **b**. The average power density increase for sample located along the X-ray probe axis is 1.2±0.2 for n=1.33 (blue), 1.2±0.3 for n=1.42 (orange) and 1.3±0.3 for n=1.47 (green). **c**. The average power density increase for sample located along the X-ray probe axis is 1.4±0.3 for n=1.33 (blue), 1.5±0.4 for n=1.42 (orange) and 1.6±0.5 for n=1.47 (green). All figure panels adapted with formatting changes from Grünbein *et al*[4].

Light propagation in micron-sized GDVN jets

For micrometre-sized GDVN jets the assumptions of geometric optics do not apply because the wavelength approaches the size of the object. To complete the analysis planar electromagnetic wave propagation was therefore employed by our collaboration partner Sylvain Lecler to determine intensity changes in small diameter jets. Significant interference effects modulate light intensity inside the jet (Figure 7.5) with the power density inside the jet reaching almost 4 times the incident power density at certain “hot spot” locations. On average the power density increases to 1.1 ± 0.5 times the incident power density. ~4 % of the volume inside the jet is subject to power densities of at least twice the incident power density. Apart from such local maxima and minima, the intensity distribution otherwise agrees well in overall form with the predictions of geometrical optics (Figure 7.5).

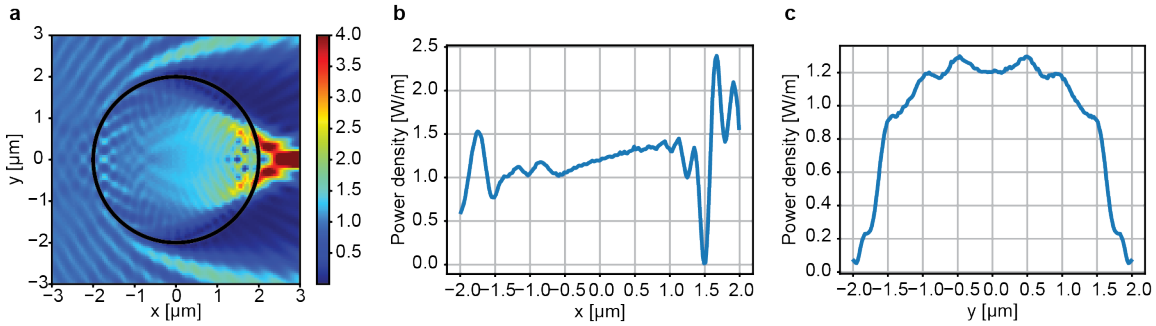


Figure 7.5 Propagation of an electromagnetic plane wave through a water jet. The following parameters were used: λ=532 nm, *n*=1.33, jet diameter 4 µm. **a**. Power density distribution. The maximum of the colour scale was set to 4 W/m, corresponding approximately to the maximum power density of 3.9 W/m within the liquid jet. **b**. Power density distribution along the *y*=0 line. **c**. Power density distribution along the *x*=0 line. **a**,**b**,**c**. Incident power density is 1 W/m. The calculations were performed in two dimensions, resulting in a power density expressed in W/m compared to W/m^2 in three dimensions. All figure panels with formatting changes from Grünbein *et al*[4].

7.5 Guidelines for biologically relevant photo-excitation conditions in SFX

Employing extraordinarily intense excitation conditions is a common approach used in all published ultra-fast time-resolved SFX experiments using femtosecond optical trigger pulses to start a reaction (Table 3). However, under such conditions it is questionable whether the biologically relevant processes are monitored, or whether structural rearrangements resemble those of biologically irrelevant multiphoton processes. Given the prevalence of such high excitation conditions, at times guided with wrong justifications why the biological importance of the experiment remains untouched[95,101], the detailed analysis of appropriate conditions described in the previous sections was performed following the time-resolved experiments on bR.

To set excitation conditions appropriately, it is mandatory to first know the excitation regime in which the desired biological reaction can be observed. This includes the spectroscopic determination of an appropriate excitation wavelength, pulse duration and peak power. Wavelength and pulse duration have to be chosen such as to both optimize efficiency in forming the excited states and photo products of interest (and concomitant low

excited state absorption) as well as to ensure a reasonable penetration depth in protein crystals and a temporal resolution sufficient to observe the desired process. By means of a power titration, the linear response regime can be determined.

Once knowledge on a suitable excitation regime is obtained, this needs to be translated to the experimental conditions of the SFX experiment, taking the intensity changes along the beam propagation path through jet and crystal into account. As described in detail in the previous sections (7.3-7.4), reflection on jet and crystal surface, as well as intensity changes due to refraction effects have to be taken into account. Thus, knowledge on the jet medium's refractive index *n*, pump laser polarization and pump laser alignment relative to jet and X-ray axis is required.

In the course of this thesis, a detailed protocol describing the practical steps required for careful preparation of a meaningful pump-probe SFX experiment was prepared based on the presented results. This protocol has been published in reference[4] and is repeated in the Appendix (chapter 11.1).

8 First MHz repetition rate experiments

Extensive parts of this chapter have been published in references[1,5,6,71]. Sections are also published in references[2,9,11].

8.1 Requirements for SFX at MHz repetition rate

Major drawbacks of SFX experiments at first-generation FELs are extraordinarily high sample consumption when using GDVN injection and limited availability of beam time at XFELs. XFELs capable of emitting X-ray pulses at MHz repetition rate have therefore been awaited eagerly. An increase in pulse rate is expected to decrease sample consumption and to speed up data collection, thereby accommodating more users and allowing the collection of enough data to study systems with very weak signals. However, increasing the peak repetition rate from first-generation XFELs[23] running at up to 120 Hz to up to 4.5 MHz at the first superconducting XFEL, the European XFEL in Hamburg[110], introduces new constraints.

The basal requirement for making use of the high repetition rate is that sample delivery needs to be sufficiently rapid to present fresh material for each pulse. The only method developed so far that is capable of doing so is liquid jet injection by GDVN jets. However, the impact of the XFEL pulse on the liquid jet brings with it new challenges that have to be overcome for successfully conducting experiments at MHz repetition rate since a segment of the sample-containing liquid jet is destroyed by the XFEL beam[8]. X-ray photons are absorbed in the liquid via the photoelectric effect and the generated photo and Auger electrons redistribute their kinetic energy via collisions with other electrons and surrounding molecules, thermalizing a small volume of liquid[8]. Within a few picoseconds, the intersected filament is isochorically heated to a temperature and pressure far beyond the boiling point[8]. The heated segment vaporizes, and the pressurised vapour expands radially outwards, pushing the jet ends apart into liquid films (Figure 8.1a)[8]. The explosion thus creates a growing gap in the jet. With the expansion of the vapour cloud, its pressure decreases, decreasing gap growth rate and enabling the gap to be flushed downstream such that a continuous jet is re-established in the interaction region after ~1 µs in case of typical GDVN jets of 3 – 7 µm diameter running at ~20 – 30 m/s[8]. Moreover, the high energy den-

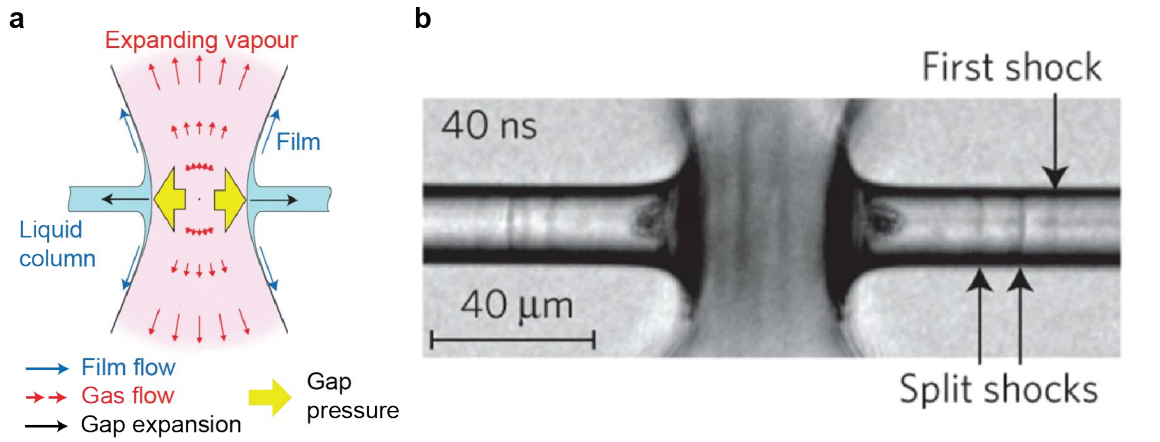


Figure 8.1 Effect of X-ray pulses on liquid jets. **a**. Physics of jet explosions after interaction of the XFEL pulse with the liquid jet. X-ray absorption leads to isochoric heating of the intersected jet fraction, leading to very high energy densities and pressures. The segment vaporizes, and the pressurized vapour expands, thereby applying pressure radially outwards, opening up a gap in the liquid jet[8]. **b**. The impact of the X-ray pulse launches a shock wave propagating along the liquid jet. During propagation shock waves split into multiple shock fronts[8]. **a**,**b**. Figures reprinted (adapted formatting in **a**) by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature Physics, Liquid explosions induced by X-ray laser pulses, Claudiu A. Stan *et al*., Copyright © 2016, Nature Publishing Group. This is reference[8] in the reference list.

sity after X-ray absorption results in the generation of a cylindrical shock wave propagating supersonically outwards along the jet (Figure 8.1) and producing ns-duration pressure jumps on the order of 0.1-1 GPa (1-10 kbar) within the jet medium[8].

Accordingly, there are two critical issues for data collection at MHz XFELs using liquid jets for sample delivery: (i) The gap precludes subsequent XFEL measurements until a continuous jet has been re-established at the interaction point, otherwise the X-ray beam does not intersect the jet. The minimum time required for native jet conditions to be regained is given by the time required for the gap to flush downstream, determining the minimum necessary time delay between two X-ray pulses indeed probing the sample jet. (ii) The shock wave propagates upstream along the jet and can damage the sample species carried by the jet. To truly exploit the increased efficiency of high-speed data collection, the experiment shall ideally be conducted at the maximum X-ray repetition rate that still probes unperturbed sample condition.

The first issue can then be addressed by using sufficiently fast jets, thereby increasing the speed with which the gap is flushed downstream, decreasing the time within which sample is replaced in the interaction region. Re-establishment of the jet and transport of damaged jet material are primarily a function of the jet speed, the recovery time decreasing with increasing jet speed[8]. However, whether the feasible jet speeds (a few tens of m/s) suffice for MHz data collection was an open question prior to first experiments.

To ensure efficient use of the scarcely available beam time, jet conditions have to be tested beforehand to ensure knowing if and how a specific sample can be injected at high speed with a given type of GDVN nozzle. For this purpose, efficient techniques and methods have to be developed to allow imaging and speed determination of high-speed, micron-sized liquid jets. This is also highly relevant for time-resolved SFX experiments to ensure probing the prepared portion of the sample jet (see section 6.4).

The issue of the potential effect of shockwaves is far less trivial as it is not a local effect but one capable of affecting samples far away from the actual exposure site. It is therefore critical to analyse that shock waves do not affect the sample under investigation at the short spacing between two X-ray pulses at MHz repetition rate (~1 µs and less).

8.2 Jet imaging and velocimetry

Imaging fast, microscopic jets requires an optical system fulfilling constraints both in terms of optical quality, enabling to visualize droplets and jets of few micrometre diameter, and in terms of temporal resolution, enabling to sharply visualize objects moving at ~100 m/s. To "freeze" the motion of a droplet, motion blurring must be reduced to a small fraction of the droplet diameter. This can be achieved with nanosecond exposure time: A 6 µm diameter droplet traveling at 100 m/s requires an exposure time of 6 ns to reduce motion blurring to 10 % of its diameter.

Sharp images of microscopic jets have been obtained using femtosecond laser pulses for illuminating the jet[8]. While producing very clean images, laser speckle, a common issue when employing coherent illumination, had to be reduced via multiple decohering stages

to obtain images of sufficient optical quality. Extending this result, the aim of the work presented here was to find simpler solutions allowing jet imaging and speed measurements that can be easily set up both in the home laboratory for beam time preparation and at the experimental end station of an XFEL without requiring extensive alignment and expensive equipment.

Pulsed illumination allowing jet imaging and velocimetry

To obtain sharp images of fast jets, few nanosecond exposure from a light source of sufficient intensity is required. To allow speed determination, the repetition rate of such exposures must be higher than 200 kHz* for at least two subsequent pulses in a row. This is easiest achieved using pulsed laser illumination. Without complicating the optical setup, it however needs to be ensured that speckle patterns created from the interference of coherent wave fronts do not compromise imaging.

Techniques to decrease speckle contrast were therefore investigated to allow improving optical quality if a laser source exhibits pronounced speckle which prohibits jet imaging. Generally, speckle contrast can be decreased by the incoherent addition of many independent speckle patterns. Similar to other results[8,111,112] it was found that only a combination of techniques is capable of efficiently improving optical quality: speckle contrast could be reduced from 68 % to 9 % (Figure 8.2a vs. b) by employing a stack of ground glass diffusers of different grain size, thereby increasing angular diversity of the beam, in combination with a 49 m long multimode fibre, thereby elongating the pulse by intermodal dispersion inside the fibre to a duration longer than its coherence time[112]. As such elements unnecessarily complicate the setup, pulsed light sources like laser diodes or fibre lasers with inherently large wavelength and or angular diversity are preferred.

Three different light sources (Table 6) capable of achieving the required repetition rate and pulse durations were tested for jet illumination (Figure 3.2). Two of them, a fibre laser and a small laser diode, allowed minimizing optical complexity since their direct output allowed essentially speckle-free imaging (Table 6). While the usable illumination laser spot

* For speed determination, the repetition rate of the laser must allow illuminating features of the jet at least twice while these traverse the field of view of the camera, which extends typically a few hundred micrometre along the length of the jet. At 100 m/s, it takes the jet 5 µs to transport features over a distance of 500 µm. To capture a feature twice, a repetition rate exceeding 200 kHz is required.

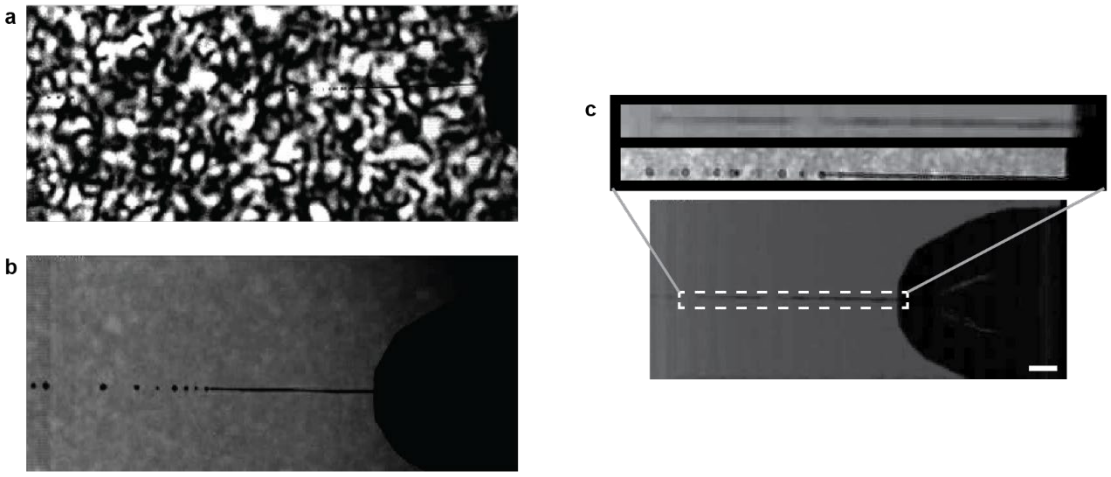


Figure 8.2 Imaging microscopic jets produced by a GDVN. **a**. Pronounced laser speckle prohibits imaging of the jet. Speckle was generated by a single diffuser (220 grit polish) in the optical path. **b**. Addition of two diffusers (1500 and 600 grit polish), a refocusing lens (25.4 mm focal length), a 49 m long multimode fibre of 300 µm inner diameter and a biconvex lens (100 mm focal length) focusing to a point just beyond the jet reduce speckle and allows obtaining clean images of the jet. **a**,**b**. Both images use the VGEN fibre laser as illumination source and show a field of view of 812x1854 µm^2. **c**. Snapshot of a water jet at 40 m/s using constant LED illumination and a camera exposure time of 250 ns. The inset shows a comparison between these conditions (top) and imaging with pulsed laser illumination of 3 ns pulse duration (bottom), illustrating the necessity for short-pulse illumination when imaging fast microscopic liquid jets. Scale bar 100 µm. Figures adapted from Grünbein *et al*[6].

	Fibre Laser VGEN-G-HE-20 (Spectra Physics)	**AlGaInP Laser Diode HL6545MG (Ushio Opto Semiconductors, Inc.)**	**Tailored Pulse Fibre Laser Pyroflex-25-GR (Eolite Lasers)**
Notable features & application target	- Direct beam allows speckle-free imaging - Up to 200 µJ pulse energy allowing illuminating large fields of view	- Direct beam allows speckle-free imaging - Cost-effective - Small & lightweight, easy to transport to and implement in XFEL experiments	- Pulse bursts can be created with tailored intensities for each pulse, interesting for advanced fluid dynamic studies[6]

Table 6 Pulsed light sources tested for their jet imaging capabilities. Optical properties are listed in Table 1.

size of the laser diode is limited to ~300 µm (FWHM) to achieve sufficient intensity, it is a viable, cost-effective solution that, given its compact size and simple design, is easily set up and aligned both in the home laboratory and when ported to the experimental chamber at the XFEL.

Automating velocimetry

To measure jet speed, features need to be tracked over at least two time points. A straightforward method to measure the speed of GDVN jets without requiring tracer particles is to measure the speed of the droplets formed after break-up of the continuous jet into droplets (Rayleigh breakup). Generally, the average speed of droplets shortly downstream of the break up region provides a good measure of jet speed.

Two approaches are feasible for this when using both a triggered light source and a triggered camera: the jet can be imaged using a high repetition-rate camera and a light source illuminating the jet once per frame (double-frame, single-exposure) or a camera of any repetition rate can be used when illuminating the jet twice per frame (single-frame, double-exposure) such that two spatially offset images of the jet are captured within one frame[113]. In both cases the temporal separation of typically 0.1 - 5 µs between exposures must be adjusted to jet speed, jet diameter and field of view such that the translated image of each droplet can be clearly assigned.

For both approaches scripts were written that allow automatic measurement of jet speed. To allow instantaneous feedback, a fast and efficient algorithm is required. Since analysis of a one-dimensional data set is much faster than evaluations of a two-dimensional data set such as an image, the ideal scenario is that of only having to evaluate a single line of pixels per image, namely those pixels in the jet centre. Since the jet's mean position remains mostly stable, changing only when sample or flow rates are changed, jet position and orientation only have to be determined once per recording. This is achieved by finding the line of minimum intensity (jet and droplets scatter light away from the camera sensor) inside the region of interest including only the free jet. To remove any non-uniform background from the intrinsic illumination profile, a third order polynomial function is fitted to the line as background and subtracted.

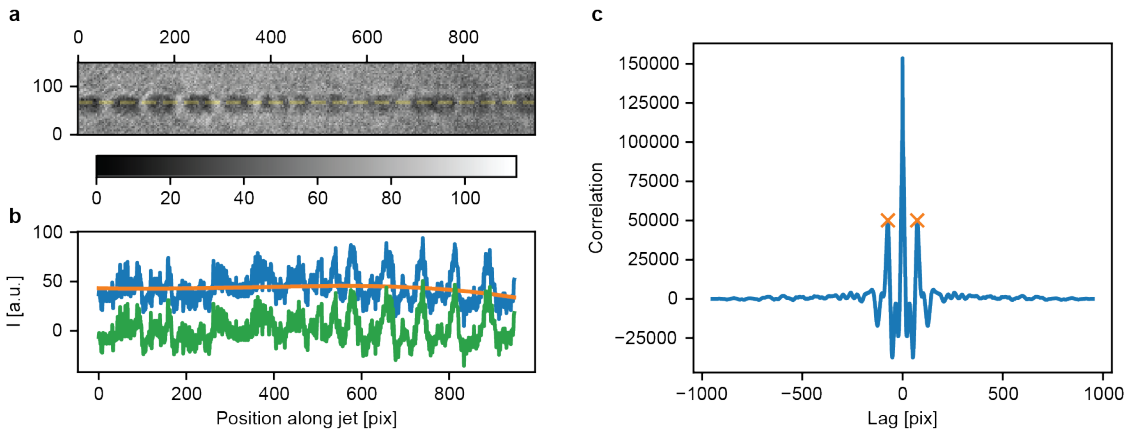


Figure 8.3 Extracting jet speed from double-exposure images. **a**. Double-exposure image taken at the SPB/SFX instrument of EuXFEL using the red laser diode described in Table 6 for illumination. The jet is imaged using two ns pulses separated by 250 ns. The dashed yellow line indicates the line of pixels at the jet centre which was used for analysis. **b**. Intensity along the jet centre (blue), third-order polynomial fit to the background intensity (orange) and background-corrected intensity (green) used to calculate the autocorrelation for each image. **c**. Average autocorrelation of all images. Correlation of the intensities with a shifted signal of itself is maximal when the shift is either zero or when it corresponds to the average distance travelled by the droplets between two exposures (marked in orange). Given a pixel size of 0.215 µm/pix, droplets move ~15.7 µm between exposures, corresponding to a jet speed of ~63 m/s.

