Viktor Ahlberg Gagnér

Electromagnetic Field Induced Out-of-Equilibrium Structural Dynamics in Protein Crystals: From Picoseconds to Milliseconds

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

Department of Chemistry and Molecular Biology
The University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg, 2022

Cover illustration: A model of bovine trypsin, based on PDB entry: 4I8G
Average B-factor differences from paper IV are coloured as $\langle \Delta B \rangle \leq -2\sigma$ (blue), $\langle \Delta B \rangle \approx 0$ (white) and $\langle \Delta B \rangle \geq 2\sigma$ (red), $\sigma = \mathrm{SD}$

Electromagnetic Field Induced Out-of-Equilibrium Structural Dynamics in Protein Crystals: From Picoseconds to Milliseconds

viktor.ahlberg.gagner@gu.se

ISBN: 978-91-8009-650-8 (PRINT)
ISBN: 978-91-8009-651-5 (PDF)
Available online at: http://hdl.handle.net/2077/70273

Department of Chemistry and Molecular Biology
Division of Biochemistry and Structural Biology
University of Gothenburg
SE-413 90 Gothenburg, Sweden
Tel. +46 (0)31-786 0000
inst.kmb@cmb.gu.se

This thesis was written in LaTeX and printed by:
Stema Specialtryck AB
Borås, Sweden, 2022

Abstract

Light-induced interactions in biomolecules are strongly varying over different model systems. Most amino acids and proteins absorb ultraviolet light, but only a few with specialised chromophores or fluorophores, are able to absorb light in the visible spectrum. In the terahertz frequency spectrum, spectroscopy have revealed protein-light interactions which are attributed to low-frequency protein vibrations.

Despite the fact that questions regarding a protein's structure and function are integrally connected, only a small fraction have been answered. With the development of ultra-fast facilities, the possibility to answer these questions at picosecond and sub-picosecond time scales are possible.

The aim of this thesis is to demonstrate results from different aspects of the structural dynamics research field. This thesis show induced structural dynamics in the model protein bovine trypsin, while irradiated by a terahertz electromagnetic field at millisecond and at femtosecond timescales, during two different X-ray crystallography experiments. At millisecond timescales, differences in averages over the model parameters B-factors and anisotropy ('ANISO') reveal structural dynamics which are not attributed to thermal vibrations. A clustering of the individual components of the anisotropic displacement parameter tensor, group atoms which have seemingly no spatial correlation. This indicates long range vibrations, oscillating over the entire protein scaffold. At femtosecond timescales, averages of distances of individual atom positions and individual B-factors, show structural differences which are distributed over the entire protein model, but localised to specific residues, or nearest neighbours.

In addition, the structural dynamics in photosyntetic reaction center protein were demonstrated in a femtosecond optical pump-X-ray probe experiment. Average distance ratios of individual C_α atoms from photo-activated and "dark" datasets were compared. The distances reveal a structural difference upon excitation with infrared light, attributed to an electron charge transfer from the special pair of chlorophyll molecules, to the menaquinone, via the cofactors of the L-subunit. The structural differences are supported by electron difference maps, time-resolved infrared spectroscopy and molecular modelling.

Finally, this thesis demonstrate X-ray data, collected at a commissioning beamtime at the FemtoMAX beamline, of the short-pulse facility at MAX IV. Despite technical and practical difficulties, high resolution X-ray diffraction data were collected with good data reduction and refinement statistics. From this data, a satisfactory protein model was obtained.

Populärvetenskaplig sammanfattning på svenska

Vi är konstant omgivna av ljus. De flesta proteiner kan absorbera och fluoroscera ultraviolett ljus men bara ett fåtal, till exempel rhodopsiner i ögonen, kan absorbera synligt ljus.

Vid längre våglängder, vid gränsen mellan infrarött ljus och mikrovågor, så har studier visat att proteiner kan absorbera så kallad terahertz strålning. Denna absorption leder till långa proteinvibrationer som verkar spela en betydande roll för proteiners funktion. Detta har visats i bland annat äggvite-proteinet Lysozym.

En vanlig metod att undersöka proteinstrukturer är med så kallad röntgenkristallografi. När man bestrålar proteinerna med röntgenstrålningen så bildar det spridda ljuset ett så kallat diffraktionsmönster. Med detta kan man återskapa den specifika proteinstrukturern via datorprogram. Denna metod kräver ett protein som kan kristalliseras, vilket alla proteiner inte gör. Eftersom proteiner överlag är flexibla, även i kristallform, så är det vanligt att defekter skapas när proteinerna kristalliseras eller fryses ned. En vanlig orsak är smuts i proteinlösningen. Dessa defekter bidrar till en diffus bakgrundsspridning av röntgenljuset.

En annan orsak till en begränsad spridning är vibrationer, bland annat från värme. Med hjälp av dagens analysprogram så kan den begränsade spridningen beräknas med hjälp av en så kallad B-faktor parameter. Varje atom i en proteinmodell associeras med en specifik B-faktor. Genom att beräkna skillnader i bland annat B-faktorer, och skillnader i atompositioner, så kan strukturella skillnader förklaras med icke-termiska proteinvibrationer. I denna avhandling görs detta.

Med hjälp av terahertz strålningen så har proteinvibrationer demonstrerats i proteinet bovint trypsin. Detta fenomen är svårförklarat, men utan koppling till en ökad temperatur. En möjlig förklaring är att proteinvibrationerna motsvarar en kollektiv vågrörelse som berör hela strukturen samtidigt.

Utöver detta så har strukturen undersökts med ultrasnabb tidsupplöst kristallografi. Detta möjliggör en detaljerad beskrivning av proteinvibrationera. Denna typ av snabba experiment har även använts för att undersöka strukturförändringarna i ett reaktionscenter-protein, som bland annat finns i kloroplasterna hos växtceller.

Modellerna från de snabba förloppen är viktiga i forskning där proteinstruktur och dynamik möts. Vet man till exempel strukturen och strukturförändringarna, så kan man bekräfta de biokemiska reaktioner som sker i proteiner.

List of publications

This thesis is based on the following publications:

Paper I: **Gagnér, V.A.**, Lundholm, I., Garcia-Bonete, MJ. et al. Clustering of atomic displacement parameters in bovine trypsin reveals a distributed lattice of atoms with shared chemical properties. Sci Rep 9, 19281 (2019). https://doi.org/10.1038/s41598-019-55777-5

Paper II: Jensen, M.*, **Ahlberg Gagner, V.***, Cabello Sanchez, J. et al. High-resolution macromolecular crystallography at the FemtoMAX beamline with time-over-threshold photon detection. (2021). J. Synchrotron Rad. 28, 64-70

Paper III: Dods, R., Båth, P., Morozov, D., **Gagnér V.A.** et al. Ultrafast structural changes within a photosynthetic reaction centre. Nature 589, 310–314 (2021). https://doi.org/10.1038/s41586-020-3000-7

Paper IV: **Ahlberg Gagnér, V.***, Jensen, M.*, Cabello Sanchez, J. et al. Ultrafast structural response of a protein crystal to a strong pulsed THz field
Manuscript

* These authors contributed equally

Related papers which I co-authored but is not included in the thesis:

Paper V: **Ahlberg Gagnér, V.**, Jensen, M., & Katona, G. Estimating the probability of coincidental similarity between atomic displacement parameters with machine learning. (2021). Machine Learning: Science and Technology, 2.