To extract jet speed, the distance travelled by the droplets between two exposures must be measured. In case of single-frame, double exposure images the dominant spatial frequency in an intensity scan along the line of droplets must arise from the translation of the jet between the two exposures. Extracted either via a fast Fourier transform (FFT) or as the second maximum in the autocorrelation of the intensity scan, this together with the temporal separation of the two exposures delivers jet speed (Figure 8.3). Similarly, for double-frame, single exposure images the correlation of the jet image of two subsequent frames is evaluated, which is maximal for the jet displacement between the two frames. For both image acquisition modes droplet speed can be determined separately along multiple sections of the field of view, thereby allowing not only to measure average droplet speed but also droplet acceleration and the initial velocity at droplet pinch-off. Even for the rare case of non-zero acceleration of the droplets which occurs only under very particular experimental conditions (as a function of nozzle geometry, liquid and gas flow rates as well as chamber pressure), the initial velocity at droplet pinch-off provides a good measure of jet speed.

Jet velocimetry at MHz repetition rate XFELs in the presence of fs X-ray pulses

To track features travelling with the jet, one can also track the movement of XFEL-induced gaps in the jet since these are flushed downstream by the subsequently injected sample at the speed of the jet itself[8]. At MHz XFELs two pulses separated by a time interval Δt inducing two distinct X-ray explosions in the jet can therefore be exploited for speed measurement by measuring the distance d between the two gap centres, from which jet speed v is obtained as $v=d/\Delta t$ (Figure 8.5a). This procedure was applied during all beam times where MHz repetition rates were employed.

8.3 MHz repetition rate experiments at European XFEL

8.3.1 Feasibility of collecting SFX data at MHz repetition rate

First experiments investigating the feasibility of MHz repetition rate experiments were conducted at European XFEL in June 2018, aiming to test whether (i) sample can be delivered sufficiently rapid into the interaction region and (ii) XFEL induced shockwaves have any effect on the collected data. The experiment was conducted on crystals of the well-characterized model protein lysozyme, and investigations were performed at 7.47 and 9.22 keV photon energy to check for differences due to different amounts of energy deposited in the sample. In addition to testing the general feasibility of MHz repetition rate experiments, the data quality obtainable at a brandnew facility operating the accelerator in an unprecedented mode of operation needs to be assessed. Therefore, a mixture of jack beam protein crystals was investigated to explore whether data quality is high enough to allow analysing uncharacterized samples.

As preparation for the experiment, GDVN nozzles particularly tuned for generating high-speed jets were fabricated and tested with the described velocimetry setup (chapter 8.2). For the experiment at European XFEL, the setup shown in Figure 2.1 was employed. Here, the more complicated femtosecond snapshot imaging as described in reference[8] was employed such that the temporal resolution theoretically allows imaging of the shock waves travelling at supersonic speeds which would not be possible when employing nanosecond laser pulses.

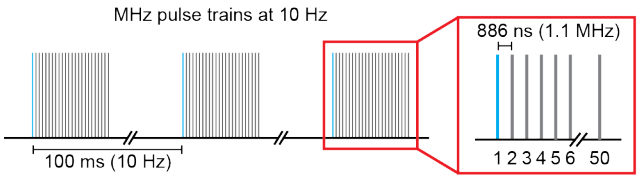


Figure 8.4 Pulse pattern at European XFEL employed during the experiment in 2018. 10 bursts (so-called 'trains') of X-ray pulses are produced per second, with each train consisting of 50 pulses at 1.1 MHz intra-train repetition rate. Within a train, pulses are thus separated by 886 ns. The first pulse in each train (coloured in blue) arrives ~100 ms after the last pulse of the preceding train.

At European XFEL, X-ray pulses arrive in 10 pulse trains per second, with MHz repetition rate within each train. Our experiment was one of the first experiments after inauguration of the accelerator when the maximum repetition rate within each train was restricted to 1.1 MHz instead of the 4.5 MHz target repetition rate. Within a single pulse train, 50 pulses separated by 886 ns were delivered (Figure 8.4). Between each train a pause of ~100 ms ensured that the first pulse of each train always probed sample not affected by shock waves. Thus, the burst structure of X-ray pulse delivery provides an intrinsic pattern for investigating potential shock wave damage by comparing data collected by the first pulse in each train to data collected by later pulses in the train.

Utilizing the femtosecond snapshot imaging, jet speed was measured *in situ* by imaging the jet shortly after the arrival of the second X-ray pulse in a train. At jet speeds of 40-50 m/s and higher, two distinct gaps in the jet were clearly visible in the optical image of the jet (Figure 8.5a). The presence of two distinct gaps separated by a section of continuous jet indicates that the second pulse indeed intercepted a recovered jet. Moreover, consecutive X-ray pulses in the same train probed different crystals (Figure 8.5b,c), proving that sufficiently fast sample delivery is possible for collecting data at 1.1 MHz repetition rate.

Shock waves could not be detected optically in the jets in the femtosecond snapshot images. This can be explained by the small jet diameter since the visibility of shocks decreases rapidly with this parameter[8]. Nevertheless, despite not being visible, appreciable shock waves may be launched by the impact of X-ray pulses, thus a potential effect has to be investigated by comparing data of unshocked sample to data of sample that may have experienced a shock wave. To keep conditions comparable over the experiment, jet speed and

diameter were kept constant during the beam time. Jet diameter determines both how much energy is absorbed by the liquid jet and thus the magnitude of the launched shock wave as well as the amount of damping within a given distance along the jet[8]. Additionally, jet speed also determines the distance displaced by the jet between two consecutive X-ray pulses, and thus the distance over which a shock wave may damp.

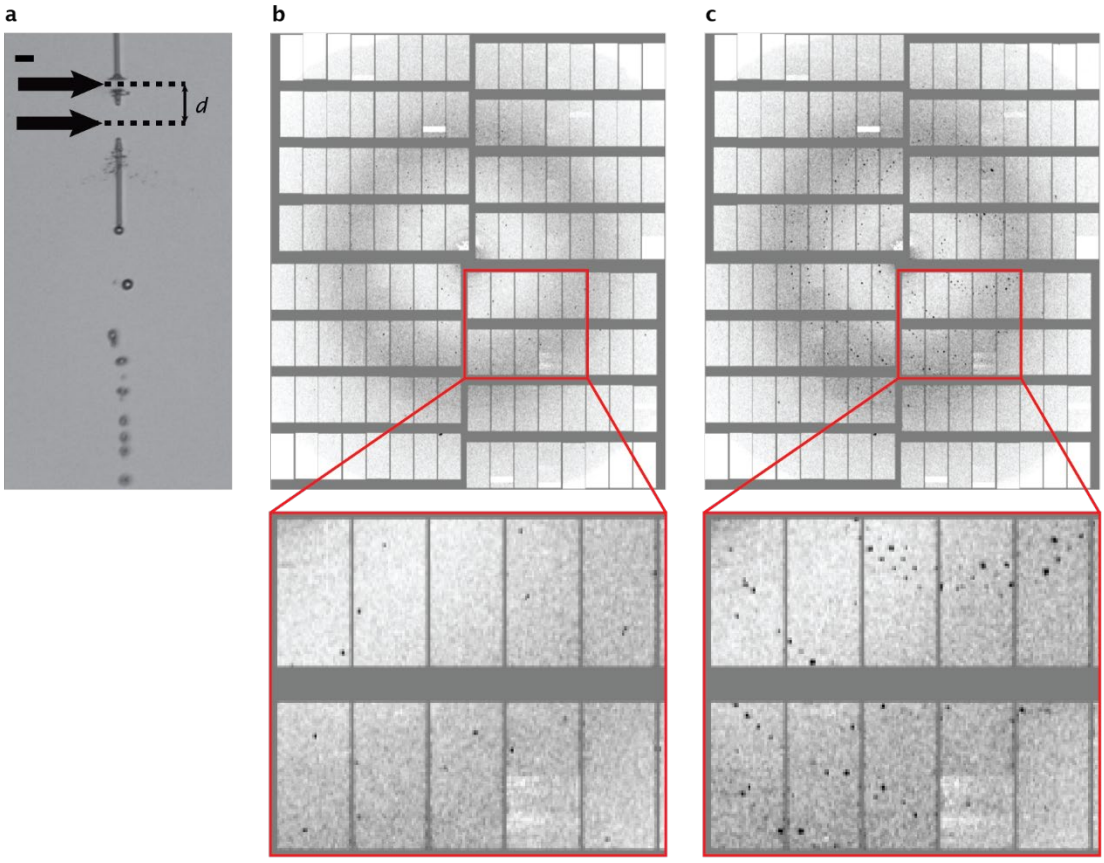


Figure 8.5 Data collection from two X-ray pulses separated by 886 ns (1.1 MHz). **a**. Femtosecond snapshot image of the 4 µm diameter liquid jet carrying lysozyme crystals after being hit by the first two X-ray pulses in a train. Both X-ray pulses lead to distinct explosions in the jet opening a visible gap (black arrows). At 45 m/s the jet is sufficiently fast to displace the gap created by one pulse in time for the consecutive pulse to probe a recovered jet segment which is also illustrated by the continous jet segment visible between the two formed gaps. The scale bar is 20 µm. **b**, **c**. Diffraction patterns of lysozyme microcrystals recorded with the first (**b**) and second (**c**) X-ray pulse in one X-ray pulse train. The inset shows a close-up view on a subset of the detector area, clearly showing two different diffraction patterns originating from two different crystals probed in different orientations. Figure reproduced from Grünbein *et al.*[1] under the CC-BY 4.0[3].

To check for shock wave damage, statistical indicators of diffraction quality typically employed in (SFX) crystallography were analysed for each pulse in the X-ray train. Parameters potentially expected to change upon pressure-induced damage are for example parameters that describe the quality of the ordered arrangement in the crystal. Such parameters include the resolution obtained in single diffraction hits, the signal-to-noise ratio (SNR) of the diffraction intensities or the fraction of diffraction patterns that can be indexed. Also, if crystalline order is strongly affected upon shock wave damage, the hit rate† may decrease. Moreover, the protein structure may change upon shock exposure.

Carefully analysing all data collected, no change in diffraction resolution was observed between sequential pulses in the train. Neither a change in the distribution of resolutions obtained for each indexed hit was observed in diffraction patterns collected by the first and second pulse in the train, respectively, nor a change in the average resolutions obtained with each pulse in the train (Figure 8.6a,b). Also R_{work}, R_{free} , CC* (Figure 8.6c) and the indexing rate (Figure 8.6d) remained constant over the pulse train‡. Other parameters, such has hit rate (Figure 8.6d), the signal-to-noise-ratio and the Wilson B factor indicate a change over the pulse train, which however can be explained by an equivalent drop in pulse energy over the pulse train. This analysis was performed for the data sets collected at 7.47 and 9.22 keV photon energy, both leading to the same conclusion[1].

† The hit rate is the fraction of detector images in which at least 10 peaks are detected. Under constant experimental parameters (X-ray pulse energy, crystal concentration, jet diameter, etc.) this should remain constant. If however the crystalline order reduces, crystal diffraction extends only to lower resolution, reducing the number of diffraction peaks which may ultimately reduce the hit rate.

‡ R_{work} and R_{free} describe the agreement between the structure factor amplitudes derived from the measured reflection intensities and those derived from the refined model with R_{free} being calculated on a subset of the data that was not used for refinement (Brünger[114]). CC* describes data quality and is derived from $CC_{1/2}$ which corresponds to the correlation of one half of the data set with the other half of the data set: $CC^* = (2CC_{1/2} / (1+ CC_{1/2}))^{0.5}$ (Karplus and Diederichs [115]).
The indexing rate is the ratio of the total number of hits that can be indexed and the total number of hits.

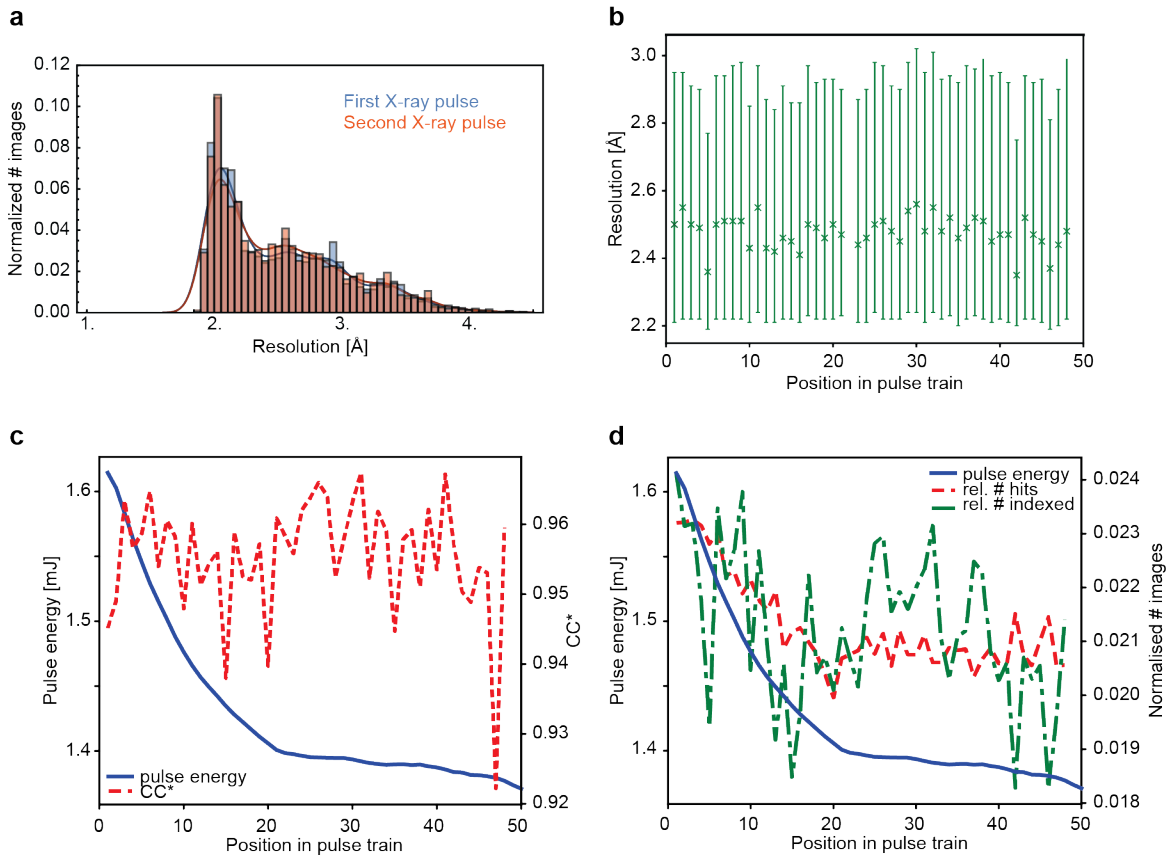


Figure 8.6 Diffraction data quality of lysozyme crystals collected at 1.1 MHz. Data was collected at 7.47 keV photon energy. **a**. Histograms of resolutions obtained for all indexed diffraction patterns collected by the first (blue) and second (red) pulse in the pulse train. **b**. Diffraction resolution as a function of position in the pulse train, showing the median resolution and the 0.25 and 0.75 quantiles. **c**. CC* of partial data sets (red) and pulse energy (blue) as a function of position in the pulse train. **d**. Hit and indexing rate (red and green, normalised by the number of images) as well as pulse energy (blue) as a function of position in the pulse train. All panels adapted from Grünbein *et al.*[1] under the CC-BY 4.0[3].

Megahertz data collection of an uncharacterized, complex system

To explore whether the data collected at MHz repetition rate using a novel X-ray detector would be of sufficient quality to permit the analysis of a more complex, uncharacterized system, further experiments were performed at European XFEL after demonstrating the general feasibility of exploiting the MHz repetition rate. For this purpose data was collected on a mixture of jack bean protein crystals, containing at least three different crystalline species that could not be characterized prior to the beam time due to the small crystal size.

Indeed, data collection at European XFEL at 1.1 MHz repetition enabled identification of crystals of the jack bean proteins urease, concanavalin A and concanavalin B. Of the latter two, sufficient amounts of high quality data could be collected to solve the structure of both proteins, showing no signs of damage when compared to previously published reference data[1].

Conclusion of first experiments

Taken together, these results show that high quality SFX data collection is possible at 1.1 MHz repetition rate, and that under the conditions used, neither protein structure nor crystal quality is affected by shockwaves launched by previous X-ray pulses.

However, while these results are highly promising for MHz data collection, it must be noted that the conditions of the experiment were less severe in terms of X-ray exposure than targeted, and in particular less severe than typical for crystallography experiments. The X-ray focus had a diameter of ~15 µm and was thus much larger than the target X-ray focus size of hundreds of nanometres (nanofocus) to a few micrometres (microfocus end station). Thus, photon density, being proportional to the beam radius squared, was significantly lower, resulting also in a lower magnitude of the XFEL-induced shock waves.

8.3.2 Investigating shock damage effects at standard X-ray fluence using an optimized pulse pattern for shock wave investigation

Upon optimizing the X-ray focusing optics at the SPB/SFX instrument[§], an increase in X-ray fluence was achieved by decreasing the X-ray focus to ~3 µm diameter[80]. Follow-up experiments were therefore performed, investigating the effect of stronger shock waves under these realistic SFX conditions of higher X-ray fluence. As a comparison to the first experiments, data was again collected on microcrystals of the model system lysozyme. Additionally, data was collected on myoglobin microcrystals to investigate effects on different protein systems with different properties. The results for the two systems led to the same

[§] Kirkpatrick-Baez mirrors instead of Beryllium focusing lenses were installed, leading to better focusing[80].

conclusion, therefore figures and data presented in this thesis refer only to the data collected on lysozyme.

To augment data collection efficiency, a new experimental design was employed, exploiting a new bunch patterning capability that allows dropping any number of X-ray pulses from the pulse train to create different pulse patterns. In the experiment presented here, each X-ray train was subdivided into 9 smaller bursts, so-called 'wagons' with 1.1 MHz intra-wagon repetition rate and a wagon separation of 13-14 µs (Figure 8.7). Each wagon consisted of four consecutive pulses separated by 0.9 µs (~1.1 MHz repetition rate) followed by a fifth pulse after 1.8 µs (~0.55 MHz) (Figure 8.7). This allows the determination of whether or not the shock effect is additive (by comparing data accumulated from 2^{nd}, 3^{rd}, 4^{th} pulses), reversible (4^{th} versus 5^{th} pulse) and possibly its damping time (by comparing all pulses). The time interval between wagons is long enough to avoid shock wave damage in the jet segment probed by all first pulses in the wagons: the full length of the jet from the interaction point to the nozzle (~200 µm) replenishes within this time. At the nozzle tip the jet diameter increases from ~5 to 75 µm such that shock pressure is reduced rapidly due to energy conservation. Moreover, the pressure jump has already decreased significantly even before the shock reaches the meniscus, given that the pressure jump damps exponentially with travel distance within the jet as shown by Blaj *et al*[116]. Thereby, the employed pulse patterning allowed increasing data collection of the undamaged reference data set collected by all first pulses in the wagon from 10 Hz (no subdivision of pulse trains) to 90 Hz (pulse trains subdivided into 9 separate wagons).

As for the previous experiment, the first analysis step included merging all data collected by equivalent wagon positions. With increasing pulse position, i.e. increasing nominal exposure to previous shock waves, the diffraction resolution did not change (indicating no damage, Figure 8.8a) but a decrease in indexing rate was observed (Figure 8.8b) which may suggest shock-wave-related damage.