Contribution report

Paper I: I performed the structural analysis of the X-ray diffracted data. I co-authored the paper together with my supervisor and our collaborators.

Paper II: I participated in the experimental planning and crystallised the protein together with Jensen M. I participated in the editing of the manuscript.

Paper III: I performed the resampling of the 2016 datasets, and I performed the structural analysis of all the resampled data. I co-created the relevant figures with Dr Båth, and I co-authored the relevant sections of the paper together with Prof. Neutze.

Paper IV: I participated in the experimental planning and crystallised the protein together with Jensen M., I analysed the X-ray diffraction data. I co-authored the manuscript with my supervisor and our collaborators.

Contents

Chapter 1

Introduction

1.1 Electromagnetic interactions with biomacromolecules

Proteins interact with light in various ways. A relatively strong interaction occur in biomolecules, such as aromatic amino acids, or proteins with photosensitive molecules (chromophores or fluorophores) in their active site [1, 2]. The photoactivated moieties and molecules contain conjugated π bonds which form transitions from the ground state to an anti-bonding excited state by $\pi \rightarrow \pi^*$. In addition, transitions from non-bonded electrons (labelled n) of the same molecules are possible: $n \rightarrow \pi^*$, however the probability is small [1, 3]. Fluorescence anisotropy, polarised infrared (IR) spectroscopy and linear and circular dichroism have shown that the strength of the interaction with polarized light can be modulated by the relative orientation of the transition dipole moment, in the case of a single chromophore, or the transition dipole moments of the donor and the acceptor molecules, in a conjugated system [4, 5, 1].

Considering only the ultraviolet (UV) - IR region of the spectrum, incoherent and polychromatic light sources (for instance halogen light bulbs), or continuous wave lasers, induce linear excitations of electronic, rotational or vibrational transitions, from usually a singlet ground state (S0) in room temperature, to either an excited singlet state (S1, S2 etc) with an electronic transition, to a higher vibrational or rovibrational excited state (both rotations and vibrations are considered) with an infrared transition, or combinations thereof [1, 6]. In the far-infrared (FIR) or terahertz (THz) spectrum, low frequency vibrations can be detected by FIR or THz spectroscopy [6].

In most proteins, aromatic residues exhibit an electronic absorbance in the UV spectrum ($\approx$ 260-300 nm), but are optically "silent" at higher frequencies, unless a protein naturally contain a photosensitive molecule [7, 2]. However, a recent experiment exhibited an albeit decaying but significant absorbance at higher wavelengths in the protein α_3C. The protein contained no aromatic residues. The tail in the spectrum is partly attributed to transitions in charged residues (Lys, Glu, Arg, Asp, and His), and to protein folding [2].

THz electromagnetic fields excite low frequency, long range vibrations, in secondary structures and subdomains of proteins [8], as torsions and librations of the backbone, and consequently, torsions and libration of side chains [9]. In tryptophan residues, THz-time domain spectroscopy (TDS) revealed a signal, attributed to bending, deformation, and torsional motion of methyl groups [10].

Furthermore, THz spectroscopy have linked a potential functional importance to the induced vibrations, such as enhancing the reaction rate of myoglobin [11], and blueshifting and increasing in intensity upon NAG3 inhibitor binding in lysozyme [9].

A key issue in THz spectroscopy is a substantial band broadening due to strong water absorption [12, 13]. Havernith's group and Markelz's group independently developed methods to mitigate the water absorption contribution from THz spectrum. Using a THz spectroscopy set up, combined with stopped-flow mixing, Dielmann-Gessner et al. determined the solvent dynamics of peptide hydrolysis in human membrane type-

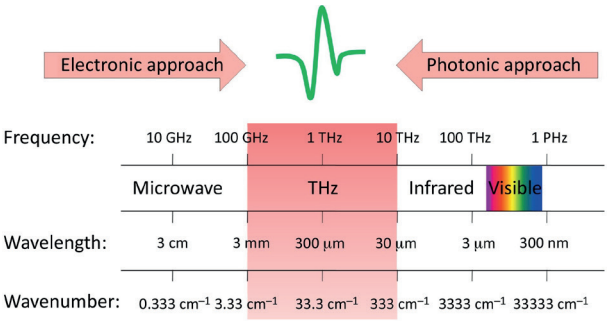


Figure 1.1: Displays the THz frequency spectrum. At this frequency, the THz signal generation from electronic wave guides meets the optics, used in the far infrared spectrum. The figure (adapted) was reprinted from [12] with permission of Elsevier.

1 matrix metalloproteinase (MT1-MMP), complimented by structural information from X-ray Absorption Spectroscopy (XAS) and Extended X-ray Absorption Fine Structure (EXAFS) [14]. Niessen et al. isolated induced THz vibrations in Chicken egg white lysozyme (CEWL) by calculating difference signals from an rotating protein crystal, with respect to a static reference signal [15].

1.2 Structural and dynamical contributions to lattice disorder

Despite the beautiful exterior, no crystal in the natural world is structurally an ideal system. A protein crystal is intrinsically dynamical ([16, 17]), and many contributions to disorder (defects) in the crystal lattice exist.

Currently, no universal theory exists which describe the nucleation and crystal growth processes in protein crystallisation [18, 19, 20]. As a first step, a real-time study of the crystallisation of bovine β-lactoglobulin attributed crystal growth to an increased fraction of immobilised protein, and a larger diffusion coefficient of

non-crystallised monomer [21].

The source of the defects vary. Most occur during crystal formation, and are attributed to impurities in the crystal growth media [22, 23]. Induced defects in lattice domains of flash-freezed crystals have been demonstrated as well [24]. In addition, static disorder include contributions from stochastic processes inside the crystal, such as small deviations from ideal periodicity (not including the defects themselves), orientational disorder [25] and conformational variation [26]. In addition, an inadequate modelling of atomic positions induces an "apparent" displacement, affecting the system [25].

In crystallographic models, the effect of disorder is parametrised by the Debye-Waller factor (DWF), which define in reciprocal space, the probability density function of atomic displacement. Assuming a trivariate Gaussian distribution, the characteristic function of the DWF, define an expression for an isotropic or anisotropic B-factor [25]. The B-factor describe the degree of attenuation of the X-ray scattering from atoms due to the atom displacement, caused by the disorder [27]. Separating the contribution of static disorder, and the so called dynamic disorder, from the B-factors, an expression for the average displacements can be described as:

$$\langle {u_i}^2 \rangle = \frac{3}{8\pi} B_i = \langle u_i \rangle_d^2 + \langle u_i \rangle_{ld}^2 \tag{1.1}$$

where the subscripts d and ld denote dynamics and lattide disorder, respectively. The dynamic disorder parameter includes contributions from thermal vibrations and phonon vibrations [26]. This expression assumes an isotropic B-factor but can easily be extended to the anisotropic case as well [25].