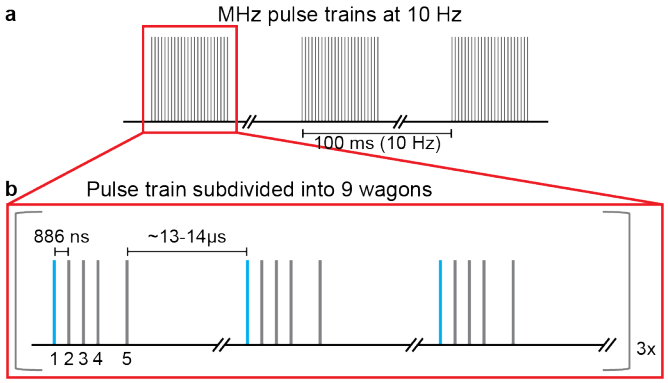


Figure 8.7 Optimized pulse pattern for investigating shock wave induced damage. **a**. At European XFEL 10 X-ray pulse trains are delivered per second. At the time of the experiment, the maximum repetition rate within each train was 1.1 MHz. **b**. Each train was subdivided into 9 smaller bursts ("wagons"). Each wagon consisted of five pulses: four pulses at 1.1 MHz were followed by a fifth pulse ~1.8 µs (~0.5 MHz) after the fourth pulse. Figure reproduced from Gorel & Grünbein *et al.*[5] under the CC-BY 4.0[3].

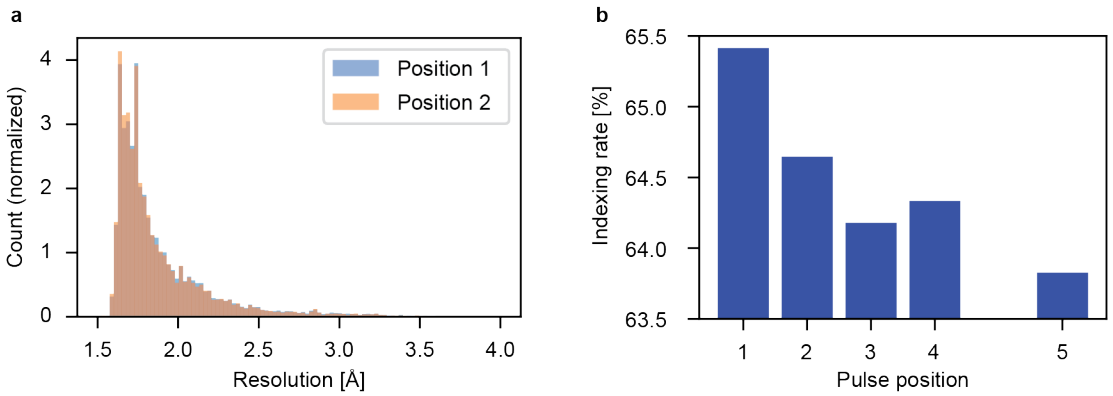


Figure 8.8 Resolution and indexing rate of shocked vs. unshocked data at 1.1 MHz repetition rate. **a**. Histogram of the resolution of all indexed diffraction patterns at wagon position 1 (blue) and 2 (orange). **b.** Average indexing rate at each pulse position. **a**, **b** .Data of wagons 1-6 at equivalent pulse positions was averaged. All panels adapted from Gorel & Grünbein *et al.*[5] under the CC-BY 4.0[3].

However, at the same time large systematic variations in experimental parameters (e.g., pulse energy, detector behaviour, X-ray alignment) were observed (Figure 8.9). While pulse energy and X-ray alignment directly impact the magnitude of the launched shock waves, all of these parameters also determine the quality of the detected diffraction pattern and might therefore lead to misleading conclusions in attributing apparent changes in data quality to shock wave damage. From the femtosecond snapshot imaging it became obvious that jet stability is an extremely important factor for the pulse-to-pulse alignment of jet and X-ray beam (Figure 8.9d-f). Since this alignment determines overlap of the X-ray beam profile with the jet, it impacts both the magnitude of the launched shock wave as well as the effective intensity with which scattering sample is probed and thus the obtainable data quality (Figure 8.9g, h).

To correctly analyse potential shock effects and disentangle effects on data quality caused by changing experimental operation conditions from those caused by shock-waves, it is therefore obligatory to work within two constraints: (i) only data collected under similar experimental conditions (absolute pulse energy, focus size, relative alignment between X-rays and jet) may be compared; and (ii) it needs to be ensured that the 'shocked' data was collected by X-ray pulses probing a section of the jet that has indeed experienced one or multiple shock waves launched by previous X-ray pulses in the same wagon. Due to instabilities of the jet and due to possible variations of the lateral pointing of the X-ray beam, a significant fraction of shots missed the jet partially or entirely and thus did not launch shock waves (Figure 8.9e,f). For an analysis of potential shock effects on upstream sample quality it therefore must be established for each pulse whether the jet was hit and how many of the preceding pulses of that wagon have also hit the jet.

To ensure that data in the later positions of a wagon actually have experienced a shock wave (i.e. to exclude shots in which previous X-ray pulses missed the jet), an algorithm was designed that uses the measured solvent scattering intensity to determine on a shot-per-shot basis whether a given pulse had interacted with the jet (see chapter 4.2). Based on this information, it was reconstructed for each shot how many previous pulses of the same wagon had interacted with the jet and may have launched shock waves affecting the current position. Furthermore, data was filtered such as to include only comparable experimental conditions: First, data collected by different wagons had to be analysed separately given the variation in pulse energy (Figure 8.9a). Second, to maximise the potential damage

signal and to ensure that the exposed and probed jet had been well-aligned with the X-ray beam, only those hits were considered that had been exposed to the maximum number of shocks launched by previous pulses (e.g. three experienced shocks at pulse position four in a wagon).

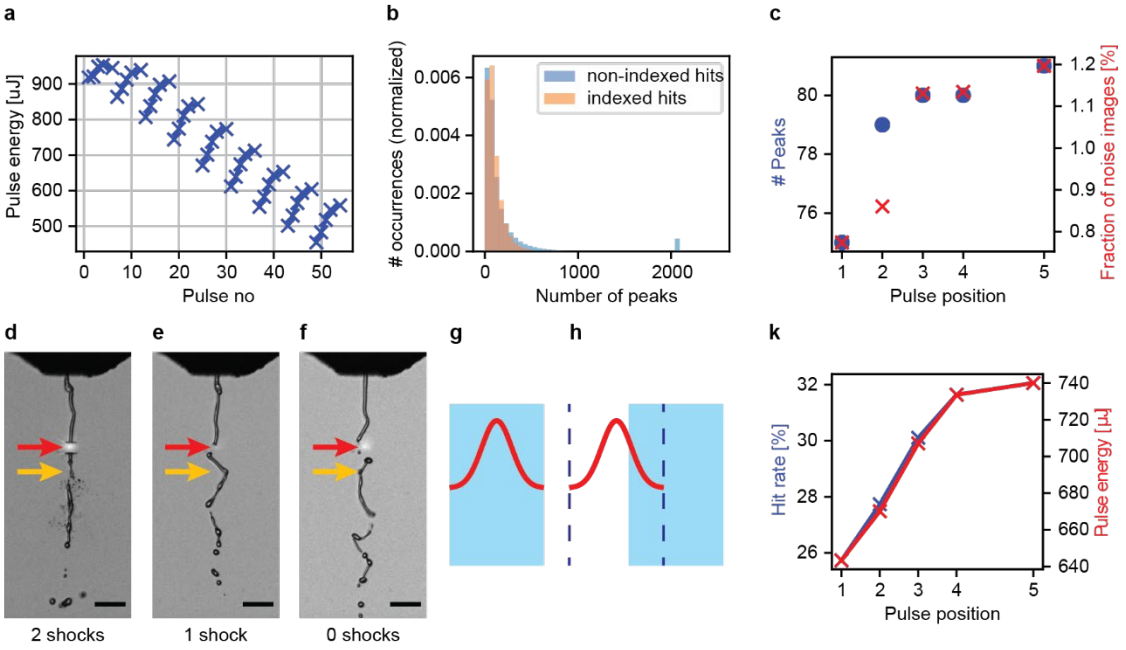


Figure 8.9 Systematic variation of experimental parameters with pulse train position and jet shape. **a**. The pulse energy decreases along the train by nearly a factor of 2, but increases within each wagon. **b**. Histogram of the number of peaks detected in diffraction patterns that can (orange) and cannot (blue) be indexed. Images with >1140 peaks could not be indexed. **c**. Average number of peaks per hit (blue circles) and fraction of hits with >1140 peaks (red crosses) as a function of pulse position in the wagon, averaging over wagons. The number of hits containing a peculiarly large number of diffraction peaks increases, lowering the indexing rate. **d**-**h**. The jet shape varies significantly, impacting data collection parameters and shock wave propagation. **d**-**f**. Femtosecond snapshot image of the jet after passage of the first two X-ray pulses in a train. The scale bar is 50 µm. The yellow and red arrows point at the segment that interacted with the first and second X-ray pulse, respectively. Due to changing jet morphology, both (a), only one (b) or none (c) of the two X-ray pulses may have interacted with the jet, leading to 2 (a), 1 (b) or 0 (c) shock waves affecting the sample segment to be probed by the subsequent X-ray pulse at pulse position 3. **g**-**h**. The relative alignment between X-rays (profile shown in red) and jet (light blue) changes upon wiggling of the jet and affects the effective probing intensity of the X-ray pulse, impacting data quality. **k**. The probability of hitting the jet (blue) and pulse energy (red) as a function of pulse position averaged over all wagons. All panels adapted from Gorel & Grünbein *et al.*[5] under the CC-BY 4.0[3].

These analyses revealed that the effect seen in the indexing rate when averaging over different wagons (Figure 8.8b) is caused by the increase in pulse energy with pulse position in the wagon, due to which the likelihood of detecting a hit increases (Figure 8.9k), while concomitantly increasing detector artefacts decrease the likelihood that a hit can be indexed (Figure 8.9b,c).

After incorporating the additional constraints, no indications of shock wave damage were found (Figure 8.10). The average diffraction resolution obtained for either tested crystalline system at a maximum repetition rate of 1.1 MHz depends neither on the number of previous pulses nor on the temporal delay to the last pulse (Figure 8.10), demonstrating that 1.1 MHz data collection is feasible also at higher X-ray fluence. While the advantages of the new pulse pattern could not be exploited in the current analysis due to instabilities in the accelerator operation (which are now solved), the presented experimental approach will be of great benefit to future experiments aiming at addressing the question of shock damage at the target X-ray repetition rate of EuXFEL of 4.5 MHz. It not only speeds up data collection, but it also allows to assess whether the shock effect is additive or reversible and possibly even extract shock damping times.

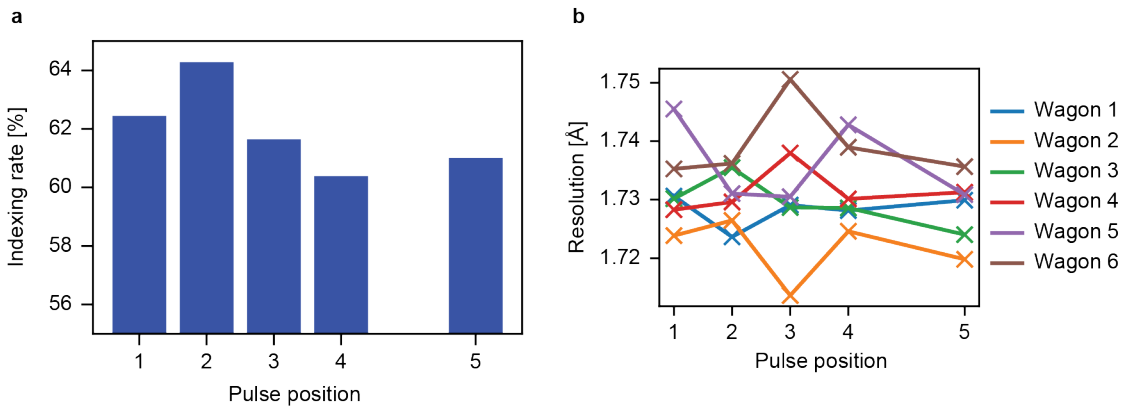


Figure 8.10 Lysozyme diffraction quality under comparable conditions at 1.1 MHz repetition rate. **a**. Indexing rate of hits in wagon 1. **b**. Median resolution of indexed crystal diffraction patterns as a function of pulse position, calculated using only data collected under comparable experimental conditions (comparable overlap between jet and X-rays, comparable pulse energy) exposed to the maximum number of possible shock waves. All panels adapted from Gorel & Grünbein *et al.*[5] under the CC-BY 4.0[3].

9 Investigating shock effects at shorter time scales

These first SFX experiments at EuXFEL employing two different X-ray fluences revealed that it is indeed possible to collect SFX data at 1.1 MHz repetition rate without any indication of shock wave damage[1,5]. The target in-train repetition rate at EuXFEL is, however, 4.5 MHz, at which pulses are only spaced by 222 ns[110]. The temporal interval in which possibly perturbed sample needs to recover between pulses is therefore much shorter, as is the jet displacement between two consecutive shocks. Since the shock wave damps with travel distance along the jet[116], the magnitude of the experienced pressure jump at the next jet segment to be probed is therefore much higher at higher repetition rate. Consequently, these more restrictive conditions have to be investigated in order to exploit the full repetition rate of EuXFEL for meaningful data collection.

To prepare for experiments at EuXFEL's target repetition rate and estimate the extent of shock effects on high-repetition rate SFX experiments, the potential effects of shock waves were investigated by using two closely spaced pulses in X-ray pump X-ray probe experiments using established techniques of pulse-pair generation at LCLS. In these experiments the first of the two pulses (the X-ray pump pulse) launches a shock wave in the liquid jet transporting the sample, the effect of which is probed by the second pulse (the X-ray probe pulse) after a set time delay.

Extensive parts of this chapter are to be published in references[2,9]. Sections may also be published in references[1,5,6,11,71].

9.1 Experimental design

Two X-ray pump X-ray probe experiments were performed in the microfocus chamber of CXI, LCLS to investigate shock wave effects at two different time delays (8.4 ns and 122.5 ns) between the two X-ray pulses. Similar to the experiments at European XFEL described in the previous section, GDVN injection was used to deliver protein microcrystals in a ~5 µm diameter liquid jet at ~50 m/s into the interaction region. To ensure that the probe pulse indeed interacted with a segment of the liquid jet instead of passing through

the gap created by the pump-pulse induced explosion, the two pulses were displaced vertically by Δx = 5 µm, such that the interaction region of the probe pulse was upstream (towards the injection nozzle) of the pump pulse (Figure 9.1). Since the detector installed at the beamline cannot separate diffraction patterns generated by two pulses separated by a nanosecond delay, the two pulses were tuned to different photon energies centred around the iron K absorption edge. An iron foil in front of the X-ray detector absorbed the pump but not the probe pulse[10,13], thereby detecting only the diffraction pattern generated by the second X-ray pulse (Figure 9.1).

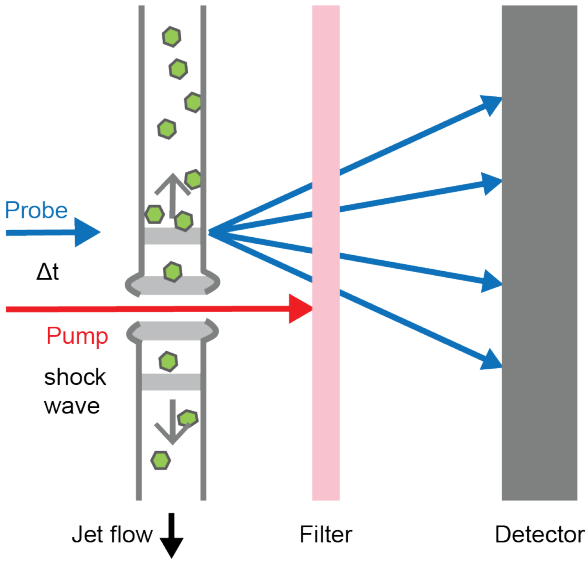


Figure 9.1 Experimental setup to investigate X-ray induced shock wave effects on SFX data collection in liquid jets. The first ~30 fs X-ray pulse, with photon energy ~40 eV above the iron K-edge (7.112 keV) and a pulse energy of up to ~0.1 mJ (0.03 mJ on average) was used as a pump, inducing ionization dynamics in the system that results in isochoric heating of the jet and the launch of a supersonic shockwave. The scattered X-rays were absorbed by a thin iron filter and did not reach the detector. After a time delay Δt (8.4 ns or 122.5 ns) a second ~30 fs X-ray pulse, with a photon energy ~40 eV below the iron K-edge and a pulse energy of ~0.9 mJ, displaced by ~ 5µm towards the nozzle, was used as a probe to measure diffraction, hitting a sample segment upstream of the pump pulse. In this case, the scattered X-rays passed through the iron filter, reaching the detector. The setup differs from previous two-colour X-ray pump X-ray probe experiments[10,13] by including a translation of the pump pulse interaction region. Figure adapted from Grünbein *et al.*[9] under the CC-BY 4.0[3].

Due to the complicated accelerator setup requiring two gun lasers to generate two nanosecond-spaced X-ray pulses of different photon energies focused to different positions in the interaction region, larger-than usual fluctuations in pulse energy and photon energy were expected than for the standard single-pulse operation mode. Therefore, as additional diagnostics, X-ray sensitive diodes fast enough to resolve both pulses separately were included in the setup, measuring the relative pulse energies of each pulse by detecting scattering off materials downstream of the experimental chamber (see chapter 4.3). Moreover, femtosecond snapshot imaging of the jet was employed to capture an image of the jet shortly after interaction of the probe pulse with the jet.

For both time delays the effect on lysozyme microcrystals was investigated, a well characterized model system. Moreover, to investigate whether a different (crystalline) protein system behaves differently, haemoglobin microcrystals were also studied which exhibit a high degree of plasticity.

To analyse effects of the shock wave launched by the first X-ray pulse on sample subsequently probed by the second X-ray pulse, diffraction data having experienced a shock wave (collected in the described X-ray pump X-ray probe) approach was compared to single-pulse data. Therefore, additional data sets were collected while suppressing the pump pulse. This resulted in single-pulse data collection at the base repetition rate of LCLS of 120 Hz. At this rate data collection cannot be impaired by shock wave damage (section 8.3).

9.2 Effective repetition rate simulated by the pump-probe experiments

The experiments were performed at pump-probe time delays of 8.4 ns and 122.5 ns, mathematically corresponding to repetition rates of ~120 MHz and ~9 MHz. However, the magnitude of the shock wave affecting the probed sample is determined by the effective spatial separation of the jet segments intersected by the two X-ray pulses since the propagation along the jet dampens the shock wave[116]. Due to the vertical offset of ~5 µm between the pump and probe pulse focus location (Figure 9.1), the segment intersected by the X-ray probe pulse would have been in the pump interaction region at a later time point than indicated by the mere temporal delay between the two pulses. With a jet speed of ~50 m/s,

it takes the jet ~100 ns to traverse the distance of 5 µm between pump and probe interaction region. Thus, the jet segment probed by the second X-ray pulse would have been in the pump interaction region after 8.4 ns + 100 ns = 108.4 ns and 122.5 ns + 100 ns = 222.5 ns for the two experiments, corresponding to repetition rates of ~9.2 MHz and ~4.5 MHz, respectively, if the pulses had been focused to the same interaction point.

9.3 Prerequisites for analysing shock wave effects

Identifying successful shock wave propagation

To analyse shock-wave effects in the pump-probe data set, it is a necessary prerequisite to ensure that a shock wave passed indeed through the segment probed by the second X-ray pulse. The femtosecond snapshot imaging of the jet revealed that fluctuations in jet shape are sufficiently large and fast that, at the larger time delay of 122.5 ns, one of the two pulses may interact with the jet while the other misses the jet (Figure 9.2). This behaviour was not observed at 8.4 ns time delay.

Therefore, for the 122.5 ns data, the femtosecond snapshot images of the jet recorded shortly after interaction of the jet with the probe pulse were used to identify shots in which (i) the pump pulse hit the jet and (ii) the jet shape supports propagation of potential shock waves to the later probed sample segment. Both are equally important, since even if the pump pulse hit the jet, the shock wave may be hindered to propagate to the jet segment to be probed by a break-up of the jet.

For this purpose a custom-written python script analysed jet shape, determining location and size of gaps in the jet as well as the size of the jet projected onto the horizontal plane. The latter indicates whether the angle of the jet to the X-ray beam axis has changed or whether the jet had a different diameter which impacts the magnitude and the decay of shock waves. The location and size of gaps in the jet may reveal if the pump pulse interacted with the jet, and whether it launched a shock wave that could propagate to the jet region probed by the second X-ray pulse. The approach is detailed in Figure 9.3.

The analysis discarded any hit in which more than two gaps were detected in the jet (jet had likely broken up, prohibiting shock wave propagation), in which no gaps were detected

(pump pulse had likely not hit the jet and thus did not launch a shock wave), in which identified gaps were less than ~4.5 µm upstream of the probe interaction region (probe pulse may have probed a singular droplet not affected by shock waves) or in which the projected jet size suggested extraordinary jet angles or diameters (Figure 9.2a-c). Thus, only those hits were used for shock analysis that do not show any abnormal jet shape, in which one or two gaps due to pump or pump and probe pulse are clearly visible (a gap due to the probe pulse does not need to exist since the recorded diffraction pattern reveals the presence of the probe pulse) and in which the jet is fully continuous upstream of the pumped segment, guaranteeing propagation of the shock wave upstream (Figure 9.2d-f).