In practice, the attenuation induces a so called diffuse scattering contribution to the background of X-ray diffraction images, in addition to other scattering sources, such as Compton scattering and bulk solvent scattering [28]. However, researchers have developed model refinement software (such as REFMAC5 [29] or Phenix [30] of the CCP4 software suite [31]) which includes the B-factor contributions to improve protein models during refinement. Although, the the disorder is a complex parameter and includes B-factor contributions to other modelling parameters

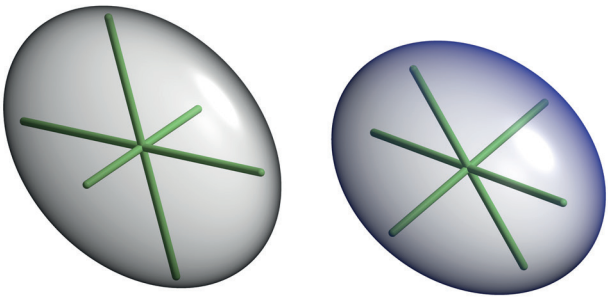


Figure 1.2: The figure displays a graphical representation of the anisotropic displacement parameter tensor as two "thermal" ellipsoids for a carbonyl carbon (gray) and a amide nitrogen (blue). The figure was reprinted from **paper I**.

as well, such as scaling parameters and solvent contributions [32] to structure factor calculations, and are not specific to atom displacement of the protein. Nevertheless, researchers have developed experimental methods to isolate the vibrational contribution of atoms, and thus obtained dynamical molecular models [33, 34, 35].

In contrast to the previous examples, [36] have indicated, and [37] have demonstrated a non-thermal contribution of phonon vibrations, induced by a weak microwave and THz electromagnetic field, respectively. The results were partly demonstrated by differences in modelled isotropic B-factors for individual atoms.

Chapter 2

Motivations and Aims

Structural dynamics is an important bridge between the fundamental research questions in biochemistry: what is a proteins structure and what is it's function? Since the dawn of structural biology, scientists have and are researching the necessary experimental tools to investigate these questions.

The scope of this thesis is to describe different aspects in the structural dynamics research field. The first aspect involves the research of low frequency protein vibrations. As described in section 1.1, contemporary research have deduced the vibrations as functionally important and could be induced by THz electromagnetic fields. However, some questions arise: in which way are the vibrations affecting the proteins? Can they be detected in a protein model?

- The first aim addressed the stated research questions. The structural differences of bovine trypsin were investigated, while being irradiated with a THz electromagnetic field. The results are presented in paper **I** and **paper IV** at millisecond (ms) and at picosecond (ps) time scale, respectively.

A second aspect involves the collection of time-resolved data at a large scale facility. Recently, The femtosecond X-ray beamline at the MAX IV short-pulse facility opened for commissioning experiments for protein crystallography users [38].

- The second aim addressed if FemtoMAX was a suitable beamline for femtosecond macromolecule X-ray diffraction experiments. By utilising bovine trypsin as a model protein, **paper II** describe the data collection methodology and results from the first protein crystallography experiment at the beamline.

A final aspect involves the processing and modelling of time-resolved X-ray crystallography data, collected at ultrafast beamlines. A characteristic of such data is a high multiplicity of the X-ray diffraction images. Which can be exploited in quantifying uncertainty and describing structural variation.

- The third aim was to investigate if structurally relevant information could be extracted by utilising two different resampling strategies: from the bootstrapping of individual Bragg reflections (addressed in **paper III**), and from resampling individual pairs odd and even diffraction images without replacement (addressed in **paper IV**).

Chapter 3

Methodology

Each paper describe different aspects of measuring time resolved structural dynamics. Therefore, for a complete description of the methodologies involved, the methods section in each paper should be read. However, with the section below, I will summarise the crystallisation part which we performed for every experiment.

3.1 Protein crystallisation

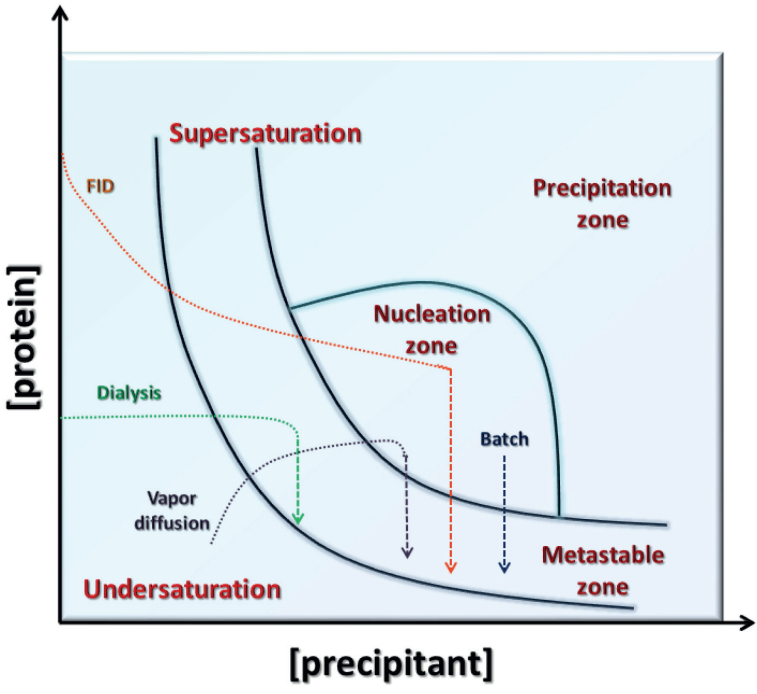
In this thesis, the diffraction pattern from bovine trypsin crystals and photosynthetic reaction center crystals were collected and analysed. In nature, only a few of protein crystallise in its own set of conditions (or sets of conditions for different varieties of crystals). Researching and developing crystallisation conditions for novel crystal structures is an extremely demanding and time consuming process [39]. However, some key concepts do exist.

The crystallisation process can be described as a phase transition which occur in

two steps: (1) nucleation, (2) crystal growth. To initiate a nucleation process, the protein solute must enter the nucleation zone in the supersaturated state. At a certain protein concentration, the nucleus leaves the nucleation zone and enters a metastable state, where only continued growth occurs up to a certain size (figure 3.1) [40].

The transition from a dissolved undersaturated protein to a supersaturated crystal nucleus can occur in multiple ways. Commonly, a diffusion type of experiment control the protein concentration or the crystallisation condition. Some examples include: i) vapour diffusion, ii) dialysis and iii) Free Interface Diffusion (FID) [40]. In addition, by controlling an environmental parameter, for instance temperature, the diffusion can be manipulated, effectively tuning the nucleation or crystal growth rate [40, 41]. Alternatively, a batch method iv) involves crystallisation in a set of conditions which initially are optimised. Often, the the solvent and the protein solution are encased in oil, thus preventing any interaction with the environment [42, 39].

Of the mentioned crystallization techniques, vapour diffusion is by far the most popular for single crystal experiments. However, due to the easy way to scale up crystallization with robots, the batch crystallisation method have proven useful for serial crystallography experiments [43].



[htbp]

Figure 3.1: Shows a simplified phase transition diagram with the popular methods included. The figure was reprinted from Krauss et al. [40]

Chapter 4

Results and discussions

4.1 Millisecond out-of-equilibrium dynamics, induced by a weak terahertz electromagnetic field

In **paper I**, we expanded upon the previous research by Lundholm et al., where structural dynamics were demonstrated in a hen egg white lysosyme, during irradiation of a THz electromagnetic field. In **paper I**, structural differences were investigated between a THz irradiated state and a non-irradiated state, in the crystallised model protein bovine trypsin. In this research, structural perturbations were induced by a millisesecond, weak (1 mW) THz electromagnetic field.

With the model parameter anisotropy (labelled as 'ANISO') an induced THz-related difference were evident when the parameter were averaged over the atoms in individual residues, and over atoms, grouped by different residue type labels. The differences were calculated by separating odd and even diffraction images and

treating them as individual datasets. In addition, a potential THz-related effect were evident, by a general reduction of average B factors from odd diffraction datasets.

In addition, it was shown that components of the tensor itself provided enough structural information to compliment the results from the previous parameters. This research provided the groundwork for future research at the sub-ps X-ray source FemtoMAX (**paper II** and **paper IV**). In addition, with the use of an hierarchical clustering method, it demonstrated the use of machine learning to connect structural information. The results of the machine learning was described in Gagnér et al. where the structural information of ADP's were learned and could be predicted with the numpy package pymc3 (not included in this thesis) [44].

4.1.1 Experimental methodology to capture protein structural dynamics

A 0.5 THz, 1 mW alternating electromagnetic field irradiated the crystals for 23.5 milliseconds (ms), during a period of a 22 ms data collection, and a 3 ms detector read out, per frame, while operating at a 50 % duty cycle (the crystals were irradiated by the THz radiation during the collection of the odd frames). The dynamical THz irradiation method ensured an approximate equal start temperature, and thermal heating and cooling profile.

The data was collected by the oscillation method over an oscillation angle of 0.01° per frame, over a total of 360°. This high sampled phi-slicing methodology enabled us to separate and reconstruct the rocking curves of the Bragg reflections, in the odd and the even data, independently. Thus, structural differences could be compared with similar experimental errors from the resulting odd and even datasets.

Comparing the discrepancy of reflection intensities over separated and independently merged odd and even datasets (R_{split}), is a standard practice with Serial Femtosecond Crystallography (SFX) analysis software [45, 46], and have proven

useful to assess Serial Synchrotron Crystallograhpy (SSX) with SFX data [47]. In addition, the self-referencing data analysis strategy was a integral component in Lundholm et al. to extract the THz-related structural differences [37].

4.1.2 Crystallographic model parameters as structural dynamics analysis tools

The bovine trypsin crystals diffracted up to 1.15 Å, thus enabling an analysis scheme based on differences in model parameters, calculated from the anisotropic displacement parameter tensor, for individual atoms. The THz induced vibrations were investigated by the differences in anisotropic equivalent B-factors (B_{eq}) and the anisotropy parameter (henceforth labelled 'ANISO') . In addition, the individual components of the anisotropic displacement parameter tensors were hierarchy clustered.

As opposed to describing phonon vibrations by thermal diffuse scattering [35], the model parameters directly transforms the vibrational contribution of the lattice disorder to structural information. Since multiple contributions affects disorder and thereby ADPs, a stable and well diffracting is model protein is essential to isolate the vibrational contribution of the ADP tensor.

An initial analysis of average difference B_{eq}, or average difference ANISO, calculated over atoms with different residue type labels, in pairs of odd - even frames, show experimental or modelling errors which affect the average of B_{eq}s over glycine, glutamate, serine and leucine atoms (figure 4.1a), and the average ANISO over atoms of most residue types (except threonines, alanines, prolines, glutamates, histidines and methionines, figure 4.1b) . However, for most difference ANISO averages (and the average difference B_{eq} over glutamate atoms), the Thz-irradiated data is separated from the reference and non-zero.

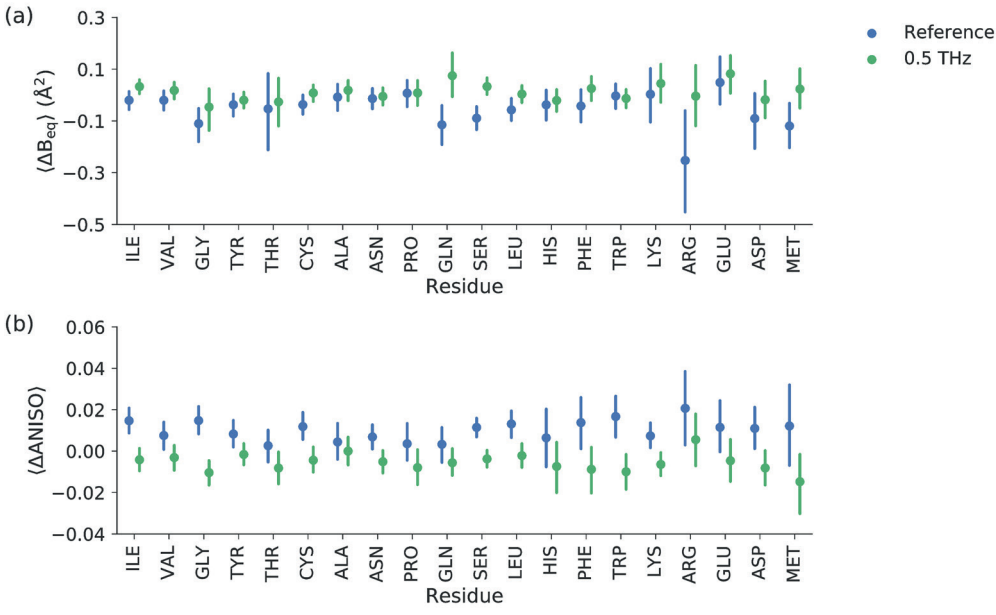


Figure 4.1: A figure of the average difference B$_e q$-factor a) and average difference ANISO b) over atoms with different residue type labels. The error bars are described with a 95 % confidence interval (CI). The figure is reprinted from **paper I**

B_{eq} factors

The analysis of the average ΔB_{eq}s over atoms with specific residue types showed either no difference, or a similar difference compared to reference data. However, an analysis of the absolute level of the B_{eq}s showed a overlapping but general reduction of averages over most residue labels (figure 4.2a) and clear reduction of average B_{eq}s over residue types (figure 4.2c). These results are consistent with the data reduction statistics, where the refined average B-factors are smaller for the THz sets, compared to the references.

Four consecutive Molecular Dynamics (MD) simulations with 10 ns temporal resolution, and one Normal Mode Analysis (NMA) simulation over all normal modes, reproduced the average B_{eq}s (and ANISO) results with a limited success (figure 4.2). Calculated Pearson correlations of the odd versus the even reference data showed correlations of 0.99 and 0.98 of the B_{eq}s and ANISO, respectively. Therefore, it appears unrealistic to assume that the simulations would be able to reproduce effects of a THz electric field. The focus of the simulations was to reproduce the experimental results of non-perturbed systems.