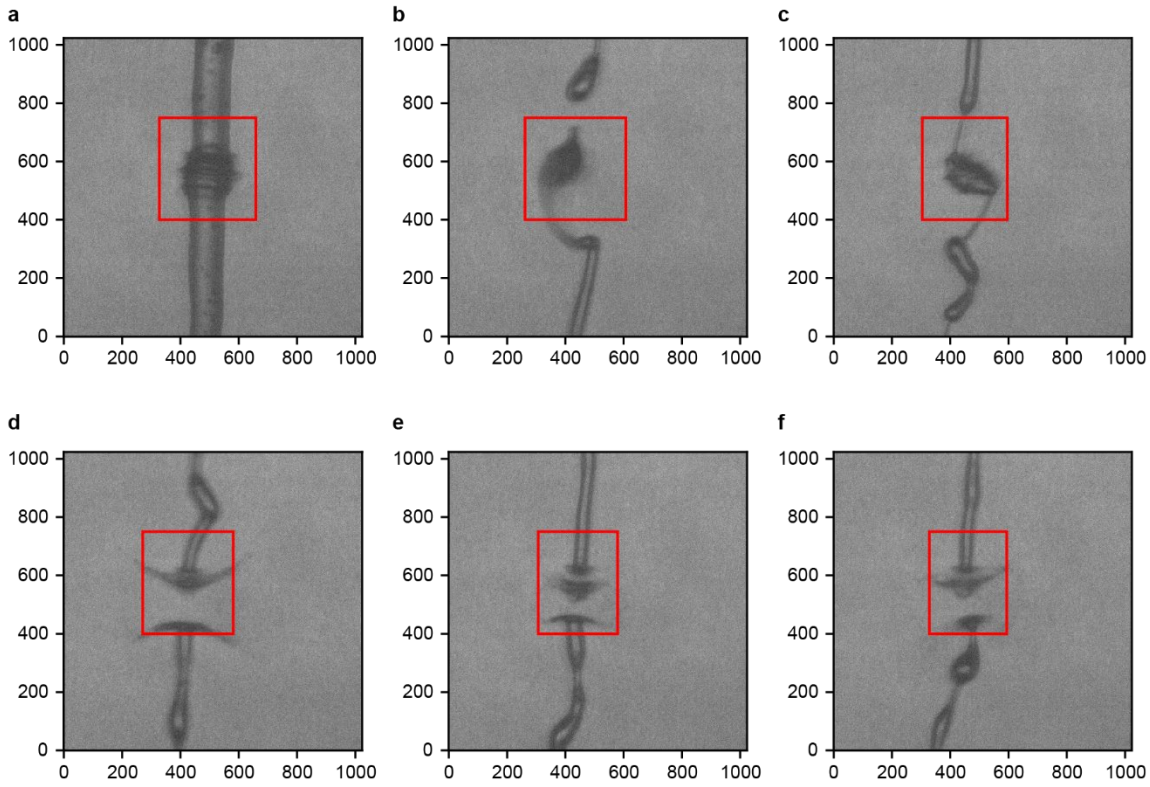


Figure 9.2 Femtosecond snapshot images of the jet in the pump-probe data of the 122.5 ns experiment. **a-c**. Exemplary jet images of data not considered in the shock analysis because of ambiguous shock wave propagation. **a**. Temporary much larger diameter jet, e.g. due to resolution of a temporary clog. Moreover, no clear pump-induced explosion is visible. **b**. Jet in the midst of breaking up. Propagation of the shock wave to the site of probing is ambiguous due to the break-up into droplets. **c**. Jet in the midst of breaking up, pump-induced explosion not visible. Most likely, no shock wave was launched. **d-f**. Exemplary images of data considered in the shock analysis. **a-f**. The magnification is 0.11 µm/pix. Figure adapted from Grünbein *et al.*[9] under the CC-BY 4.0[3].

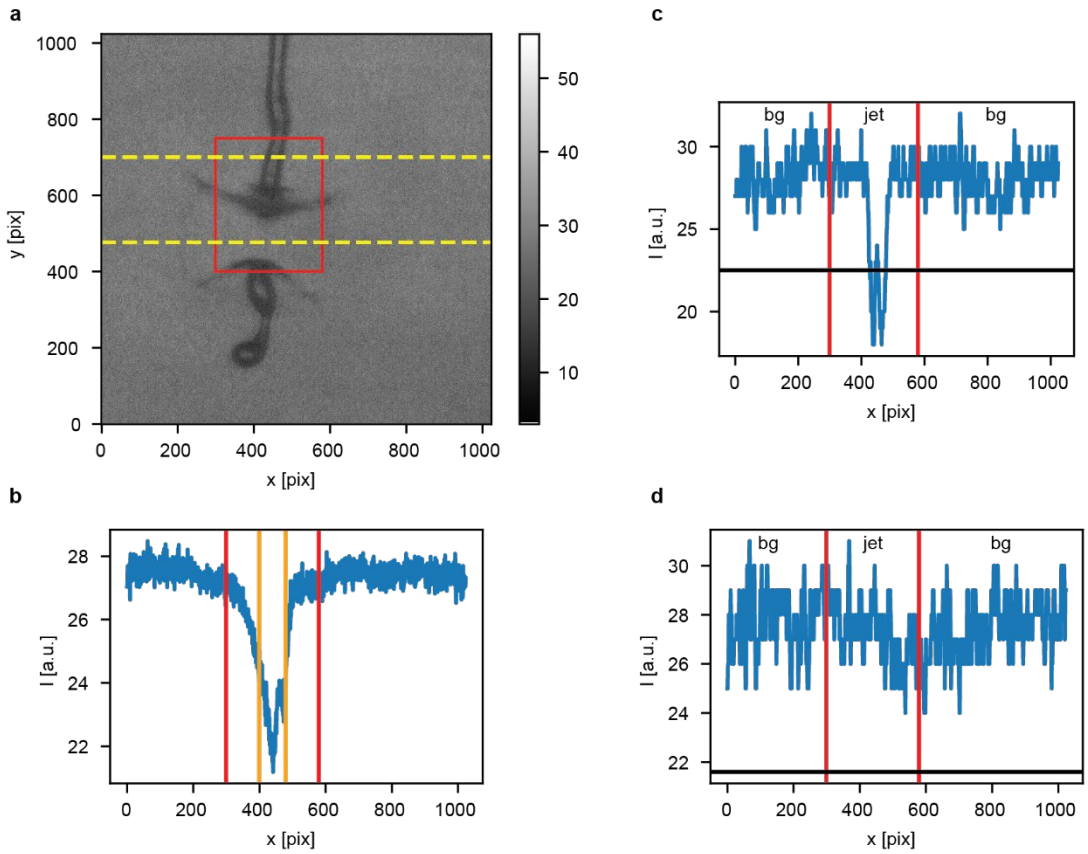


Figure 9.3 Classifying hits based on femtosecond snapshot images. Hits of the jet were characterized to identify pump-probe shots in which a well-defined shock wave has been launched by the pump pulse that affects the jet segment interrogated by the probe pulse. Such jets have neither abnormal morphology (e.g. very skewed), nor break up into droplets which prohibits shock wave propagation. The latter means that jet images must show 1 or 2 gaps in the jet (indicating being hit by the pump or pump and probe pulse) and the gaps must be located in the X-ray interaction region. **a**. Snapshot of the jet a few nanoseconds after interaction with the probe pulse. The gap due to the pump pulse interaction is clearly visible. The red rectangle marks the region of interest in which gaps within the jet are analysed, spanning a region from ~10 µm downstream to ~22 µm upstream of the pump interaction region. The magnification is 0.11 µm/pix. **b**. Projection of the mean pixel intensity onto the *x*-axis. The jet (orange boundaries) was defined as the region in which the projected pixel intensity was smaller than the median projected pixel intensity minus half the difference between the median and minimum intensity of the projection. Adding 100 pixels left and right to the centre (red boundaries) delivers the region of interest in *x* in which the jet is analysed. Hits with jets in which the projected sizes deviates by more than one standard deviation from the median projected jet size are discarded. **c**, **d**. Line scan intensities of the 5x5 median filtered image at y = 700 pix and y = 477 pix, indicated as dashed yellow lines in (**a**). By comparing the minimum intensity of the image in the background region (bg) to the minimum intensity inside the region of interest containing the jet (confined by the red boundaries), positions along the jet axis y can be identified in which a jet is present (**c**) or which represents a gap (**d**). The black horizontal line indicates the threshold (corresponding to 90 % of the minimum pixel intensity of the background regions) that pixel intensities within the jet region have to cross in order to be identified as a jet segment. All pixel rows are identified as “gap” or “jet”. To decrease vulnerability to noise, gaps closer than 2 pixels are merged and only gaps of at least 2 pixels in size are considered. To identify number and location of gaps, connected gap regions are counted and located. Figure with formatting changes from Grünbein *et al.*[9] under the CC-BY 4.0[3].

Ensuring that measurement of the pump diffraction signal is efficiently suppressed

Unavoidable fluctuations of the photon energy of the two pulses may lead to changes in photon energy for some pump-probe pairs such that also the photon energy of the pump pulse shifts below the iron K absorption edge. In this case also the pump pulse is transmitted through the iron foil in front of the detector, leading to the recorded scattering pattern being a superposition of the scattering signal originating from both pump and probe pulses. Since it is unclear to which extend the diffraction intensities of such hits originate from the first pulse, i.e. interrogating a sample segment that has not experienced a shock wave, these pump-probe pairs need to be excluded from the analysis. In the first experiment collecting data at 8.4 ns time delay, the X-ray photon energy spectrum was measured for each pulse, indicating that for most pump pulses ~1 % of photons had energies below the iron K absorption edge. To exclude cases where this may lead to measurable diffraction on the detector, additional simpler photon energy diagnostics were implemented in the experiment collecting data at 122.5 ns time delay,

For this purpose one of the two fast photodiodes used for measuring relative pulse energies was covered with a 25 µm iron foil. The second diode remained unmasked. Comparison of the signal of both diodes allowed checking whether the pump pulse photon energy was indeed above the iron K absorption edge, thus not erroneously contributing to the measured diffraction signal. For pump-probe data, the diode signals were therefore analysed to exclude hits where the pump pulse signal on the masked diode was high, potentially indicating (partial) drift of the pump photon energy below the iron K absorption edge. To find this threshold, the full trace of the diode signal was plotted for different pump pulse signals. Hits in which the pump pulse signal was within ~1.5 standard deviations of the mean pump pulse signal of all shots centred around 0 did not exhibit any visible pump pulse signal and were used for analysis.

Ensuring comparability between data sets

In addition to the described possible fluctuations in photon energy, the measured total pulse energy of both pump and probe pulse were distributed over a wide range. Changes in probe pulse energy can lead to changes in the observed diffraction quality. To disentangle whether a change in diffraction quality is thus due to a mere change in probe pulse

energy or indeed a change induced by shock waves, the distribution of probe pulse energies between pump-probe and single-pulse data sets needs to be comparable. To this end subsets of each data set were created in which the distribution of probe pulse energies was equalized between a given pump-probe data set and its corresponding single-pulse reference as described in detail in section 3.3. Using this method, samples are drawn with replacement. The data subsets thus contain some data duplicates, which are required for shifting the distribution of a variable. All subsets contain at least 65 % unique indexed hits.

9.4 Shock wave effect on crystalline order

Effect on hit rate, indexing rate and unit cell constants

For both lysozyme data sets at both time delays, as well as for the haemoglobin data set collected at 122.5 ns time delay, the average hit rate dropped significantly relative to the single-pulse reference data set (Table 7). To investigate whether this might be due to the pump-induced explosion of the jet, potentially leading to the probe pulse passing through the gap (indicating too low jet speed), the pump pulse energy was reduced until it was so low that no explosion was observed*. Even at this weak pump pulse energy the induced shock wave was strong enough to lead to a significant decrease in hit rate. While it cannot be excluded that changes in hit rate are due to a small difference (10 – 30 %) in probe pulse energy between pump-probe and single-pulse data collection (Table 7), such a dramatic drop most likely indicates that a corresponding fraction of crystals was damaged to such an extent that their diffraction patterns were degraded so much that they no longer qualified as a hit, meaning that less than 10 Bragg peaks could be detected. In contrast, the likelihood that a given hit can be indexed does not differ significantly between pump-probe and single-pulse data. Also, no significant changes in the values of the unit cell constants are observed†.

* This approach was performed during the pump-probe experiment employing an 8.4 ns time delay.

† A small shift (≤0.5 %) in unit cell constants is observed between the pump-probe and the single-pulse data set, with the unit cell apparently shrinking for the pump-probe data collection. This can be explained by a difference in the value of the X-ray wavelength used by the indexing program for the two data sets.
The indexing program extracts for each diffraction image the X-ray wavelength that is estimated from the properties of the electron bunches in the accelerator. This measurement cannot resolve the two pulses, and thus returns the average value for all closely spaced pulses. For the single-pulse reference data, this yields accurate photon energy estimates. For the two-pulse data, the average photon energy is

	Pump-probe	Single-pulse reference
Lysozyme, Δt=8.4 ns		
# hits (hitrate)	98 650 (11 %)	3 049 (40 %)
# indexed (indexing rate)	32 687 (33 %)	900 (30 %)
Probe pulse energy [a.u.]	0.09 ± 0.04	0.10 ± 0.04
Lysozyme, Δt=122.5 ns		
# hits[$] (hitrate)	96 146 (17%)	62 865 (37 %)
# indexed[€] (indexing rate)	12 793 (35 %)	20 097 (35 %)
Probe pulse energy [a.u.]	3.0 ± 0.5	2.8 ± 0.4
Haemoglobin, Δt=122.5 ns		
# hits[$] (hitrate)	43 003 (13 %)	25 742 (19 %)
# indexed[€] (indexing rate)	3 531 (24 %)	5 541 (23 %)
Probe pulse energy [a.u.]	3.3 ± 0.5	3.0 ± 0.5

Table 7 Hit rate, indexing rate and probe pulse energies for pump-probe experiments investigating shock wave effects on protein crystal diffraction. Probe pulse energies were measured by X-ray sensitive diodes. The values are comparable for data sets performed at the same time delay.

$ Total number of detected hits (diffraction images with ≥ 10 peaks). Hitrate is the ratio of the toal number of detected hits to the total number of recorded X-ray probe pulses.
€ Number of indexed hits that have passed the filtering conditions on jet shape and diode signal described in chapter 9.3. Indexing rate is the ratio of indexed images to number of hits, calculated only from indexed images and hits that have passed the filtering conditions.

Effect on diffraction resolution

Reduced crystalline order can be clearly detected as a loss in diffraction resolution if other data collection parameters, like probe pulse energy, remain constant. Analogously to the analysis of shock wave damage in the EuXFEL experiments (chapter 8.3), diffraction resolution was determined for each diffraction pattern at a fixed signal-to-noise threshold and

~40 eV larger than the targeted photon energy of the probe pulse (7.07 keV), such that the photon energy estimate used for indexing is ~0.6 % too high, leading to the determination of correspondingly smaller unit cell constants.

resolution distributions were compared between pump-probe and single-pulse data of equal probe pulse energy distributions.

For all three data sets, i.e. both time delays and both samples, a significant drop in resolution was observed in the pump-probe data set. At 8.4 ns time delay the resolution of lysozyme microcrystals decreased by ~0.6 Å (Figure 9.4a). Even at 122.5 ns time delay, probing a segment that experienced a far weaker shock wave, resolution had decreased by still 0.2 Å for lysozyme and 0.3 Å for haemoglobin (Figure 9.4b,c).

Pressure-dependent damage

The initial magnitude of the launched shock wave scales with the pulse energy of the pump pulse. Since the pump pulse energy was measured for each pulse by the fast X-ray sensitive diodes, shock wave damage can be analysed as a function of shock wave pressure (Figure 9.5). In the two-bunch mode tuning of the pulse energy of one pulse generally also affects the pulse energy of the other pulse, leading to possible (anti-)correlations in the behaviour of the two pulse energies (Figure 9.5). Due to the diffraction quality depending on probe pulse energy, only a limited range of the pump pulse energy bins is comparable to each other in the lysozyme data sets (Figure 9.5a,b). In case of haemoglobin, the positive correlation between pump and probe pulse energy does not even allow a differentiation

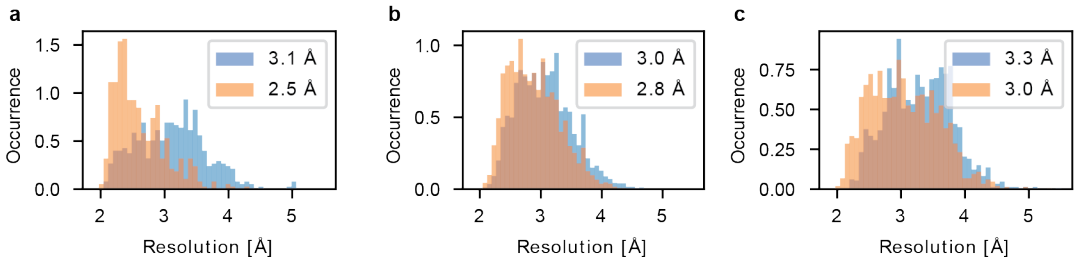


Figure 9.4 Crystals impacted by a shock wave diffract to lower resolution. The figure shows normalized histograms of the diffraction resolution at a signal-to-noise ratio of 4, comparing pump-probe data (blue) to the single-pulse reference (orange). For each pair, subsets of data with equal probe pulse energy distributions were used. The legend indicates the median resolution of each data set. **a**. Lysozyme, Δt = 8.4 ns. **b**. Lysozyme, Δt = 122.5 ns. **c**. Haemoglobin, Δt = 122.5 ns, panel adapted from Grünbein *et al.*[9] under the CC-BY 4.0[3].

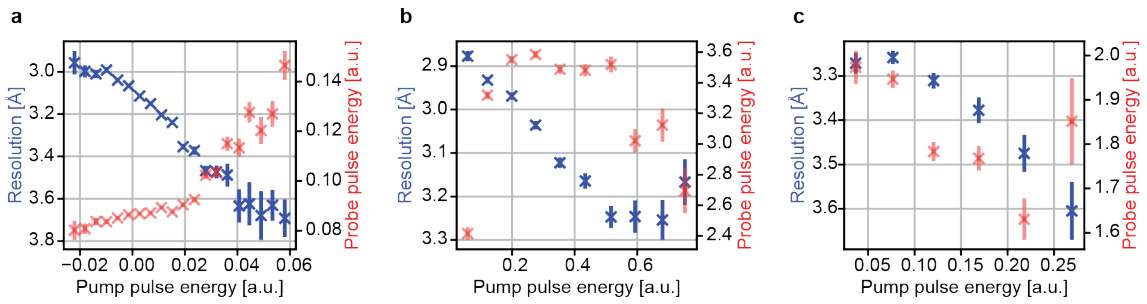


Figure 9.5 Shock wave effect as a function of pump pulse energy. Diffraction resolution at a signal-to-noise ratio of 4 (blue, left axis) and probe pulse energy (red, right axis) as a function of pump pulse energy for **a**. Lysozyme at Δt = 8.4 ns **b**. Lysozyme at Δt = 122.5 ns and **c**. Haemoglobin at Δt = 122.5 ns. x and y error bars represent the standard error of the mean. Panel **c** adapted from Grünbein *et al.*[9] under the CC-BY 4.0[3]. Absolute pump pulse energy values are comparable between **b** and **c**, while a different set up was used for measuring pump pulse energy in **a**.

whether the observed decrease in resolution as a function of pump pulse energy is caused by shock wave degradation or by the decrease in probe pulse energy (Figure 9.5c).

To allow a better comparison between bins at different pump pulse energy, subsets of each bin were assembled such that probe pulse energy distributions were equalized across bins (Figure 9.6). For lysozyme, at both time delays diffraction resolution decreases with increasing pulse energy of the pump pulse, resulting in larger damage (lower diffraction resolution) due to the stronger shock waves (Figure 9.6a,b). In line with the higher shock pressures at shorter time delay (at which subsequently closed jet segments are more closely spaced, leading to less pressure decay between pulses), the extend of shock damage is indeed stronger at 8.4 ns time delay than at 122.5 ns time delay. For haemoglobin, (Figure 9.6c) a statement on pressure-dependent behaviour cannot be made with confidence: Due to the generally low number of indexed hits in the data set and the strong correlation between pump and probe pulse energy, too few hits have comparable probe pulse energy conditions leading to a poor sampling of probe pulse energies especially at higher pump pulse energies (Figure 9.7). A similar analysis was performed on both lysozyme data sets. This showed that the last three (Figure 9.6b) as well as the first and last five bins (Figure 9.6a) for the 8.4 and 122.5 ns data sets, respectively, are not representative samples of the target probe pulse energy distribution.