The MD simulations partially reproduced the average B_{eq} in figure 4.2a, but overestimated the results of disordered residues (large modelled CI), and underestimated the results of the well-ordered structure. Figure 4.2c show a similar trends, where Tyr, Asn, Gln, Phe and Lys atoms greatly overestimated the experimental data, but only Ile, Cys, Ala, Pro, Trp and Met slightly underestimated the data. In a general case, the MD simulations showed a higher probability to model equal or higher B_{eq}s, compared to the experimental data. Therefore, a general reduction of the average B_{eq}s due to THz-irradiation seems not unrealistic. However, the MD simulations shared only a 0.61 - 0.33 Pearson correlation with the reference crystals.

In comparison, the B_{eq}s from NMA simulations were two orders of magniutude lower than the experimental and MD simulated data and did not provide any information to the question.

In the supporting information to **paper I**, we address the issues of the simulations. For example, MD simulations lack the functionality to follow the anharmonicity of low frequency vibrational modes, which the B factors of the experimental data partially reflects. In the case of NMA simulations, the eigenvectors from collective vibrations are often need to be heavily extrapolated to accurately describe conformational changes.

Without the self-referencing system, attributing a specific cause to the reduced B_{eq}s is difficult, but providing estimates B-factors from potential perturbing sources is possible. Two examples include perturbations due to radiation damage and due to temperature differences.

We begin by estimating the B-factor differences due to radiation damage. An estimation of the radiation damage, as an weighed average over the number of diffraction images, generated the doses 0.36 ± 0.04 MGy and 0.32 ± 0.04 MGy, for the reference and THz-irradiated datasets, respectively. Thus, the dose difference between the sets were calculated as 0.04 ± 0.11 MGy. For small X-ray doses, Shimizu et al. showed that the increase in B-factors is approximately linear [48]. Therefore, corresponding B-factor increase becomes 0.12 ± 0.34 Å^2 between the sets. The difference in average B-factors, obtained from the model refinement of each respective set, was calculated as ≈1.08 ± 0.56 Å^2, thus the difference of average B_{eq}s between the sets are not likely caused by radiation damage alone.

The THz source share traits with microwave radiation, where water molecules are strong absorbers and start to vibrate (section 1.1). Therefore, the THz radiation produces a thermal response during irradiation. If a protein crystal is modelled as ice of an approximate equivalent size as the proteins (300 μm × 300 μm × 300 μm), and we assumed the THz radiation only produced a pure thermal response with final state in a thermal equilibrium, a simple method to estimate the increase in B-factors due to heating is by the equipartition theorem. The equation have been used to successfully model protein dynamics using elastic neutron scattering [33]. With two squared position coordinates (both potential and kinetic energy are

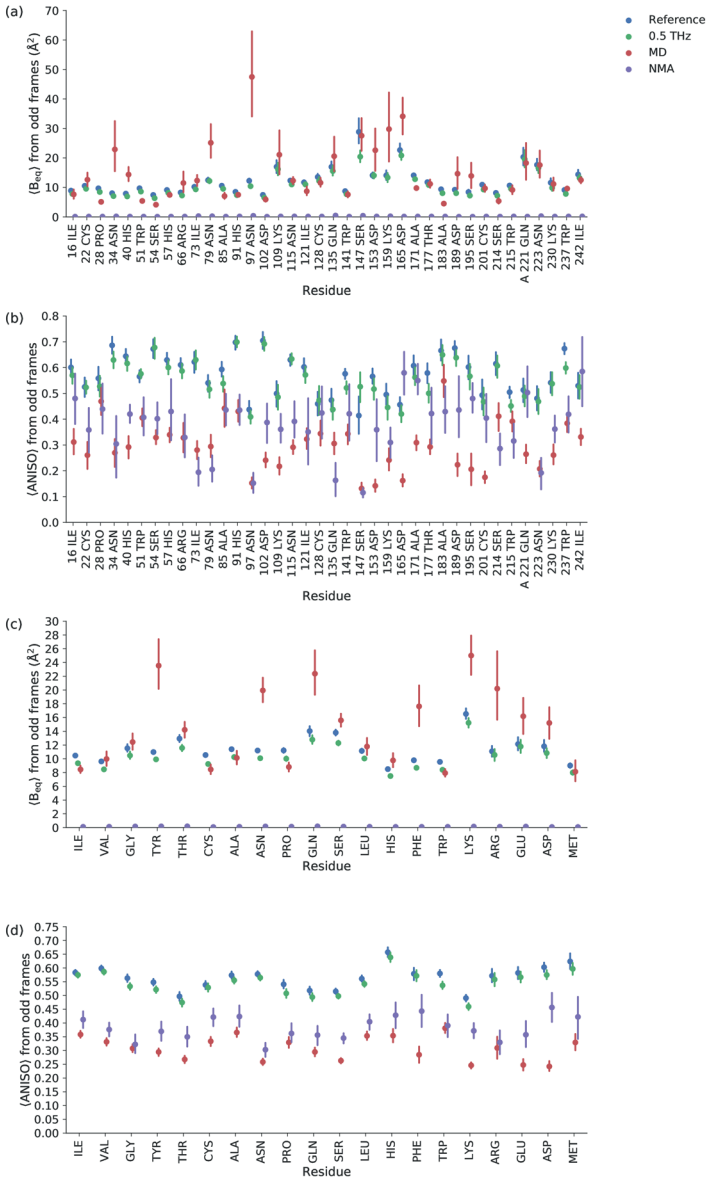


Figure 4.2: a) The average B_{eq} over atoms in odd datasets with specific residue labels b) The average ANISO parameter over atoms in odd datasets with specific residue labels, c) The average B_{eq} over atoms in odd datasets with specific residue types, c) The average ANSIO over atoms in odd datasets with specific residue types. The error bars are described with a 95 % confidence interval (CI). The figure is reprinted from **paper I**

included) the equation becomes [49]):

$$k_B T = \frac{1}{2} k \langle x^2 \rangle \tag{4.1}$$

where k_B describes the Boltzmann constant, T denote the temperature (in Kelvin), and k is a spring or force constant. The model assumed all average energy to be expressed as potential energy. For a system of $3N_A$ harmonic oscillators in a solid (N_A corresponds to the Avogadro constant), the expression for thermal energy becomes $3N_A k_B T$. Using COMSOL Multiphysics® [50], the radiation was simulated to increase the temperature by 54 μK. The force constant can be measured via various means [33, 51]. Accodring to anisotropic network analysis, AFM and neutron scattering, the force constant for proteins is measured and simulated as $\approx$ 0.1 - 1 N/m. By using the definition of B-factors in the isotropic case equation [25], the estimated temperature change gives a rise of about 10^{-4} Å^2 - 10^{-5} Å^2 of the B-factors. Hence, the change is not likely related to temperature differences.