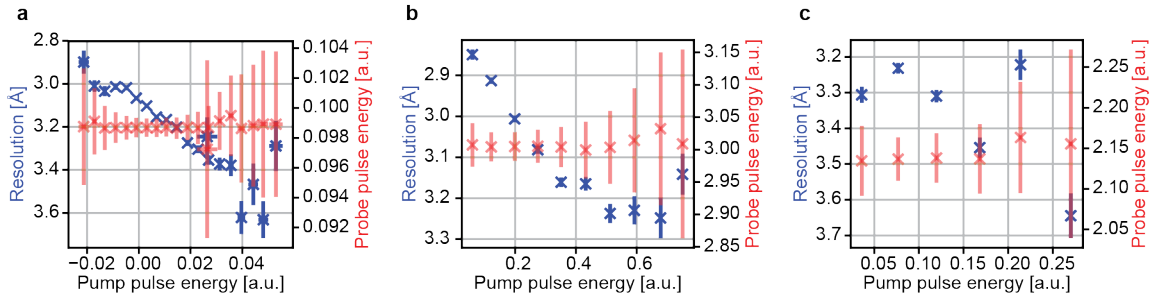


Figure 9.6 Pump-dependent shock wave effect after equalizing probe pulse energy distributions between bins. Diffraction resolution at a signal-to-noise ratio of 4 (blue, left axis) and probe pulse energy (red, right axis) as a function of pump pulse energy. N samples were drawn from a Gaussian target probe pulse energy distribution centred on μ with a standard deviation σ. **a**. Lysozyme at Δt = 8.4 ns, (N=20000, μ=0.099, σ=0.04). **b**. Lysozyme at Δt = 122.5 ns (N=25000, μ=3.02, σ=1.22). **c**. Haemoglobin at Δt = 122.5 ns (N=7000, μ=2.15, σ=0.94). Panel adapted from Grünbein *et al.*[9] under the CC-BY 4.0[3]. **a**,**b**,**c**. x and y error bars represent the standard error of the mean. Absolute pump pulse energy values are comparable between **b** and **c**, while a different set up was used for measuring pump pulse energy in **a**.

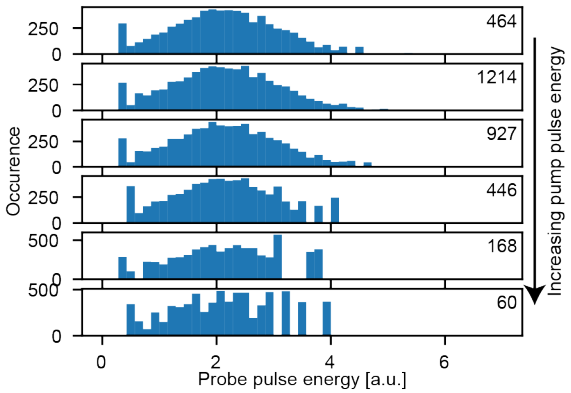


Figure 9.7 Probe pulse energy histograms for the data subsets analyzing the pressure-dependence of shockwave effects in haemoglobin after equalization of probe pulse energy distributions. The 6 histograms show the probe pulse energy distribution for the 6 bins in Figure 9.6. The number in the upper right corner of each histogram indicates the number of unique hits within each data set. Figure adapted from Grünbein *et al.*[9] under the CC-BY 4.0[3].

To extract the absolute magnitude of the launched shock wave as a function of pump pulse energy measured by the diodes, the diode signal was first calibrated to the pulse energy at the X-ray source point in mJ based on a series of single-pulse measurements. In a second step, this pulse energy was translated to a shock wave magnitude based on previous calibrations under similar experimental conditions published by Blaj *et al.*[116]. These calibrations were performed by Claudiu Stan and are further described in reference[2]. After translating the measured diode signal to absolute shock wave pressures, the results of experiments on lysozyme crystals at different time delays can be compared directly (Figure 9.8). Despite the different temporal delay between pump and probe pulse, the scaling of damage with shock pressure is similar in both experiments.

In conclusion, all pump-probe data sets display a loss in resolution with respect to the single-pulse reference data sets. The extent of damage depends on sample type but is similar for lysozyme and haemoglobin at comparable time delays (and shock wave pressures). Overall, the loss in resolution scales with the magnitude of the experienced shock wave.

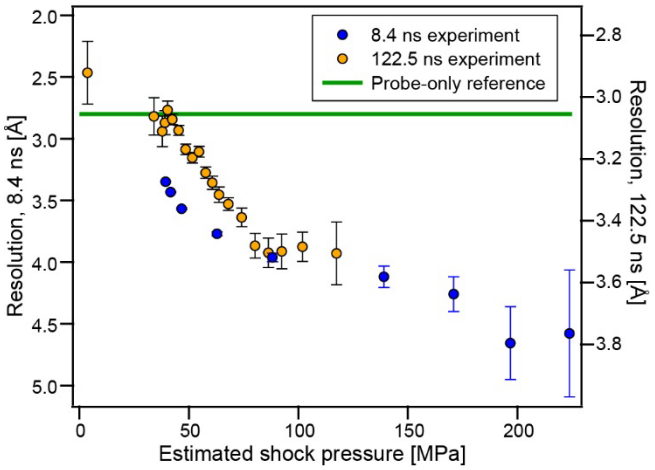


Figure 9.8 Shock wave effect on lysozyme crystals as a function of shock pressure. Indexed hits from both pump-probe experiments of lysozyme (Δt = 8.4 ns, blue and Δt = 122.5 ns, yellow) were binned by the experienced shock wave pressure and the average diffraction resolution (at a fixed signal-to-noise ratio of 5) evaluated. Errorbars represent the standard error of the mean. Figure by Claudiu Stan published in Grünbein *et al.*[2] under the CC-BY 4.0[3].

9.5 Shock wave effect on protein structure

For lysozyme microcrystals, no significant structural changes were observed neither for the 8.4 ns nor the 120 ns pump probe time delay. In contrast, small changes were observed in haemoglobin structures obtained from the pump-probe and single pulse data set respectively. Due to the small number of diffraction images and the limited resolution (Table 7), the significance of the atomic rearrangement was validated using a bootstrapping procedure (implemented by Thomas Barends and Alexander Gorel).

When overlaying the structures derived from the single-pulse and pump-probe data, respectively, small but significant differences in the peptide backbone are apparent (Figure 9.9). Importantly, many of the changes are correlated along helices and connecting loops with several amino acid residues being displaced in a similar direction. These correlated structural displacements include for example movement of helices towards the haem as well as compressive movements. Overall a small but significant compaction of the molecule and a decrease in the radius of gyration is observed.

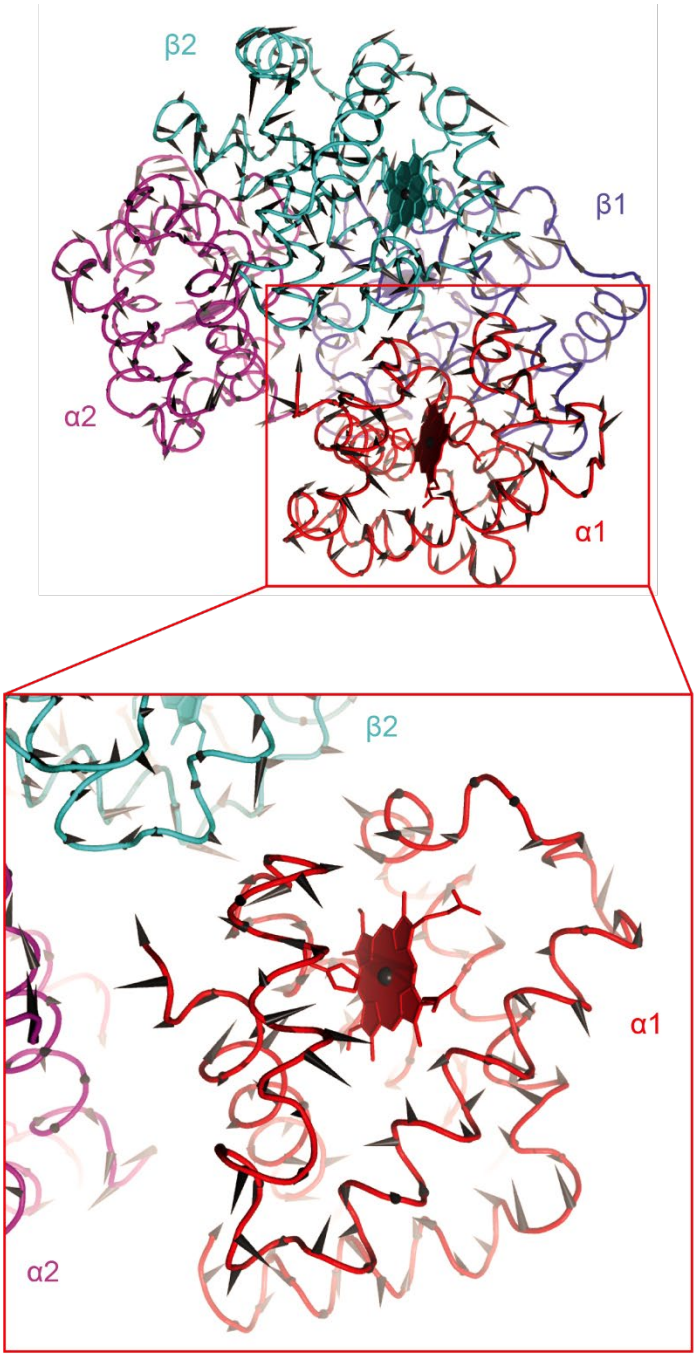


Figure 9.9 Shock-induced structural changes in haemoglobin. Haemoglobin consists of two α- and two β subunits, each containing a covalently bound haem cofactor, arranged as a dimer of α/β dimers. The alpha subunits are shown in red and magenta, beta subunits in blue and cyan; haem planes are depicted as filled planes. The figure compares structures obtained in the single-pulse and pump-probe data collection. The displacement between respective Cα positions is indicated by black arrows. The magnitude of the displacement is illustrated by the length of arrows (multiplied by a factor of 10). The fact that clusters of arrows point in similar directions shows that the displacements are correlated both within and between secondary structure elements. The inset shows a magnified view of the α1 subunit. Figure adapted from Grünbein *et al.*[9] under the CC-BY 4.0[3].

10 Discussion

In structural biology X-ray free-electron lasers offer unique experimental opportunities. The most transforming novelty introduced by XFELs is the possibility to perform time-resolved serial femtosecond crystallography (SFX) experiments observing molecular structural changes with atomic resolution at sub-picosecond time scales. In this thesis, such experiments were prepared and conducted to study the ultrafast isomerization in the photoreceptor bacteriorhodopsin. This required tackling numerous challenges encountered in this and similar experiments. Concomitantly, limitations arising when conducting experiments at next-generation XFELs capable of generating pulses with much higher repetition rate were investigated.

Parts of this discussion are published in references[1,2,4-7,9,11,71,88].

10.1 Time-resolved SFX experiment on ultrafast dynamics in bacteriorhodopsin

The time-resolved SFX experiment on bacteriorhodopsin (bR) described in this thesis was conducted at a time when experiments on ultrafast dynamics in protein crystals requiring a viscous environment, such as many membrane proteins like bR, had not yet been performed. From spectroscopy it was known that the light-induced isomerization reaction of the retinal cofactor in bR takes place on the femtosecond time scale[24], and that the quantum yield and dynamics of retinal in solution differ from those of retinal bound to rhodopsin[117]. The aim of the time-resolved SFX experiment on bR was therefore to investigate how the protein controls quantum yield, dynamics and specificity of the reaction and to elucidate the role of the coupling between the two in driving the isomerization reaction. Critical for the time-resolved SFX experiment and the interpretation of its results was the development of techniques allowing both the preparation of the correct starting state before the reaction, as well as the characterization of the excitation conditions employed to trigger the reaction. These steps were performed as part of this thesis.

The retinal in ground state bR naturally assumes one of two isomeric configurations ($bR_{13\text{-}cis}$ or $bR_{\text{all-}trans}$), and while only the all-*trans* retinal can undergo the functional photocycle,

the equilibrium between the two forms depends on the light conditions. Therefore, to maximise the fraction of molecules that can be triggered into the desired reaction and thereby maximise the measured signal, suitable illumination conditions have to be established *in situ* during the SFX experiment. For this purpose sample injection equipment was adapted such that continuous illumination of the sample stream upstream of the X-ray interaction region allowed increasing the all-*trans* fraction by a factor of 2 from 40-45 % to 65-80 %. Given that the half-life of light-adaptation in bR microcrystals is much shorter than the time required to install and measure a fresh sample batch, performing the light adaptation *in situ* greatly enhances signal and is the only way to ensure a constant high level of light-adapted bR capable of undergoing the functional photocycle. Therefore, *in situ* light adaptation is highly preferable over simple offline light adaptation outside of the experimental chamber prior to injector loading as done in similar experiments by other groups[94,95]. This simpler but inferior approach has the additional disadvantage that sample composition changes during data collection due to the decay of the light adapted state. If collection of the reference dark state structure fails to capture exactly the same change in sample composition, undefined systematic errors are introduced that distort the light-induced differences between pump-probe and dark data sets.

To achieve defined and reproducible starting states for each individual excitation, any illuminated sample needs to be removed from the interaction region in time before the subsequent excitation pulse. Failure to do so either results in re-excitation and probing of an ill-defined state, or in probing a previously excited crystal instead of the dark ground state. Due to large variations in jet speed, sometimes even including complete halts of the jet, it is critical that jet speed is measured *in situ* during the experiment when working with crystals embedded in a viscous material. Jet behaviour is highly dependent on sample consistency which varies from batch to batch. In particular, the overly simplistic procedure applied in other experiments[95] of measuring jet speed of similar samples offline before the beam time or calculating average jet speed based on sample flow rate and the expected diameter of the jet is not sufficient as shown in this thesis. Online jet speed measurements have to be performed.

During the bR beam time measurements of jet speed were performed manually by tracking features carried in the jet over time. While this successfully allowed excluding data in which jet speed variation was too large for reliable pump excitation conditions, the method

can be improved by automating the procedure to allow faster feedback. First steps towards automation have been made during a later beam time at SwissFEL also employing viscous jets. Interaction of the jet with the X-ray pulse changes the consistency of the viscous matrix in a localized region around the site of impact, appearing as a dark stripe on the jet if suitable jet imaging conditions are established. Consecutive X-ray pulses intersecting the moving jet thus leave behind multiple such stripes, appearing as a "ladder steps" in the brighter jet with the step size depending on both jet speed and X-ray repetition rate. The distance between these marks of impact is easily obtained from jet images and provides, together with the known X-ray pulse repetition rate, immediate feedback on jet speed.

A significant issue for functional interpretation of the collected SFX data on bR is the high pump power density used for photoexcitation, likely leading to multiphoton excitation. The choice was guided by pump power densities employed in previous time-resolved experiments on other proteins[16], and the as always tightly allocated beam time did not allow for a power titration like it is typically performed in spectroscopic pump-probe experiments. With 630 GW/cm^2 the incident power density was much larger than the tens of GW/cm^2 typically employed in spectroscopic experiments investigating similar samples. In particular, follow-up spectroscopic investigations have shown that bR's response starts to deviate from a linear behaviour already at ~30 GW/cm^2, indicating multiphoton events[7].

Under the described experimental conditions the ultrafast retinal isomerization in bR could be followed at 1.8 Å spatial resolution, uncovering retinal twisting already at the earliest measured time point of 240 fs[7]. Both protein structure and water network change on the sub-ps time scale, and an oscillatory behaviour in the retinal and surrounding residues was observed, suggesting vibrational coupling between those components[7]. However, given the high pump power excitation all results have to be interpreted in a multiphoton framework and it remains unclear whether the observed changes are triggered by biologically relevant single-photon or multiphoton events.

10.2 Controlled photo-excitation in light-sensitive protein crystals

Not only our experiments on bR, but in fact all published ultrafast time-resolved SFX experiments have been carried out in a high intensity regime, using 360-500 GW/cm^2 or more[16,20,21,95,102,103] and thus orders of magnitude higher than pump-probe experiments in

spectroscopy. While pump-probe-SFX and –spectroscopy are conceptually similar, the two methods differ significantly in their signal collection strategy: given unique spectral bands, spectroscopy is highly sensitive to small concentration differences (< 1 %) of particular species. It can therefore work at low excitations densities, exciting only a fraction of all molecules into another state. In contrast, SFX measures the superimposed signal of all species, requiring maximisation of the occupancy of the desired intermediate to facilitate structural interpretation. Thus, the general experimental tendency is to maximise the nominally excited volume fraction even when the thickness of crystals probed by the X-ray beam significantly exceeds the light penetration depth. This however requires very large intensities in order for a significant number of photons to reach the rear end of the crystal, thereby inevitably resulting in excessively high laser powers at the front of the crystal, likely inducing multiphoton absorption.

Multiphoton excitation can amplify nuclear motion (see Supplement of Barends *et al.*[16]), lead to additional radical intermediates[7,20] and open non-productive higher excited state relaxation channels, decreasing the single photon reaction yield[7,99,104]. For proof-of-principle experiments showing that time resolved SFX can successfully measure structural changes with femtosecond time resolution, signal maximisation under multiphoton conditions may be a pragmatic approach. However, it likely does not probe the biologically relevant reaction which is almost exclusively the response to a single-photon absorption event.

In view of the omnipresence of multiphoton issues in time-resolved SFX experiments, the topic of appropriate light excitation in such experiments was investigated systematically in this thesis. In particular, previously published statements claiming that the power density available for photoexcitation is reduced by >80 % upon scattering on the crystal-carrying jet[95] were inspected, and a detailed protocol for preparing time-resolved pump-probe experiments was established based on all findings[4].

First, appropriate excitation conditions under which the signal scales linearly with incident laser power, indicating a single-photon absorption regime, have to be identified. Power titrations evaluating this are well-established in spectroscopy. Given the scarcity of XFEL beam time and the complexity of SFX data interpretation, the power titration is therefore most easily done spectroscopically. In a second step, this power density needs to be trans-

lated into the required laser power in SFX experiments, in which the incident power density may change due to refraction and reflection at interfaces of the crystal-carrying jet and the crystal itself, as well as due to possible absorption by the jet medium.

Therefore, interactions which may change the incident power density were quantified individually in this thesis using a combination of measurements and simulations where appropriate. My investigations showed that absorption by common jet media is low (<2 %), as is reflection at jet and crystal interface (< 10 % at each interface). Reflection at the crystal interface can be estimated only in an approximate fashion since it depends on crystal shape, orientation and refractive index, with crystal orientation even varying for each XFEL exposure. The numbers presented here were therefore obtained as an estimate for the upper limit on reflectance. A sample-specific alternative to this generic upper bound would seem to be the measurement of the total light scattering cross section of the crystals, which however is not useful because it cannot quantify the fraction of light penetrating the crystal (see Supplement of Grünbein *et al.*[4]). In contrast to the small power density changes due to absorption and reflection, refraction at the jet changes the incident power density significantly, leading to inhomogeneous illumination inside the jet. The focusing effect of the cylindrically shaped jet surface in fact causes an increase in power density in large regions of the interior volume (on average a 10 – 60 % increase is observed depending on jet type (large diameter viscous jet vs. small diameter GDVN jet) and relative alignment of pump and probe optical axes).

These results are in stark contrast to the previously published claims[95] (which did not provide the underlying data) that scattering on the jet reduces the power density effectively available for triggering a reaction inside the jet by 80 %. However, the authors measured scattering on an LCP jet which only quantifies the sum of all scattering events that re-direct the incident beam (reflection, diffraction, refraction). While refraction at the cylindrical jet surface leads to strong focusing of the incident beam, it does not result in intensity losses in the jet interior. Solely the intensity measured at a downstream detector is reduced due to the large divergence of the beam after passage through the jet.

The quantification of effective power densities in crystal-containing jets presented here is based on a number of idealised assumptions and simplifications. For example, in most calculations changes in jet shape, interference effects and back reflection from rear jet and

crystal interface are neglected, while the incident beam is assumed to have a uniform intensity profile. Therefore only the idealised average situation is quantified, not the true intensity change for all pump-probe shots. Introducing these simplifications however allows approximating excitation regimes in SFX experiments in jets quantitatively. Based on these quantifications, a step by step protocol was established[4], guiding others to identify suitable excitation regimes and translate these into the required experimental laser parameters for SFX measurements. This includes both general guidelines as well as scripts[118] in python and Excel that were made publicly available to help design experiments and avoid misconceptions like in previous experiments[95].

10.3 Towards MHz repetition rate measurements

The occupancies of intermediate states populated in pump probe experiments are generally low. Even in the case of unity quantum yield maximally 50 % of the molecules will be in the excited state (saturation state, equilibrium of stimulated emission and absorbance) in a two-level system. In the linear regime one should not expect more than 30 % excited molecules. Given the generally small light-induced difference signal many diffraction images need to be collected for good data statistics and high signal-to-noise ratio. This data collection strategy, however, requires very long beam times which is currently a limiting factor because beam time is typically only granted for a few days per experiment.

The possibility of performing experiments at MHz repetition rate would increase the rate of data collection by orders of magnitude and facilitate measurement of weak signals. Moreover, sample delivered in a continuous stream would be more efficiently used.

The first XFEL capable of running at MHz repetition rate is the European XFEL that accommodated first user experiments in 2017. The EuXFEL is designed to generate X-ray pulses in pulse trains at a maximum peak repetition rate of 4.5 MHz, generating up to 27 000 pulses per second which is a tremendous increase from the previously highest repetition rate of an XFEL of 120 Hz at LCLS[23]. In view of the XFEL beam-induced jet explosions and shock waves affecting sample delivering liquid jets[8], such high repetition rates require two main questions to be answered in order for SFX experiments to become feasible: on the one hand, sample needs to be supplied for each pulse at sufficiently fast speed. On the other

hand, the X-ray-induced shock waves propagating along the liquid jet must not alter the sample's state.

First SFX experiments at EuXFEL employing two different fluences revealed that it is indeed possible to collect SFX data at 1.1 MHz repetition rate without any indication of damage by the shock waves[1,5]. This holds true for both lysozyme and myoglobin microcrystals and agrees with later results obtained by other groups in similar experiments[119,120]. These experiments were conducted during the early stages of accelerator operation when the EuXFEL design specification of 4.5 MHz repetition rate[69] was not yet available. With increasing repetition rate the shock wave pressure to which sample is subjected becomes larger because subsequent X-ray pulses probe segments of the jet that are separated by a smaller distance, allowing less decay in shock wave pressure.

Therefore, to investigate the effect of larger shock wave pressure under situations equivalent to higher X-ray repetition rates, two X-ray pump X-ray probe experiments were conducted at LCLS using two vertically offset pulses separated by 8.4 and 122.5 ns respectively. Mathematically this temporal separation corresponds to repetition rates of 119 and 8 MHz, respectively. However, it is the effective spatial separation of the jet segments probed by the two pulses that is relevant because the shock pressure and therefore the sample damage depends on this distance[8]. The additional vertical offset of the two pulses thus allows to simulate an experiment of effectively lower repetition rate than expressed by the pure temporal separation of the two pulses. In the described experiments at LCLS, the shock wave pressure experienced by sample probed by the second X-ray pulse is equivalent to the pressure experienced in 9.2 and 4.5 MHz experiments in which the foci of both pulses overlap.

Jet velocity is an important factor for the quantification of the actual shock wave pressure experienced by the probed sample since it determines the jet displacement between subsequent X-ray pulses. Moreover, jet speed is relevant for transporting sample sufficiently fast into the X-ray region such that each X-ray pulse probes a fresh jet segment. Alongside with the investigation of shock wave effects, techniques were therefore developed to generate, image and automatically characterize the high-speed micrometre sized jets required for sample injection at MHz repetition rate facilities[6].

In the 4.5 MHz and 9.2 MHz repetition rate surrogate experiments at LCLS, the shock wave launched by the pump pulse damaged microcrystals probed by the subsequent probe pulse. The extent of damage depended on time delay, pulse energy and sample. Damage was higher for shorter time delays and more intense X-ray pump pulses, which agrees with the larger shock wave pressure affecting the probed sample segment. At shorter time delays but constant jet speed, the jet translates less between shocks, thereby shortening the distance along which the shock wave may decay. Shock wave damage was investigated on lysozyme microcrystals at both time delays, as well as haemoglobin microcrystals at the longer time delay. For both samples at the longer time delay a similar loss of resolution was observed in the diffraction patterns. Remarkably, for haemoglobin, the protein structure also changed upon exposure of the microcrystals to the shock wave.

Given otherwise constant parameters, a loss in diffraction resolution is caused by a loss in crystalline order. This can either be due to random translations and/or rotations of the molecules constituting the crystal, or by random changes in molecular structure, or a combination of the two. For lysozyme a loss in resolution is observed at both time delays, but no changes in molecular structure are detected. In contrast, for haemoglobin the loss in resolution is accompanied by coordinated structural changes in the protein. The absolute changes are small, but a bootstrapping procedure quantifying errors in the determination of atomic positions confirmed their significance. Importantly, many of the changes are correlated along helices and connecting loops with several amino acid residues being displaced in a similar direction towards the haem cofactor, i.e. the active site of the protein. For investigative studies trying to elucidate binding pockets or molecular mechanisms, even such small changes may affect the interpretation and thus need to be avoided in clean experiments.

The shock wave pressure to which crystals are subjected in the described pump-probe experiments is on the order of a few tens of MPa. Importantly, the XFEL-induced pressure wave only transiently propagates through the microcrystals in the jet, which are then probed after the shock wave has passed and the pressure is released. This is different from experiments studying protein structure under static pressure. Static pressure experiments on Lysozyme[121] show a monotonic compression of molecular and internal cavity volumes up to a pressure of 710 MPa, as well as structural changes at high pressure (several hundred MPa) and a phase transition of the crystal symmetry at ~900 MPa. In contrast, no

similar behaviour, neither in the compression of the crystal nor in the structural changes, is observed in the diffraction data collected on lysozyme after shock wave pressure, which could either be due to the difference between transient and static pressurisation or due to the large difference in the absolute maximum pressure.

The machine operation mode for experiments at LCLS to generate two pulses of different photon energy separated by up to 122.5 ns in time and imaged to two vertically offset interaction regions is highly unusual and complicated to run. This not only led to a strongly reduced pulse energy compared to standard XFEL operation modes, but may also cause larger variations in photon energy and X-ray focus size, both of which may lead to apparent changes in diffraction resolution. In the two experiments, the photon energy was therefore monitored using different diagnostics and it was excluded that changes in photon energy are the cause of the changes in diffraction resolution. In contrast, a change in X-ray fluence by changing focus size cannot be detected as easily. The reference data sets were collected in single-pulse runs in which the pump pulse was suppressed. This changes the orbit of the electron bunch due to missing wakefields of the other bunch, leading to slightly different X-ray pulse properties. Based on the clear trend in increased damage with increasing pump pulse energy it is however unlikely that such potential pulse property changes are the cause for the observed difference between the pump-probe and single-pulse data sets.

Despite the close approximation of the X-ray pump X-ray probe experiments performed at LCLS simulating 9.2 and 4.5 MHz repetition rate experiments by using vertically offset pulses separated by 8.4 and 122.5 ns, there are differences to experiments truly running at these repetition rates. Most obvious, while the shock wave pressure to which sample was exposed was comparable, the actual time delay was shorter, leaving less time for hypothetical relaxation processes. However, since the sample was subjected to shocks for only a few nanoseconds, any changes detected at 122.5 ns delay are either irreversible or slowly decaying, and are therefore likely to be also detected at a 222 ns time delay. Due to the unusual machine set up at the LCLS, the pump pulse energy in the X-ray pump X-ray probe experiments was for the majority of data below 0.1 mJ, much lower than in standard SFX experiments. In a 4.5 MHz SFX experiment the single-pulse energies are typically ~1 mJ, possibly leading to higher shock wave pressures and thus larger damage. However, the separation between consecutive hits along the jet would also be bigger since the jets need to be flown more rapidly to clear the larger gaps at 1 mJ. The additional decay of the

shock pressure over the larger separation may compensate the increase in the pulse energy, leading to a degree of damage similar to the one in our experiment. Theoretically, the pump-probe surrogate experiments investigating shock wave effects may even underestimate damage because potential additive effect of shock waves launched by continuously successive pulses at MHz repetition rate cannot be investigated.

XFEL-induced shock waves may affect not only crystalline samples but also proteins or other molecules in solution. While SFX experiments are typically conducted in thin microjets (diameter 3-5 μm), spectroscopy experiments[122,123] as well as small and wide angle solution scattering experiments (SAXS, WAXS)[122,124] typically use large diameter Rayleigh jets (~50-100 μm) for sample delivery. Since the decay rate of shock becomes smaller in larger diameter jets[116], higher pressure is expected for experiments using Rayleigh jets at MHz data collection rates. During the X-ray pump X-ray probe beam time at LCLS, first attempts to investigate the shock wave effects in a ~5 μm liquid jet on apo-Myoglobin solution at 8.4 ns time delay (again with vertical offset of the two pulses) were made. Only limited preliminary data could be collected due to shortage of beam time. While the results are therefore not fully conclusive, the SAXS scattering curves of the pump-probe data set are suggestive of potential protein unfolding upon exposure to the shock wave.

As detrimental as shock waves may be for collecting structurally native data, this approach may well open up a novel experimental regime that allows ns time-resolved studies of, for example, pressure-induced phase transitions in liquids, protein unfolding, and generation of pressure-induced pH jumps to trigger chemical reactions on ns time-scales. For example, a sub-microsecond (0.7 μs) pressure jump instrument[125] achieved pressure jumps of 0.25 GPa (2.5 kbar), with first experiments showing refolding times of 2.1 μs in a genetically engineered lambda repressor mutant. Molecular dynamics simulations supporting this result[126] and similar results on other systems[127] predict processes on the nanosecond time-scale. However, since no clean experimental method currently exists to explore this regime the ns time scale has remained largely unexplored, limiting our understanding of the fast initial steps in folding or unfolding. This experimental gap might be closed by X-ray pump X-ray probe experiments as described here.

It remains to be tested whether the changes in order and structure of protein microcrystals observed in the described two-pulse experiments truly extrapolate to full 4.5 MHz repetition rate at European XFEL. While this presents both novel opportunities and challenges,

the prospect of being able to measure SFX data already at 1.1 MHz repetition rate is promising, resulting in an over 50-fold increase in effective repetition rate at EuXFEL compared to LCLS. This, together with the techniques and procedures presented in this thesis, opens the door to time-resolved SFX experiments studying biologically meaningful reactions at atomic resolution in many systems.

11 Appendix

11.1 Guidelines for time-resolved pump-probe SFX experiments

These guidelines are originally published in reference[4].

11.1.1 Experimental protocol

1. Spectroscopic characterization of the photoreaction of protein solution. This includes determination of appropriate
    - Photoexcitation wavelength λ. The compromise is between a) a low molar absorption coefficient* (relevant for crystals, in order to increase penetration depth (see point 2 below)) and b) a high efficiency formation of the excited states and photoproducts of interest (and concomitant low excited state absorption, thereby maximizing population of those states leading towards the process of interest). For the chosen wavelength, determine the participating species, quantum yields, and time constants†.
    - Pulse duration τ (shorter than reaction time of process under study, peak power)
    - Peak power density P (see below): Perform a power titration to determine the linear regime for the chosen pulse duration and wavelength (abate competing multiphoton (de-)excitation, avoid formation of higher electronic excited states that can open new relaxation pathways and minimize the formation of ionized species)[128].
        - Identify the onset of nonlinear processes. The concentration of photoproduct (change in absorbance) should vary linearly with P up to a maximum

* The molar absorption coefficient applies to the case of natural (non-polarised) light and isotropically oriented molecules. In SFX experiments it thus does not describe the true penetration depth in each shot but rather an average assuming all crystal orientations to be equally probable (no flow alignment). The molar absorption coefficient corresponds to the molar extinction coefficient in case of negligible scattering.

† If several wavelengths are favorable and if femtosecond excitation is desired, the lowest energy (red-shifted) excitation wavelength should be chosen to excite the lowest excited state and to reduce formation of ionized and higher electronic states.

peak power P_{max}. For $P \geq P_{max}$; the absorption saturates as nonlinear (multiphoton) processes ensue. P_{max} must be determined for the chosen photoexcitation wavelength λ and pulse duration τ.

- For a Gaussian laser beam the peak power density [W m^{-2}] is given by $P = 2 \cdot E/(\pi \cdot w^2 \cdot \tau)$ where E is the total pulse energy [J],τ the pulse duration (FWHM) [s] and 2w the $1/e^2$ beam diameter [m] at the interaction point. The $1/e^2$ beam diameter 2w is related to the full width half maximum FWHM of the beam as $2w = 1.699 \cdot FWHM$.
    - For a Gaussian beam profile, the peak power density is twice that of a flat-top beam. The effective beam area is therefore $\pi \cdot w^2/2$ and accordingly only one-half the geometric beam area $\pi \cdot w^2$.
- As a rule of thumb, absorption saturates when the photon density per laser pulse equals the inverse of the absorption cross section σ (in cm^2); see below section "General remarks". This gives a maximum energy density per pulse of $(P \cdot \tau)_{max} = h\nu/\sigma(\lambda) = \frac{h\nu N_A}{2303\epsilon(\lambda)}$. Here, $(P \cdot \tau)$ is in [J cm^{-2}] (i.e. P in [W cm^{-2}] and τ in [s]), σ in [cm^2] and ϵ in [l mol^{-1} cm^{-1}]. For a large molar absorption coefficient $\epsilon = 5 \cdot 10^4$ l mol^{-1} cm^{-1} at λ = 500 nm ($\nu = 5.996 \cdot 10^{14}$ Hz), this yields $(P \cdot \tau)_{max} \approx 2.1$ mJ cm^{-2}, corresponding to only ~0.23 µJ focused to a spot with 100 µm (FWHM) diameter ($1/e^2$ diameter 2w = 170 µm). We stress that the above formula is only a guideline for the experiment and cannot replace an experimental power titration. It is a simple estimation for the power density at which ground state absorption saturates and takes no other processes (e.g. excited state absorption) into account.
    - For a given $1/e^2$ beam diameter 2w, the pulse energy E_{max} at which a given $(P \cdot \tau)_{max}$ is reached is obtained as $E_{max} = 0.5 \cdot 10^{-8} \cdot (P \cdot \tau) \cdot \pi \cdot w^2$. Here, E_{max} is in [mJ], $(P \cdot \tau)$ in [mJ cm^{-2}] and w in [µm].
    - The absorption cross section σ (in [cm^2]) and the molar absorption coefficient ϵ (in [l mol^{-1} cm^{-1}]) are related via Avogadro's number N_A: $\sigma = 2303 \cdot \epsilon / N_A$

- Determine the longest laser pulse duration τ compatible with the desired time-resolution of the TR-SFX experiment (taking into account lifetime of probed

states, resolution of timing tool, ability to address pump probe laser jitter, temporal binning of diffraction data, feasibility of pulse length stretching). Repeat the previous three steps at this longer τ to test whether the fraction of the desired state and/or the formation of undesired states change. The power titration taken as a reference for the SFX experiments should ultimately be performed with a pulse duration similar to that of the SFX experiment itself. To reduce multiphoton effects, the pump pulse duration τ should be kept as long as possible, but shorter than the fastest reaction time constant and formation of re-absorbing photo-products. In any case, the pulses should be Fourier transform-limited in order to avoid complications due to chirped pulse excitation. More advanced experiments relying on pump pulse shaping techniques (e.g. coherent control) require separate characterization beyond the scope of the present work.

2. Analysis and preparation of protein microcrystals: Calculate protein concentration c and penetration depth δ. Use molar absorption coefficient ϵ [l $mol^{-1}cm^{-1}$] at the pump wavelength λ determined in (1) to determine the penetration depth δ [µm] (the depth at which the incoming intensity has dropped to 1/e ≈ 37 % of its initial intensity). For a sample of concentration c [mol l^{-1}] this is given by
   - $\delta[\mu m] = -10^6 \cdot \log_{10}(1/e)/(\epsilon[l\ mol^{-1}\ m^{-1}] \cdot c[mol\ l^{-1}]) \approx 4343/(\epsilon[l\ mol^{-1}\ cm^{-1}] \cdot c[mol\ l^{-1}])$

   Produce crystals of appropriate dimensions for efficient photoexcitation (not significantly larger than penetration depth δ), either by growth or by fractionation/milling of overly large crystals.

3. Carry out spectroscopic characterization of the crystalline protein suspension. Test whether the parameters listed in step 1 (wavelength, pulse duration, and power density range for linear regime) are also suitable for and applicable to the crystalline system. Determine kinetic rates and yields.
   - The absorption maxima of the crystalline protein may differ from those in solution. If this is the case, the ideal wavelength λ for an SFX-pump-probe experiment

might not be that of the λ determined in step (1) with protein solution. All spectroscopic characterizations (steps 1,2) on the crystalline system should then be repeated for the corrected λ.

- Based on the spectroscopically determined rates, determine suitable time delays for the TR-SFX experiment.

In protein crystals, as opposed to proteins in solution, chromophore molecules are not randomly oriented with respect to the electric field vector of the pump beam (see footnote concerning the molar absorption coefficient). Photoselection rules may then enter. In contrast to traditional rotation data collection, in SFX the largely random orientation of the microcrystals in 3D delivered in jets with respect to pump and probe laser beams helps mitigate this issue. Nevertheless, the use of circularly polarized pump laser light is advisable.

4. Determine the effective pump intensity inside the jet as determined by
   - The medium's refractive index n at the pump wavelength
   - The pump laser polarization
   - The pump laser geometry

This comprises the following steps:

I. Measure the refractive index of the embedding medium at the pump wavelength and employ this to determine the reflectance R_{med} from the jet surface using Table 5 for the pump laser polarization and geometry,
II. Assume a reflectance R_{xtal} at the crystal/medium interface of $R_{xtal} \lesssim 10$ %,
III. Determine the increase β in power density inside the jet based on Figure 7.4 and Figure 7.5 (chapter 7.4), making use of the refractive index of the medium and pump laser geometry; in case of liquid jet injection estimate β~1.1 for GDVN jets (Figure 7.5 and chapter 7.4),
IV. For input intensity P_{in} the effective pump laser intensity P inside the jet is then $P = P_{in} \cdot (1-R_{med}) \cdot (1-R_{xtal}) \cdot \beta$,
V. Set a pump laser intensity $P \leq P_{max}$
where P_{max} denotes the termination of the linear regime as defined above.

If measurement of the refractive index of the jet medium is not possible, use the following rules of thumb:

I. Estimate reflectance at jet interface to be $R_{med} \lesssim 2\%\text{-}10\%$
II. Estimate reflectance R_{xtal} at crystal interface to be $R_{xtal} \lesssim 10\ \%$
III. Estimate increase β in power density inside the jet to be $\beta \sim 1.2$ (~ 1.5) for perpendicular (parallel) pump probe geometry for large-diameter jets and $\beta \sim 1.1$ for GDVN jets
IV. As above, $P = P_{in} \cdot (1\text{-}R_{med}) \cdot (1\text{-}R_{xtal}) \cdot \beta$, with
    a) $P \sim P_{in}$ for a perpendicular pump probe geometry and a viscous jet experiment
    b) $P \sim 1.2 \cdot P_{in}$ to $1.3 \cdot P_{in}$ for a parallel pump probe geometry and a viscous jet experiment
    c) $P \sim 0.9 \cdot P_{in}$ to $1 \cdot P_{in}$ for a GDVN jet

For a detailed description of spectroscopic experiments as outlined in our brief protocol above, see e.g. publications on the characterization of photoactive yellow protein[104,128,129].

11.1.2 Power titration during the SFX experiment

Together with a preceding spectroscopic power titration, our data and approaches presented here permit to set the pump laser intensity properly for use in pump probe SFX experiment to study the light-sensitive crystalline system in the linear single photon excitation regime. It is nonetheless conceivable that the light-induced changes in the electron density are so weak as to be unsuitable for refinement, even by increasing the number of diffraction images constituting the light data. This could be due to *inherently* low occupancies of intermediate states and/or due to the (potentially perceived) need of using large crystals (e.g., to ensure high resolution diffraction) so that the penetration depth of the pump laser is exceeded, resulting in "*experimentally-inflicted*" low occupancies. In this case, an SFX power titration should be performed to establish the lowest required photoexcitation power density that yields interpretable electron density maps. Which photophysical and photochemical processes are thereby induced and thus characterized by the time-resolved pump probe SFX experiment, must then be established by performing an additional (retrospective) spectroscopic power titration over a large and common range of pump

power densities. This, possibly together with quantum chemical analysis, may then reveal which processes were in fact observed in the SFX experiment at the used power density[7].

SFX power titrations performed within the optically established linear photoexcitation regime are highly desirable. However, this not only requires additional beam time and sample but often does not allow for fast enough feedback during the beam time, given the time required to analyse the data, particularly in case of low occupancy intermediates.

References

1 Grünbein, M.L., Bielecki, J., Gorel, A., Stricker, M., Bean, R. *et al.* Megahertz data collection from protein microcrystals at an X-ray free-electron laser. *Nat. Commun.* **9**, 3487, doi:10.1038/s41467-018-05953-4 (2018).