In addition to providing estimates for changes in B-factor, we can also qualitatively estimate the cause of the difference, based on experimental parameters, modelling parameters or crystal production. Trivial causes include differences due to the independent modelling of the datasets, by non-isomorphism between the crystals and by the fact of the inconsistent removal of shadows between each crystal. By judging from the similar mean R-free values and their corresponding standard errors, and the fact that the same initial model was used for the refinement, a difference due to modelling differences seems unlikely. Since the self-referencing system show a large reduction in average B_{eq}s of glutamate and serine atoms of the reference data, this indicates that there are differences in the datasets which are related to crystal inhomogeneities, perhaps caused by different conditions during flash freezing or a inhomogeneous temperature profile of the cryostat, or systematic errors in the X-ray flux or beam profile during data collection. However, these systematic errors are unlikely, again, due to the similar data reduction statistics. Despite this fact of potential systematic, or crystallographic differences could cause the differences, a decrease in isotropic B-factors was demonstrated in Lundholm et

al. [37] and in Weissenborn et al. [36], or in systems where EM radiation induces higher order states. At sub-ps timescales, theory and experimental observations (**paper IV**) predict and show a local rearrangement of isotropic B-factors on the protein model, where certain atoms gain positive ΔB_{iso} and certain atoms gain negative ΔB_{iso}. This indicates that a general trend is difficult to predict and much depend on either local (or global) atom arrangement or individual atom characteristics. However, the oscillation is expected to exponentially decrease until a global minima is reached.

Anisotropy

Average ΔANISOs over atoms with different residue types (figure 4.1b) showed a separated overlap of confidence intervals between reference and THz-irradiated crystals, where the largest differences are measured in Ile, Gly, Asn, Cys, Ser, Leu, Trp, Lys and Asp residues and the averages of Gly, Trp, Lys and Asp are seperated from 0. Of the specified residue types, the ΔANISO decreased in the odd frames upon THz irradiation. Hence the atoms became more anisotropic.

The location of the amino acids in the structure seem to have an importance. The average ΔANISO over atoms in all Gly, Trp and His residues reveal a tendency where atoms in residues close to the N-terminus become more anisotropic (figure 4.3). None of the selected residues were modelled with alternative conformations. The trend is highlighted with Gly-43, Gly-78, Trp-51 and the catalytic residue His-57.

The trend of a reduction of the ANISO parameter upon THz irradiation is strengthened by the comparison of average ANISO over atoms of different residues and residue types (figure 4.2b,d). Similar to the average ΔANISOs, the average ANISOs in figure 4.2d show a higher anisotropy (a lower ANISO) in all types of residues, but in tryptophanes in particular. Similarly, to figure 4.3, figure 4.2b reveal a general reduction of the average ANISO, but the amount fluctuates strongly, depending on which residue atoms that are averaged over, with either overlapping, limited

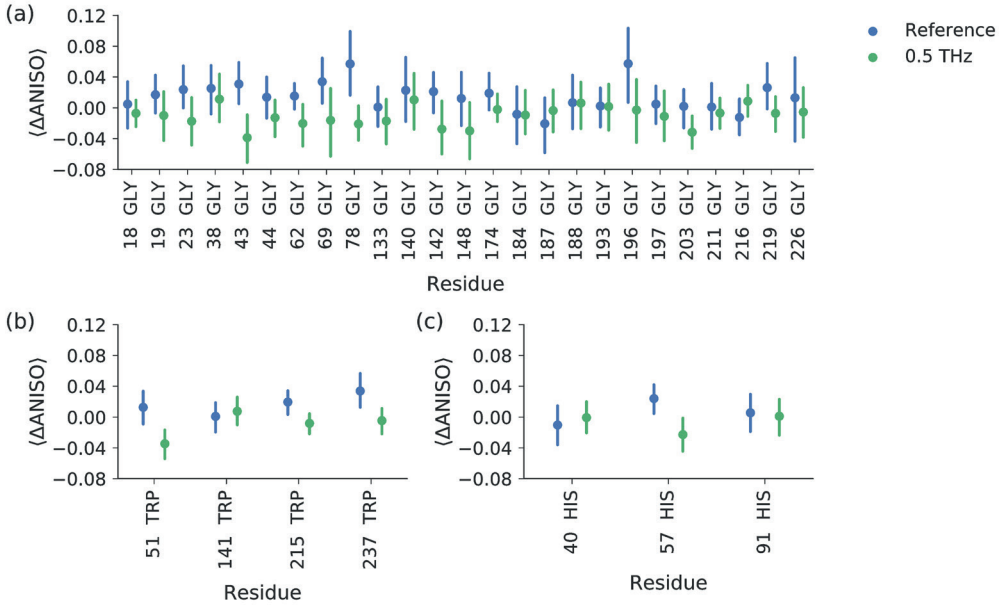


Figure 4.3: The average difference ANISO over atoms in selected residues. a) Gly, b) Trp, c) His. The error bars are described with a 95 % confidence interval (CI). The figure is reprinted from **paper I**

overlapping (in Asn-34, Ala-85 and Thr-177) or separated confidence intervals (in Trp-141, Trp-213 and Trp-237). For the average of Ser-147, an opposite relation is evident (the average is less anisotropic compared to the reference data).

Anisotropic displacement parameters

Given the evidence of a seemingly higher anisotropy in THz-irradiated crystals, a clustering of the ADP tensor components of the individual atoms revealed a previously unknown structural connection between atoms which share similar ADPs. For instance, of THz-irradiated data, the clustering grouped a C_β, a C_α and a $C_{\epsilon 2}$ atom of Trp 51, a O and a $C_{\zeta 2}$ of Trp 237, and three $C_{\epsilon 1}$ atoms of His 57, in a total of ≈15900 atoms. In addition, the atoms were well separated in the protein model, and shared no apparent physical or chemical interaction (figure 4.4).

In another representation, the ADP tensors can be visualised as 'thermal' ellipsoids [25, 52], with the B_eq-factors, defining the radii of the ellipsoids, and the eigenvectors of the ADPs as the directions of the ellipsoid's semi-axii. Hence, the clustering showed THz-irradiated thermal ellipsoids of which are connected by a similar orientation, shape and size.

This information show a mosaic of atoms, which are, or become connected upon THz-irradiation. In **paper I**, we propose a theory, arguing that an intrinsic property of the atoms (for instance partial charge) is an important component in allosteric interactions of residues, and the atoms' response to the electromagnetic field. The teory is partially supported by the results of Hekstra et al. [53]. They showed that partial or formal charges can be utilised to structurally manipulate protein with no known voltage dependence in strong electric fields. In Ahlberg Gagnér et al., we demonstrated an example how the information, provided from the ADP tensors, could be utilised for structural prediction in an machine learning context [44]. Future research would include mutation studies of selected clustered residues, to determine the implication of structural modification on the potential information transfer.