2 Grünbein, M.L., Foucar, L., Gorel, A., Hilpert, M., Kloos, M. *et al.* Observation of shock-induced protein crystal damage during megahertz serial femtosecond crystallography. *Phys. Rev. Research* **3**, 013046, doi:10.1103/PhysRevResearch.3.013046 (2021).

3 Creative Commons Attribution 4.0 International License, <https://creativecommons.org/licenses/by/4.0/>.

4 Grünbein, M.L., Stricker, M., Nass Kovacs, G., Kloos, M., Doak, R.B., Shoeman, R.L., Reinstein, J., Lecler, S., Haacke, S. & Schlichting, I. Illumination guidelines for ultrafast pump–probe experiments by serial femtosecond crystallography. *Nat. Methods* **17**, 681-684, doi:10.1038/s41592-020-0847-3 (2020).

5 Gorel, A., Grünbein, M.L., Bean, R., Bielecki, J., Hilpert, M. *et al.* Shock Damage Analysis in Serial Femtosecond Crystallography Data Collected at MHz X-ray Free-Electron Lasers. *Crystals* **10**, 1145, doi:10.3390/cryst10121145 (2020).

6 Grünbein, M.L., Shoeman, R.L. & Doak, R.B. Velocimetry of fast microscopic liquid jets by nanosecond dual-pulse laser illumination for megahertz X-ray free-electron lasers. *Opt. Express* **26**, 7190-7203, doi:10.1364/OE.26.007190 (2018).

7 Nass Kovacs, G., Colletier, J.-P., Grünbein, M.L., Yang, Y., Stensitzki, T. *et al.* Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin. *Nat. Commun.* **10**, 3177, doi:10.1038/s41467-019-10758-0 (2019).

8 Stan, C.A., Milathianaki, D., Laksmono, H., Sierra, R.G., McQueen, T.A. *et al.* Liquid explosions induced by X-ray laser pulses. *Nat. Phys.* **12**, 966-971, doi:10.1038/nphys3779 (2016).

9 Grünbein, M.L., Gorel, A., Foucar, L., Carbajo, S., Colocho, W. *et al.* Effect of X-ray free-electron laser-induced shockwaves on haemoglobin microcrystals delivered in a liquid jet. *Nat. Commun.* **12**, 1672, doi:10.1038/s41467-021-21819-8 (2021).

10 Ferguson, K.R., Bucher, M., Gorkhover, T., Boutet, S., Fukuzawa, H. *et al.* Transient lattice contraction in the solid-to-plasma transition. *Sci. Adv.* **2**, e1500837, doi:10.1126/sciadv.1500837 (2016).

11 Grünbein, M.L. & Nass Kovacs, G. Sample delivery for serial crystallography at free-electron lasers and synchrotrons. *Acta Crystallogr. D* **75**, 178-191, doi:10.1107/S205979831801567X (2019).

12 Wickstrand, C., Nogly, P., Nango, E., Iwata, S., Standfuss, J. & Neutze, R. Bacteriorhodopsin: Structural Insights Revealed Using X-Ray Lasers and Synchrotron Radiation. *Annu. Rev. Biochem.* **88**, 59-83, doi:10.1146/annurev-biochem-013118-111327 (2019).

13 Nass, K., Gorel, A., Abdullah, M.M., V. Martin, A., Kloos, M. *et al.* Structural dynamics in proteins induced by and probed with X-ray free-electron laser pulses. *Nat. Commun.* **11**, 1814, doi:10.1038/s41467-020-15610-4 (2020).

14 Atkins, P.W., De Paula, J. & Keeler, J. *Physical Chemistry*. 11th edn, (Oxford University Press, 2018).

15 Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCrJ* **2**, 246-255, doi:10.1107/S205225251402702X (2015).

16 Barends, T.R.M., Foucar, L., Ardevol, A., Nass, K., Aquila, A. *et al.* Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350**, 445-450, doi:10.1126/science.aac5492 (2015).

17 Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406**, 752–757, doi:10.1038/35021099 (2000).
18 Kirian, R.A., Wang, X., Weierstall, U., Schmidt, K.E., Spence, J.C.H., Hunter, M., Fromme, P., White, T., Chapman, H.N. & Holton, J. Femtosecond protein nanocrystallography—data analysis methods. *Opt. Express* **18**, 5713-5723, doi:10.1364/OE.18.005713 (2010).
19 Boutet, S., Lomb, L., Williams, G.J., Barends, T.R.M., Aquila, A. *et al.* High-resolution protein structure determination by serial femtosecond crystallography. *Science* **337**, 362-364, doi:10.1126/science.1217737 (2012).
20 Coquelle, N., Sliwa, M., Woodhouse, J., Schirò, G., Adam, V. *et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nat. Chem.* **10**, 31-37, doi:10.1038/nchem.2853 (2018).
21 Pande, K., Hutchison, C.D.M., Groenhof, G., Aquila, A., Robinson, J.S. *et al.* Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science* **352**, 725-729, doi:10.1126/science.aad5081 (2016).
22 Miller, R.J.D., Paré-Labrosse, O., Sarracini, A. & Besaw, J.E. Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin in the multiphoton regime and biological relevance. *Nat. Commun.* **11**, 1240, doi:10.1038/s41467-020-14971-0 (2020).
23 Emma, P., Akre, R., Arthur, J., Bionta, R., Bostedt, C. *et al.* First lasing and operation of an ångstrom-wavelength free-electron laser. *Nat. Photonics* **4**, 641-647, doi:10.1038/nphoton.2010.176 (2010).
24 Ernst, O.P., Lodowski, D.T., Elstner, M., Hegemann, P., Brown, L.S. & Kandori, H. Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chem. Rev.* **114**, 126-163, doi:10.1021/cr4003769 (2014).
25 Pellegrini, C., Marinelli, A. & Reiche, S. The physics of x-ray free-electron lasers. *Rev. Mod. Phys.* **88**, 015006 (2016).
26 Pellegrini, C. X-ray free-electron lasers: from dreams to reality. *Phys. Scr.* **2016**, 014004 (2016).
27 Geloni, G., Huang, Z. & Pellegrini, C. *The Physics and Status of X-ray Free-electron Lasers* in *X-Ray Free Electron Lasers: Applications in Materials, Chemistry and Biology* Ch. 1, 1-44 (2017).
28 Margaritondo, G. & Rebernik Ribic, P. A simplified description of X-ray free-electron lasers. *J. Synchrotron Radiat.* **18**, 101-108, doi:10.1107/S090904951004896X (2011).
29 Andruszkow, J., Aune, B., Ayvazyan, V., Baboi, N., Bakker, R. *et al.* First Observation of Self-Amplified Spontaneous Emission in a Free-Electron Laser at 109 nm Wavelength. *Phys. Rev. Lett.* **85**, 3825-3829, doi:10.1103/PhysRevLett.85.3825 (2000).
30 Bostedt, C., Boutet, S., Fritz, D.M., Huang, Z., Lee, H.J., Lemke, H.T., Robert, A., Schlotter, W.F., Turner, J.J. & Williams, G.J. Linac Coherent Light Source: The first five years. *Rev. Mod. Phys.* **88**, 015007 (2016).
31 Aquila, A., Barty, A., Bostedt, C., Boutet, S., Carini, G. *et al.* The linac coherent light source single particle imaging road map. *Struct. Dyn.* **2**, 041701, doi:10.1063/1.4918726 (2015).
32 Colletier, J.-P., Sawaya, M.R., Gingery, M., Rodriguez, J.A., Cascio, D. *et al.* De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. *Nature* **539**, 43-47, doi:10.1038/nature19825 (2016).
33 Stauch, B. & Cherezov, V. Serial Femtosecond Crystallography of G Protein–Coupled Receptors. *Annu. Rev. Biophys.* **47**, 377-397, doi:10.1146/annurev-biophys-070317-033239 (2018).

34 Zhang, H., Unal, H., Gati, C., Han, Gye W., Liu, W. *et al.* Structure of the Angiotensin Receptor Revealed by Serial Femtosecond Crystallography. *Cell* **161**, 833-844, doi:10.1016/j.cell.2015.04.011 (2015).

35 Kang, Y., Zhou, X.E., Gao, X., He, Y., Liu, W. *et al.* Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* **523**, 561-567, doi:10.1038/nature14656 (2015).

36 Zhou, X.E., He, Y., de Waal, P.W., Gao, X., Kang, Y. *et al.* Identification of Phosphorylation Codes for Arrestin Recruitment by G Protein-Coupled Receptors. *Cell* **170**, 457-469.e413, doi:10.1016/j.cell.2017.07.002 (2017).

37 Ishchenko, A., Wacker, D., Kapoor, M., Zhang, A., Han, G.W. *et al.* Structural insights into the extracellular recognition of the human serotonin 2B receptor by an antibody. *Proc. Natl. Acad. Sci.* **114**, 8223-8228, doi:10.1073/pnas.1700891114 (2017).

38 Young, I.D., Ibrahim, M., Chatterjee, R., Gul, S., Fuller, F.D. *et al.* Structure of photosystem II and substrate binding at room temperature. *Nature* **540**, 453-457, doi:10.1038/nature20161 (2016).

39 Suga, M., Akita, F., Sugahara, M., Kubo, M., Nakajima, Y. *et al.* Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature* **543**, 131-135, doi:10.1038/nature21400 (2017).

40 Corbett, M.C., Latimer, M.J., Poulos, T.L., Sevrioukova, I.F., Hodgson, K.O. & Hedman, B. Photoreduction of the active site of the metalloprotein putidaredoxin by synchrotron radiation. *Acta Crystallogr. D* **63**, 951-960, doi:10.1107/S0907444907035160 (2007).

41 Meharenna, Y.T., Doukov, T., Li, H., Soltis, S.M. & Poulos, T.L. Crystallographic and Single-Crystal Spectral Analysis of the Peroxidase Ferryl Intermediate. *Biochemistry* **49**, 2984-2986, doi:10.1021/bi100238r (2010).

42 Kern, J., Alonso-Mori, R., Tran, R., Hattne, J., Gildea, R.J. *et al.* Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature. *Science* **340**, 491-495, doi:10.1126/science.1234273 (2013).

43 Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H. & Shen, J.-R. Native structure of photosystem II at 1.95 A resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99-103, doi:10.1038/nature13991 (2015).

44 Kern, J., Chatterjee, R., Young, I.D., Fuller, F.D., Lassalle, L. *et al.* Structures of the intermediates of Kok's photosynthetic water oxidation clock. *Nature* **563**, 421-425, doi:10.1038/s41586-018-0681-2 (2018).

45 Suga, M., Akita, F., Yamashita, K., Nakajima, Y., Ueno, G. *et al.* An oxyl/oxo mechanism for oxygen-oxygen coupling in PSII revealed by an x-ray free-electron laser. *Science* **366**, 334-338, doi:10.1126/science.aax6998 (2019).

46 Fukuda, Y., Tse, K.M., Nakane, T., Nakatsu, T., Suzuki, M. *et al.* Redox-coupled proton transfer mechanism in nitrite reductase revealed by femtosecond crystallography. *Proc. Natl. Acad. Sci.* **113**, 2928-2933, doi:10.1073/pnas.1517770113 (2016).

47 Schmidt, M. Mix and Inject: Reaction Initiation by Diffusion for Time-Resolved Macromolecular Crystallography. *Adv. Condens. Matter Phys.* **2013**, 10, doi:10.1155/2013/167276 (2013).

48 Woodhouse, J., Nass Kovacs, G., Coquelle, N., Uriarte, L.M., Adam, V. *et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nat. Commun.* **11**, 741, doi:10.1038/s41467-020-14537-0 (2020).

49 Wang, D., Weierstall, U., Pollack, L. & Spence, J. Double-focusing mixing jet for XFEL study of chemical kinetics. *J. Synchrotron Radiat.* **21**, 1364-1366, doi:10.1107/S160057751401858X (2014).

50 Calvey, G.D., Katz, A.M., Schaffer, C.B. & Pollack, L. Mixing injector enables time-resolved crystallography with high hit rate at X-ray free electron lasers. *Struct. Dyn.* **3**, 054301, doi:10.1063/1.4961971 (2016).

51 Kupitz, C., OlmosJr., J.L., Holl, M., Tremblay, L., Pande, K. *et al.* Structural enzymology using X-ray free electron lasers. *Struct. Dyn.* **4**, 044003, doi:10.1063/1.4972069 (2017).

52 Stagno, J.R., Liu, Y., Bhandari, Y.R., Conrad, C.E., Panja, S. *et al.* Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature* **541**, 242-246, doi:10.1038/nature20599 (2017).

53 Olmos, J.L., Pandey, S., Martin-Garcia, J.M., Calvey, G., Katz, A. *et al.* Enzyme intermediates captured "on the fly" by mix-and-inject serial crystallography. *BMC Biol.* **16**, 59, doi:10.1186/s12915-018-0524-5 (2018).

54 Chapman, H.N., Caleman, C. & Timneanu, N. Diffraction before destruction. *Phil. Trans. R. Soc. B* **369**, 20130313, doi:10.1098/rstb.2013.0313 (2014).

55 Botha, S., Nass, K., Barends, T.R.M., Kabsch, W., Latz, B. *et al.* Room-temperature serial crystallography at synchrotron X-ray sources using slowly flowing free-standing high-viscosity microstreams. *Acta Crystallogr. D* **71**, 387-397, doi:10.1107/S1399004714026327 (2015).

56 Sugahara, M., Mizohata, E., Nango, E., Suzuki, M., Tanaka, T. *et al.* Grease matrix as a versatile carrier of proteins for serial crystallography. *Nat. Methods* **12**, 61-63, doi:10.1038/nmeth.3172 (2015).

57 Sierra, R.G., Laksmono, H., Kern, J., Tran, R., Hattne, J. *et al.* Nanoflow electrospinning serial femtosecond crystallography. *Acta Crystallogr. D* **68**, 1584-1587, doi:10.1107/S0907444912038152 (2012).

58 Roessler, C.G., Kuczewski, A., Stearns, R., Ellson, R., Olechno, J., Orville, A.M., Allaire, M., Soares, A.S. & Heroux, A. Acoustic methods for high-throughput protein crystal mounting at next-generation macromolecular crystallographic beamlines. *J. Synchrotron Radiat.* **20**, 805-808, doi:10.1107/S0909049513020372 (2013).

59 Zarrine-Afsar, A., Barends, T.R.M., Muller, C., Fuchs, M.R., Lomb, L., Schlichting, I. & Miller, R.J.D. Crystallography on a chip. *Acta Crystallogr. D* **68**, 321-323, doi:10.1107/S0907444911055296 (2012).

60 Weierstall, U., Spence, J.C.H. & Doak, R.B. Injector for scattering measurements on fully solvated biospecies. *Rev. Sci. Instrum.* **83**, 035108, doi:10.1063/1.3693040 (2012).

61 Hunter, M.S., Segelke, B., Messerschmidt, M., Williams, G.J., Zatsepin, N.A. *et al.* Fixed-target protein serial microcrystallography with an x-ray free electron laser. *Sci. Rep.* **4**, 6026, doi:10.1038/srep06026 (2014).

62 Thomas D. Murray, Artem Y. Lyubimov, Craig M. Ogata, Huy Vo, Monarin Uervirojnangkoorn, Axel T. Brunger & Berger, J.M. A high-transparency, micro-patternable chip for X-ray diffraction analysis of microcrystals under native growth conditions. *Acta Crystallogr. D* **71**, 1987-1997, doi:10.1107/S1399004715015011 (2015).

63 Doak, R.B., Nass Kovacs, G., Gorel, A., Foucar, L., Barends, T.R.M. *et al.* Crystallography on a chip - without the chip: sheet-on-sheet sandwich. *Acta Crystallogr. D* **74**, 1000-1007, doi:10.1107/S2059798318011634 (2018).

64 DePonte, D.P., Weierstall, U., Schmidt, K., Warner, J., Starodub, D., Spence, J.C.H. & Doak, R.B. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Phys. D: Appl. Phys.* **41**, 195505, doi:10.1088/0022-3727/41/19/195505 (2008).

65 Weierstall, U., James, D., Wang, C., White, T.A., Wang, D. *et al.* Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat. Commun.* **5**, 3309, doi:10.1038/ncomms4309 (2014).

66 Vega, E.J., Montanero, J.M., Herrada, M.A. & Gañán-Calvo, A.M. Global and local instability of flow focusing: The influence of the geometry. *Phys. Fluids* **22**, 064105, doi:10.1063/1.3450321 (2010).

67 Herrada, M.A., Gañán-Calvo, A.M., Ojeda-Monge, A., Bluth, B. & Riesco-Chueca, P. Liquid flow focused by a gas: jetting, dripping, and recirculation. *Phys. Rev. E* **78**, 036323, doi:10.1103/PhysRevE.78.036323 (2008).

68 Kovacsova, G., Grünbein, M.L., Kloos, M., Barends, T.R.M., Schlesinger, R., Heberle, J., Kabsch, W., Shoeman, R.L., Doak, R.B. & Schlichting, I. Viscous hydrophilic injection matrices for serial crystallography. *IUCrJ* **4**, 400-410, doi:10.1107/S2052252517005140 (2017).

69 Decking, W., Abeghyan, S., Abramian, P., Abramsky, A., Aguirre, A. *et al.* A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator. *Nat. Photonics* **14**, 391-397, doi:10.1038/s41566-020-0607-z (2020).

70 Beale, J.H., Bolton, R., Marshall, S.A., Beale, E.V., Carr, S.B., Ebrahim, A., Moreno-Chicano, T., Hough, M.A., Worrall, J.A.R., Tews, I. & Owen, R.L. Successful sample preparation for serial crystallography experiments. *J. Appl. Crystallogr.* **52**, 1385-1396, doi:10.1107/S1600576719013517 (2019).

71 Grünbein, M.L., Bielecki, J., Gorel, A., Stricker, M., Bean, R. *et al.* MHz data collection of a microcrystalline mixture of different jack bean proteins. *Sci. Data* **6**, 18, doi:10.1038/s41597-019-0010-0 (2019).

72 Conrad, C.E., Basu, S., James, D., Wang, D., Schaffer, A. *et al.* A novel inert crystal delivery medium for serial femtosecond crystallography. *IUCrJ* **2**, 421-430, doi:10.1107/S2052252515009811 (2015).

73 Kuster, M., Boukhelef, D., Donato, M., Dambietz, J.S., Hauf, S., Maia, L., Raab, N., Szuba, J., Turcato, M., Wrona, K. & Youngman, C. Detectors and Calibration Concept for the European XFEL. *Synchrotron Radiation News* **27**, 35-38, doi:10.1080/08940886.2014.930809 (2014).

74 Fangohr, H., Aplin, S., Barty, A., Beg, M., Bondar, V. *et al.* in *16th International Conference on Accelerator and Large Experimental Control Systems.* 245-252 (JACoW, 2018).

75 Foucar, L. CFEL-ASG Software Suite (CASS): usage for free-electron laser experiments with biological focus. *J. Appl. Crystallogr.* **49**, 1336-1346, doi:10.1107/S1600576716009201 (2016).

76 Foucar, L., Barty, A., Coppola, N., Hartmann, R., Holl, P. *et al.* CASS—CFEL-ASG software suite. *Comput. Phys. Commun.* **183**, 2207-2213, doi:10.1016/j.cpc.2012.04.023 (2012).

77 White, T.A., Kirian, R.A., Martin, A.V., Aquila, A., Nass, K., Barty, A. & Chapman, H.N. CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* **45**, 335-341, doi:10.1107/S0021889812002312 (2012).

78 Bionta, M.R., Lemke, H.T., Cryan, J.P., Glownia, J.M., Bostedt, C. *et al.* Spectral encoding of x-ray/optical relative delay. *Opt. Express* **19**, 21855-21865, doi:10.1364/OE.19.021855 (2011).

79 Blaj, G., Caragiulo, P., Carini, G., Carron, S., Dragone, A. *et al.* X-ray detectors at the Linac Coherent Light Source. *J. Synchrotron Radiat.* **22**, 577-583, doi:10.1107/S1600577515005317 (2015).

80 Mancuso, A.P., Aquila, A., Batchelor, L., Bean, R.J., Bielecki, J. *et al.* The Single Particles, Clusters and Biomolecules and Serial Femtosecond Crystallography instrument of the European XFEL: initial installationThis article will form part of a virtual special issue on X-ray free-electron lasers. *J. Synchrotron Radiat.* **26**, doi:10.1107/S1600577519003308 (2019).

81 Allahgholi, A., Becker, J., Delfs, A., Dinapoli, R., Goettlicher, P. *et al.* The Adaptive Gain Integrating Pixel Detector at the European XFEL. *J. Synchrotron Radiat.* **26**, doi:10.1107/S1600577518016077 (2019).

82 Maltezopoulos, T., Dietrich, F., Freund, W., Jastrow, U.F., Koch, A., Laksman, J., Liu, J., Planas, M., Sorokin, A.A., Tiedtke, K. & Grunert, J. Operation of X-ray gas monitors at the European XFELThis article will form part of a virtual special issue containing papers

presented at the PhotonDiag2018 workshop. *J. Synchrotron Radiat.* **26**, 1045-1051, doi:10.1107/S1600577519003795 (2019).

83 F. Obier, W.D., M. Hüning, J. Wortmann. in *39th International Free-Electron Laser Conference.* (ed Dong-Eon Kim Ivan Andrian, Natalia Juszka, Michaela Marx, Raphael Müller, John Poole, Ruth Rudolph, Volker RW Schaa, Jens Völker) 353-356 (JACoW Publishing).

84 Stricker, M. *Addressing challenges in sample injection at high repetition free electron lasers* Master of Science thesis, University of Heidelberg, (2018).

85 Lomb, L., Steinbrener, J., Bari, S., Beisel, D., Berndt, D., Kieser, C., Lukat, M., Neef, N. & Shoeman, R.L. An anti-settling sample delivery instrument for serial femtosecond crystallography. *J. Appl. Crystallogr.* **45**, 674-678, doi:10.1107/S0021889812024557 (2012).

86 Pergament, M., Palmer, G., Kellert, M., Kruse, K., Wang, J. *et al.* Versatile optical laser system for experiments at the European X-ray free-electron laser facility. *Opt. Express* **24**, 29349-29359, doi:10.1364/OE.24.029349 (2016).

87 Decker, F.-J., Bane, K.L.F., Colocho, W.S., Lutman, A.A. & Sheppard, J.C. in *International Free Electron Laser Conference.* (eds Kip Bishofberger, Bruce Carlsten, & Volker RW Schaa) (JACoW, 2018).

88 Grünbein, M.L., Kovacs, G.N., Kloos, M., Gorel, A., Doak, R.B., Shoeman, R.L., Barends, T.R.M. & Schlichting, I. *Crystallographic studies of rhodopsins –structure and dynamics* in *Rhodopsin: Methods and Protocols* in *Methods in Molecular Biology, Springer Science+Business Media* (ed Valentin Gordeliy) (Humana Press, Springer Protocols, in press).

89 Västberg, A. & Lundborg, B. Signal intensity in the geometrical optics approximation for the magnetized ionosphere. *Radio Sci.* **31**, 1579-1588, doi:10.1029/96RS02630 (1996).

90 Nogly, P., James, D., Wang, D., White, T.A., Zatsepin, N. *et al.* Lipidic cubic phase serial millisecond crystallography using synchrotron radiation. *IUCrJ* **2**, 168-176, doi:10.1107/S2052252514026487 (2015).

91 Wand, A., Friedman, N., Sheves, M. & Ruhman, S. Ultrafast Photochemistry of Light-Adapted and Dark-Adapted Bacteriorhodopsin: Effects of the Initial Retinal Configuration. *J. Phys. Chem. B* **116**, 10444-10452, doi:10.1021/jp2125284 (2012).

92 Scherrer, P., Mathew, M.K., Sperling, W. & Stoeckenius, W. Retinal isomer ratio in dark-adapted purple membrane and bacteriorhodopsin monomers. *Biochemistry* **28**, 829-834, doi:10.1021/bi00428a063 (1989).

93 Stoeckenius, W., Lozier, R.H. & Bogomolni, R.A. Bacteriorhodopsin and the purple membrane of halobacteria. *Biochim. Biophys. Acta - Reviews on Bioenergetics* **505**, 215-278, doi:10.1016/0304-4173(79)90006-5 (1979).

94 Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C. *et al.* A three-dimensional movie of structural changes in bacteriorhodopsin. *Science* **354**, 1552-1557, doi:10.1126/science.aah3497 (2016).

95 Nogly, P., Weinert, T., James, D., Carbajo, S., Ozerov, D. *et al.* Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science* **361**, eaat0094, doi:10.1126/science.aat0094 (2018).

96 Muders, V., Kerruth, S., Lórenz-Fonfría, V.A., Bamann, C., Heberle, J. & Schlesinger, R. Resonance Raman and FTIR spectroscopic characterization of the closed and open states of channelrhodopsin-1. *FEBS Lett.* **588**, 2301-2306, doi:10.1016/j.febslet.2014.05.019 (2014).

97 Schmidt, B., Sobotta, C., Heinz, B., Laimgruber, S., Braun, M. & Gilch, P. Excited-state dynamics of bacteriorhodopsin probed by broadband femtosecond fluorescence spectroscopy. *Biochim. Biophys. Acta - Bioenergetics* **1706**, 165-173, doi:10.1016/j.bbabio.2004.10.008 (2005).

98 Florean, A.C., Cardoza, D., White, J.L., Lanyi, J.K., Sension, R.J. & Bucksbaum, P.H. Control of retinal isomerization in bacteriorhodopsin in the high-intensity regime. *Proc. Natl. Acad. Sci.* **106**, 10896-10900, doi:10.1073/pnas.0904589106 (2009).

99 Prokhorenko, V.I., Halpin, A., Johnson, P.J., Miller, R.J. & Brown, L.S. Coherent control of the isomerization of retinal in bacteriorhodopsin in the high intensity regime. *J. Chem. Phys.* **134**, 085105, doi:10.1063/1.3554743 (2011).

100 Kraack, J.P., Buckup, T. & Motzkus, M. Resonant Two-Photon Excitation Pathways During Retinal-Isomerization in Bacteriorhodopsin. *EPJ Web of Conferences* **41**, 07019 (2013).

101 Arnlund, D., Johansson, L.C., Wickstrand, C., Barty, A., Williams, G.J. *et al.* Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser. *Nat. Methods* **11**, 923, doi:10.1038/nmeth.3067 (2014).

102 Claesson, E., Wahlgren, W.Y., Takala, H., Pandey, S., Castillon, L. *et al.* The primary structural photoresponse of phytochrome proteins captured by a femtosecond X-ray laser. *eLife* **9**, e53514, doi:10.7554/eLife.53514 (2020).

103 Skopintsev, P., Ehrenberg, D., Weinert, T., James, D., Kar, R.K. *et al.* Femtosecond-to-millisecond structural changes in a light-driven sodium pump. *Nature* **583**, 314-318, doi:10.1038/s41586-020-2307-8 (2020).

104 Hutchison, C.D.M., Kaucikas, M., Tenboer, J., Kupitz, C., Moffat, K., Schmidt, M. & van Thor, J.J. Photocycle populations with femtosecond excitation of crystalline photoactive yellow protein. *Chem. Phys. Lett.* **654**, 63-71, doi:10.1016/j.cplett.2016.04.087 (2016).

105 Mayerhöfer, T.G., Mutschke, H. & Popp, J. Employing Theories Far beyond Their Limits—The Case of the (Boguer-) Beer–Lambert Law. *ChemPhysChem* **17**, 1948-1955, doi:10.1002/cphc.201600114 (2016).

106 He, G.S., Qin, H.-Y. & Zheng, Q. Rayleigh, Mie, and Tyndall scatterings of polystyrene microspheres in water: Wavelength, size, and angle dependences. *J. Appl. Phys.* **105**, 023110, doi:10.1063/1.3068473 (2009).

107 Cervelle, B., Cesbron, F., Berthou, J. & Jolles, P. Morphologie et propriétés optiques des cristaux de lysozyme de poule de type quadratique et orthorhombique. *Acta Crystallogr. A* **30**, 645-648, doi:10.1107/s0567739474001550 (1974).

108 Demtröder, W. *Experimentalphysik 2: Elektrizität und Optik*. 6., überarb. u. akt. Aufl. 2013 edn, Online-Ressource (XVI, 482 S. 690 Abb., 300 Abb. in Farbe, digital) (Springer, 2013).

109 Kubo, M., Nango, E., Tono, K., Kimura, T., Owada, S. *et al.* Nanosecond pump-probe device for time-resolved serial femtosecond crystallography developed at SACLA. *J. Synchrotron Radiat.* **24**, 1086-1091, doi:10.1107/S160057751701030X (2017).

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

111 Kohler, D., Seitz, W.L., Loree, T.R. & Gardner, S.D. Speckle reduction in pulsed-laser photographs. *Opt. Commun.* **12**, 24-28, doi:10.1016/0030-4018(74)90065-0 (1974).

112 Manni, J.G. & Goodman, J.W. Versatile method for achieving 1% speckle contrast in large-venue laser projection displays using a stationary multimode optical fiber. *Opt. Express* **20**, 11288-11315, doi:10.1364/OE.20.011288 (2012).

113 Raffel, M., Willert, C.E., Wereley, S. & Kompenhans, J. *Particle Image Velocimetry: A Practical Guide*. XX, 448 (Springer, 2007).

114 Brünger, A.T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475, doi:10.1038/355472a0 (1992).

115 Karplus, P.A. & Diederichs, K. Linking Crystallographic Model and Data Quality. *Science* **336**, 1030-1033, doi:10.1126/science.1218231 (2012).

116 Blaj, G., Liang, M., Aquila, A.L., Willmott, P.R., Koglin, J.E., Sierra, R.G., Robinson, J.S., Boutet, S. & Stan, C.A. Generation of high-intensity ultrasound through shock

propagation in liquid jets. *Phys. Rev. Fluids* **4**, 043401, doi:10.1103/PhysRevFluids.4.043401 (2019).

117 Song, L., El-Sayed, M.A. & Lanyi, J.K. Protein Catalysis of the Retinal Subpicosecond Photoisomerization in the Primary Process of Bacteriorhodopsin Photosynthesis. *Science* **261**, 891, doi:10.1126/science.261.5123.891 (1993).

118 Grünbein, M.L. Illumination conditions in pump-probe experiments, <https://github.com/MGruenbein/pump-probe-experiments> (2020).

119 Wiedorn, M.O., Oberthür, D., Bean, R., Schubert, R., Werner, N. *et al.* Megahertz serial crystallography. *Nat. Commun.* **9**, 4025, doi:10.1038/s41467-018-06156-7 (2018).

120 Yefanov, O., Oberthür, D., Bean, R., Wiedorn, M.O., Knoska, J. *et al.* Evaluation of serial crystallographic structure determination within megahertz pulse trains. *Struct. Dyn.* **6**, 064702, doi:10.1063/1.5124387 (2019).

121 Yamada, H., Nagae, T. & Watanabe, N. High-pressure protein crystallography of hen egg-white lysozyme. *Acta Crystallogr. D* **71**, 742-753, doi:10.1107/S1399004715000292 (2015).

122 Khakhulin, D., Otte, F., Biednov, M., Bömer, C., Choi, T.-K. *et al.* Ultrafast X-ray Photochemistry at European XFEL: Capabilities of the Femtosecond X-ray Experiments (FXE) Instrument. *Appl. Sci.* **10**, 995 (2020).

123 Lemke, H.T., Bressler, C., Chen, L.X., Fritz, D.M., Gaffney, K.J. *et al.* Femtosecond X-ray Absorption Spectroscopy at a Hard X-ray Free Electron Laser: Application to Spin Crossover Dynamics. *J. Phys. Chem. A* **117**, 735-740, doi:10.1021/jp312559h (2013).

124 Levantino, M., Schirò, G., Lemke, H.T., Cottone, G., Glownia, J.M., Zhu, D., Chollet, M., Ihee, H., Cupane, A. & Cammarata, M. Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nat. Commun.* **6**, 6772, doi:10.1038/ncomms7772 (2015).

125 Dumont, C., Emilsson, T. & Gruebele, M. Reaching the protein folding speed limit with large, sub-microsecond pressure jumps. *Nat. Methods* **6**, 515-519, doi:10.1038/nmeth.1336 (2009).

126 Liu, Y., Prigozhin, M.B., Schulten, K. & Gruebele, M. Observation of Complete Pressure-Jump Protein Refolding in Molecular Dynamics Simulation and Experiment. *J. Am. Chem. Soc.* **136**, 4265-4272, doi:10.1021/ja412639u (2014).

127 Piana, S., Sarkar, K., Lindorff-Larsen, K., Guo, M., Gruebele, M. & Shaw, D.E. Computational Design and Experimental Testing of the Fastest-Folding β-Sheet Protein. *J. Mol. Biol.* **405**, 43-48, doi:10.1016/j.jmb.2010.10.023 (2011).

128 Lincoln, C.N., Fitzpatrick, A.E. & Thor, J.J.v. Photoisomerisation quantum yield and non-linear cross-sections with femtosecond excitation of the photoactive yellow protein. *Phys. Chem. Chem. Phys.* **14**, 15752-15764, doi:10.1039/C2CP41718A (2012).

129 van Thor, J.J. Advances and opportunities in ultrafast X-ray crystallography and ultrafast structural optical crystallography of nuclear and electronic protein dynamics. *Struct. Dyn.* **6**, 050901, doi:10.1063/1.5110685 (2019).

Lists

List of Abbreviations

AGIPD	Adaptive Gain Integrating Pixel Detector
bR	Bacteriorhodopsin
CC BY 4.0	Create Commons Attribution 4.0 International Public License
CW	Continuous wave
CXI	Coherent X-ray Imaging
EuXFEL	European XFEL
FEL	Free-electron laser
FFT	Fast Fourier transform
GDVN	Gas-dynamic virtual nozzle
GPCR	G-protein coupled receptor
HEWL	Henn egg-white lysozyme
HPLC	High performance liquid chromatography
HVE	High viscosity extrusion
ID	Inner diameter
LCLS	Linac Coherent Light Source
LCP	Lipidc cubic phase
Mb	Myoglobin
NaCMC	Sodium carboxymethylcellulose
SFX	Serial femtosecond crystallography
SNR	Signal-to-noise ratio
SPB	Single Particles, Clusters and Biomolecules
TR	Time-resolved
UV-vis	Ultraviolet visible
XFEL	X-ray free-electron laser

List of Figures

List of Tables

Acknowledgements

The work presented in this thesis would not have been possible without the support, encouragement, inspiration and contribution of many other people, whom I would like to acknowledge here:

An erster Stelle möchte ich mich bei **Ilme Schlichting** für die außergewöhnliche Möglichkeit bedanken, diese Arbeit in der FEL Gruppe unter ihrer Leitung durchführen zu können. Ihre fortwährende, großzügige Unterstützung geht weit über exzellenten fachlichen Input hinaus. Besonders prägend war für mich die Erfahrung einer Chefin, die jederzeit tatkräftig mitten im Team mitwirkt und sich immer Zeit für alle Anliegen nimmt. Darüber hinaus bin ich außerordentlich dankbar für die Gelegenheiten, durch viele verschiedene Experimente an FELs weltweit sowie durch internationale Konferenzen meinen Horizont wissenschaftlich und kulturell erweitert haben zu dürfen.

Ebenso danke ich **Robert Bittl** herzlich für die Unterstützung meines Dissertationsprojektes am Max-Planck-Institut für medizinische Forschung und für die großzügige Möglichkeit, meine Ergebnisse an der FU Berlin zu vertreten. Die Seminare und Diskussionen in Berlin waren äußerst stimulierend, oft orthogonal zu denjenigen in Heidelberg und ich danke Ihnen und der ganzen Arbeitsgruppe für die einladende Atmosphäre.

I am sincerely grateful to **Bruce Doak** for having introduced me to the FEL group originally for my Master's thesis, and for the rewarding collaboration that developed from thereon. I highly appreciate learning the art of injection from the expert as well as the freedom with which I could perform experiments in the physics lab that came hand in hand with invaluable advice whenever I needed it. I am grateful for having learned from someone inventing and designing solutions with such passion, expertise and extremely high standards, as well as for the many hikes with the two four legged friends through Odenwald.

I sincerely appreciate the always open door of **Bob Shoeman**, who was always ready for discussions and patient explanations on any technique, be it night or day, beginning or end of a shift, at any beam line or place in the world. I am amazed by his detailed knowledge on what feels like everything, and am grateful to have worked in a team with him learning from his very practical approach to solving issues.

I am also grateful to **Claudiu Stan** for the fruitful collaboration on the pressure jump projects, for the discussions and for the careful and critical approach to research and analysis without which the experiments would not have been successful. Also, I am thankful for learning from his imaging capabilities and for taking and providing the high quality jet images.

Furthermore, I want to thank everyone from the **FEL group** whom I have not mentioned so far: **Thomas Barends, Lutz Foucar, Alexander Gorel, Elisabeth Hartmann, Mario Hilpert, Marco Kloos, Gabriela Nass Kovacs, Chris Roome, Miriam Stricker**. Thank you

for all the discussions on science and beyond and for making both everyday life at the institute and the FEL trips an exciting and fun time. In particular: **Thomas** thank you for your crystallographic expertise and structural evaluations; **Lutz** and **Mario** thank you for always ensuring that everything works and that the data is "there"; **Alex** thank you for the insight into computational analyses, the joint investigations and the enjoyable atmosphere in our office; **Elisabeth** thank you for the huge effort of producing all sample which is the foundation of all our work; **Marco**, **Gabriela** and **Miriam** thank you for being the very best injection team and for always being ready to help no matter what; **Chris** thank you for excellent IT solutions. To all of you: thank you for being great colleagues and travel companions, I would always join forces with you again.

None of our FEL experiments would have been possible without the excellent **beam line and accelerator staff** of **European XFEL** and the **LCLS**, all of whom are gratefully acknowledged here.

Discussions with **Stefan Haacke**, **Michel Sliwa** and **Jochen Reinstein** have been illuminating. I very much appreciate their insight and expertise, which were invaluable to investigating how to find appropriate excitation conditions for light-triggered experiments on protein crystals.

Auch der **ganzen Verwaltung des MPIs** bin ich sehr dankbar, die durch ihre Organisationshilfe bei Einkauf, Finanzen und co. eine große Last von den Schultern nehmen und eine noch stärkere Fokussierung auf die Forschung erlauben. Insbesondere bedanke ich mich herzlich bei **Martina Blochmann** für die perfekte Organisation aller Reisen zu Strahlzeiten und Konferenzen. Ohne ihre Hilfe wären wir sicher nicht immer gemeinsam zur selben Zeit am richtigen Ziel angekommen. Gleichfalls danke ich **Ulrike Sanders** für die perfekte und vorausschauende Abrechnung der Reisen.

Ebenso sehr danke ich allen Mitarbeitern der **Feinmechanik- und** der **Elektronikwerkstatt des MPI**, die mir stets mit Rat und Tat zur Seite standen und selbst alle last-minute Sonderwünsche vor Messzeiten fix und ausgezeichnet umgesetzt haben.

I am thankful for all my colleagues generating a great working environment at the institute and in particular within the **BMM department**. Tatiana, Kazu, Rob, Udo, Clara, Mirek, Diana, Akram, Andreas, Sarah, Melanie, Sabine, Kerstin, Egle, Monica, Matthias, Tiia, Melanie, Lukas, Anna, Tadeo, Alexa, Ulli, Alex C., Jolanta, Paola, Victor, Evgeniy, Franziska, Karina, Fiona, Katrin, Thomas H.: it is a pleasure to work and discuss with all of you, be it in seminars, over lunch or during hikes and get-togethers. Also thank you to my **fellow PhDs** from all departments for putting a different perspective to work, inside and outside of the institute.

During my PhD, several experiments and beamtimes were carried out together with the **Grenoble teams** led by Martin Weik and Jacque-Philippe Colletier. To you two and to Giorgio Schirò, Marco Cammarata, Guillaume Tetreau, Kyprianos Hadjidemetriou, Elena An-

dreeva, Nicolas Coquelle thank you for your support with some of the "Heidelberg experiments", for letting me join your "French" experiments and for the welcoming reception in Grenoble.

Meinen **Schul- und Studienfreunden**, allen voran Kerstin und Solvejg, die mich über die Zeit meiner Doktorarbeit begleitet und unterstützt haben, danke ich für den unerschöpflichen Vorrat an Motivation und Ablenkung, jederzeit in der richtigen Dosierung.

Ganz besonderer Dank gilt **meinen lieben Eltern und meinem Bruder Paul** für ihre grenzenlose Unterstützung. Ohne Eure Hilfe, Prägung und Zuwendung wäre ich nie dort angekommen, wo ich heute stehe.

Meinem Mann **Daniel Schwarz** danke ich von ganzem Herzen für den felsenfesten Rückhalt, die rege Diskussionsfreude und das selbstverständliche Begleiten aller Höhen und Tiefen.

Selbstständigkeitserklärung

Name: Grünbein
Vorname: Marie Luise

Ich erkläre gegenüber der Freien Universität Berlin, dass ich die vorliegende Dissertation selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe. Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht. Diese Dissertation wurde in gleicher oder ähnlicher Form noch in keinem früheren Promotionsverfahren eingereicht.

Mit einer Prüfung meiner Arbeit durch ein Plagiatsprüfungsprogramm erkläre ich mich einverstanden.

Datum: _____ Unterschrift: _____