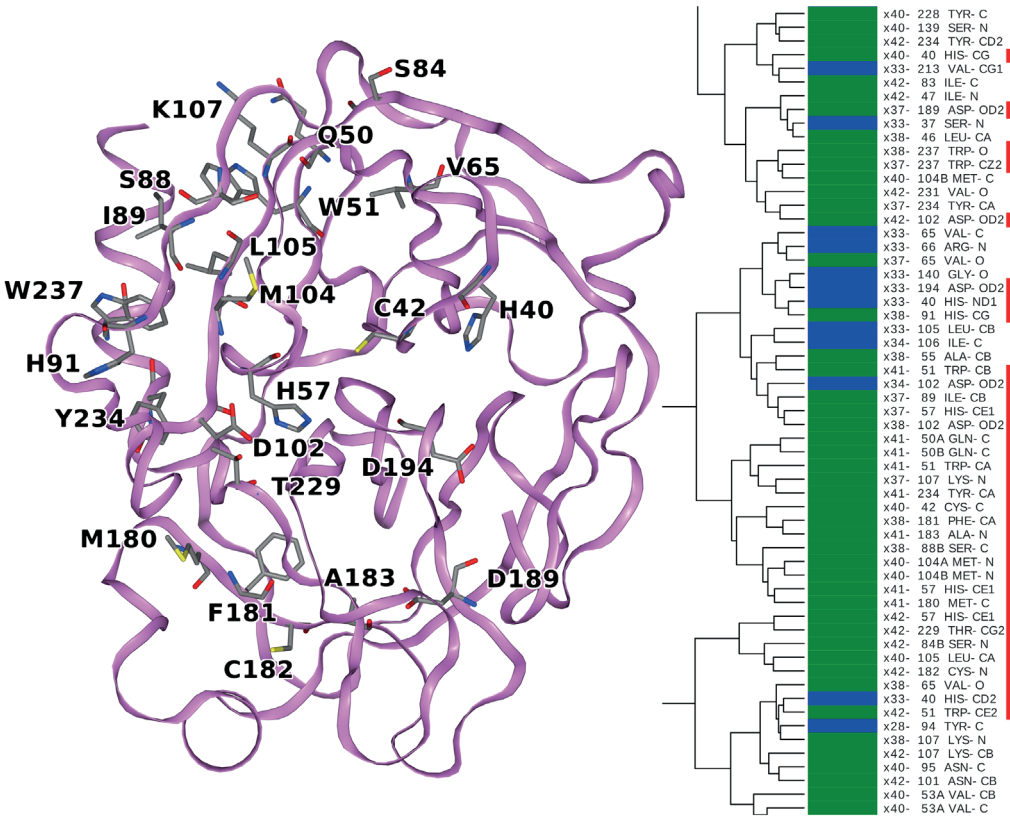


Figure 4.4: A cluster of structurally important atoms which have been radiated with THz irradiaton (red marker). The figure is reprinted from **paper I**

4.2 FemtoMAX: a state-of-the-art beamline for detecting fast structural dynamics in proteins

Paper II addresses the data collection and processing, during a commissioning experiment at the FemtoMAX beamline at Max IV. The data was collected with a combination of the Laue (stationary) method, and the oscillation method, but is evaluated by considering either one, or multiple diffraction images per rotation step. Both methods are compared by assessing the collected diffraction patterns, the data collection and refinement statistics, and the resulting protein models.

This research describe the potential of using the FemtoMAX beamline in a Macromolecule Crystallography (MX) context. Historically, this beamline have demonstrated its use in materials science, of which time resolved processes, such as laser heating [54], photoacoustic activated transducing [55] and radiative relaxation [56], can be excited with either a optical laser pulse or the X-ray radiation. However, due to a low X-ray flux at the required X-ray energies from a previous undulator [38], the beamline haven't been operational in the context of macromolecular crystallography. In this research, a fully functional MX experiment is demonstrated for the first time at the beamline. This research opens up possibilities for picosecond (ps) time resolved MX experiments, of which the data could rival established MX and Serial Femtosecond Crystallography (SFX) beamlines.

Future plans for FemtoMax include upgrading to a 100 Hz repetition rate, thus making it comparable to SFX beamlines at SACLA and LCLS.

4.2.1 Data collection

The diffraction pattern of crystallised bovine trypsin was collected in room temperature, using a 130 μm $\times$ 160 μm (horizontal and vertical full widths at half maximum) X-ray beam, operating with a 2 Hz repetition rate and a 150 fs pulse duration. The data was collected by rotating the cystal in steps of 0.1°, over a total

of 128.3°. A multilayer monochromator selected photons at 11.15 keV (1.112 Å) at a flux of 2×10^7 photons per second. The low photon flux was compensated by collecting 100 diffraction images (corresponding to 100 X-ray pulses) per rotation step. Its relatively wide bandwidth ($\Delta E/E \approx 0.01$; typical values are in the order of 10^{-4} - 10^{-5} for crystal monochromators [57]) ensured the possibility of collecting the diffraction pattern from the still crystal orientations.

Due to the ultrafast data collection timescale per frame, the estimated levels of total maximum absorbed X-ray dose, and average absorbed X-ray dose, during the entire (18 h) data collection, was 27 kGy and 7 kGy, respectively [58]. In comparison, the Garman limit of absorbed X-ray dose in room temperature is reported as 100 – 400 kGy [59].

The collection (and subsequent analysis) of the diffraction data provided many technical and practical challenges. For instance, during refill of the MAXIV storage ring, the X-rays from linear accelerator dropped at infrequent intervals, during the long data collection. Therefore, the collection stopped and had to be resumed, resulting multiple frames with no diffraction patterns. The problem was mitigated by summing the photon count in each individual pixel over the 100 the diffraction images.

To prevent the crystal from drying out during the long data collections, the crystals were sealed in MiTeGen plastic capillaries. The plastic material was chosen over amorphous or quartz glass to not interfere with the THz pump pulse, described in **paper IV**. A 1 µl mother liquor at the end of the capillary provided a sufficiently constant vapor pressure over the crystal. The capillaries were sealed with vacuum grease at the goniometer base.

An inherent limitation with hybrid photon counting detectors is the maximum achievable count rate (10^7 per second and pixel for Pilatus3 and Eiger detectors [60, 61]. At a typical MX beamline, the average X-ray flux is in the order of 10^{12} photons/s [58], irradiating a sample as pulses with an approximate 50-100 ps pulse duration [62, 63, 64] and an approximately constant 10–500 MHz repetition rate (resulting in a 2-100 ns pulse separation) [65]. Expressed as photons per pulse and

pixel for a Pilatus3 detector with the smallest active area (PILATUS3 X 100K-A), the X-ray flux is approximately 0.01-1 photon per pulse and pixel. Due to the fact that the pulses hit the detector with an approximately constant time interval, the detector can usually handle the estimated photon fluxes [66].

At FemtoMAX, approximately 104 X-ray photons per pulse and pixel hit the detector with quasi-instantaneous fs time intervals. In this condition, single photon counting becomes insufficient. Instead, this research utilised a Pilatus3 1.2M detector with a Time-over-Threshold data collection mode. This method relies on integrating the resulting current from each photon absorption event as an energy (in eV). As long as the energy level is above a certain threshold (the gain), the detector is integrating the pixel values. The number of photons is obtained by a calibration of the energy values in single pixels. In addition, the generated current (or electron cloud) can be shared by adjacent pixels. Therefore, more than one photon per pixel per pulse can be integrated. A measurement have reported a 2.5 MeV maximum detection limit with a <10 % absolute error per pulse over the active area [38], well above the estimated number of detected photons of 104 photons (1.2 MeV) in this experiment.

The detection event of a photon on a detector screen can be described with a stochastic Poisson distribution [67], and the temporal aspect of the instant detection shares characteristics to the detection events at an XFEL [66]. A well characterised dynamic range is essential in order to properly differentiate the low and the high photon absorption events.

The size limitation of the detectors active area constricted the collection of high resolution photons, with a satisfactory amount of low resolution photons. A balance was achieved by centering the detector screen behind the goniometer, but offsetting it vertically with respect to the direct x-ray beam.

4.2.2 Data reduction, model building and refinement

Using conventional data reduction, scaling and merging (XDS, XSCALE, XDSCONV, respectively) [68], model building (Coot) [69] and subsequent refinement (Phenix) [30] software, the diffraction data was processed into a 1.5 Å resolution structure. The refinement was performed using the standard settings for rigid body and isotropic refinement, respectively without alternative conformations. In the case of the 100 images per rotation step, the intensities of individual pixels were summed over the entire stack into a single $n \times m$ frame (thus keeping the original shape of a single frame), prior to the data reduction.

Figure 4.5 show a sample of the Bragg diffraction patterns from a single image (4.5a) and the summed images (4.5b). In both panels, only a weak background contribution is evident. Figure 4.5a show distinguishable diffraction peaks from only a single X-ray pulse. Figure 4.5b show streaks in the diffraction pattern, likely caused by a convolution of contributions. For instance, one contribution could be the relatively wide energy bandwidth during the diffraction at the sample [70, 71]. It is represented by the ratio of $d\lambda/\lambda$, where an expression for λ is given by Bragg's law. After a differentiation with respect to the diffraction angle θ, the ratio becomes: $\frac{\delta\lambda}{\lambda} = cot(\theta)\delta\theta \Rightarrow \frac{\delta E}{E} = cot(\theta)\delta\theta$. Other contributions to the streaked pattern includes a beam broadening of the diffracted X-ray from the monochromator (as a consequence of the wide energy bandwidth) [72, 71], and a sum of contributions from different mosaic domains, lattice deformations or other defects [70, 24, 73, 71].

Despite the fact that the reflections exhibited the streaked profile and the large energy bandwidth, XDS were able to sufficiently process the data. The wavelength was set as a constant value, and the variation in spot profiles was compensated with a larger calculated beam divergence and a larger standard deviation of the reflecting range, compared to the 1x summation.

A Wilson plot over resolution (figure 4.6a), and a cumulative Wilson distribution for centric and acentric reflections (figure 4.6b), both calculated with the software *TRUNCATE* [74] of the CCP4 suite [75] show no significant deviations of the

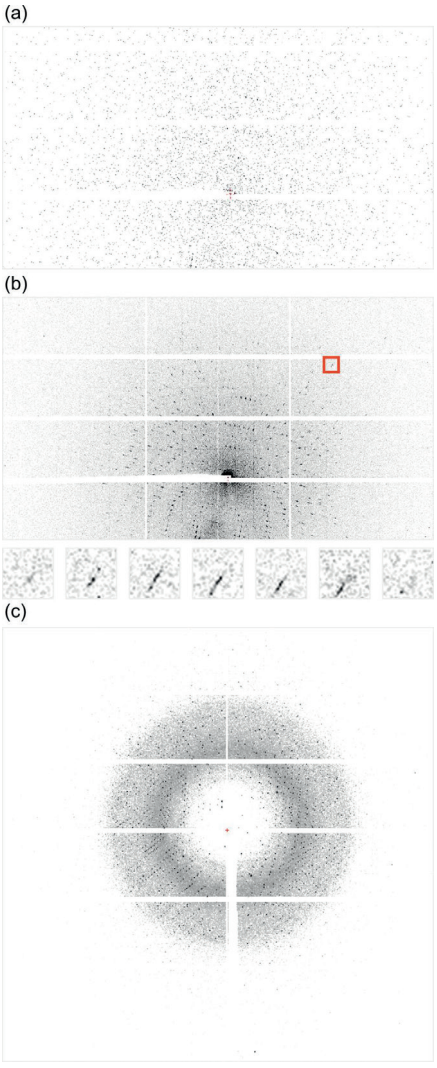


Figure 4.5: An image of sample diffraction, collected at FemtoMAX. a) Shows the diffraction from one frame, b) The diffraction from the sum of 100 frames with insets (below) of a spot profile over several frames, c) A comparison of a diffraction image from the BioMAX beamline. 15 images from room-temperature bovine trypsin crystal, in MiTeGen plastic capillaries were collected at every 11 ms by an Eiger 16M detector in a shutterless mode. The X-rays were exposing the crystal over 1.5 s, over a rotation of 1.5° oscillation angle. The figure is reprinted from **paper II**

100x summed diffraction from previously reported data [27] and from theoretical predictions. The deviation from the straight line at the low resolution is most likely related to deviations from the randomness of the probability distribution, due to called Debye effects (for instance stereochemical restraints of atom bonds), the uneven binning of low resolution shells, or from a lower amount of low resolution reflections (or combinations thereof) [27].

With the softwares *POINTLESS* and *AIMLESS* [76] of the CCP4 suite [75], a plot of R_{meas} versus intensity of the summed diffraction data (figure 4.6c) show that the weakest Bragg reflections contribute the most to the R_{meas} parameter. An opposite relationship would indicate dynamic range issues of the detector.

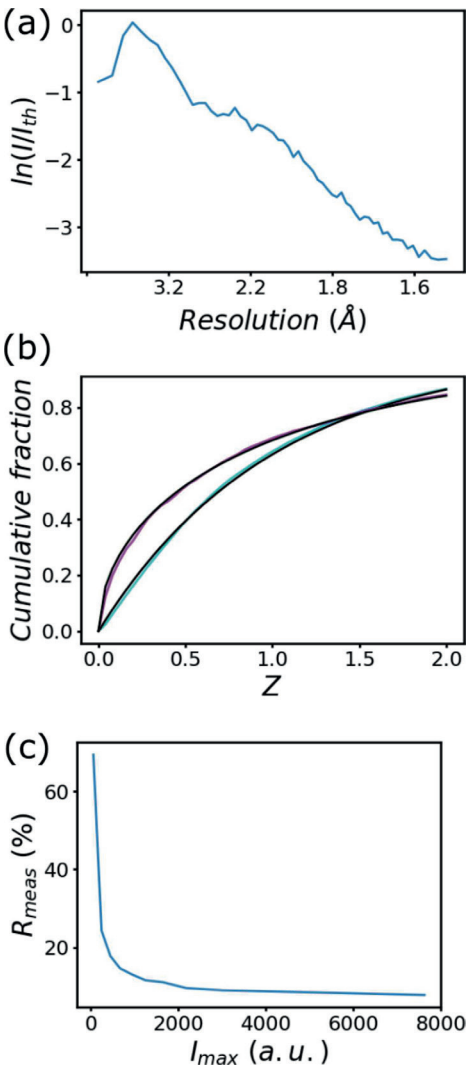


Figure 4.6: A figure over the Wilson plot. a), displays the cumulative distribution function b) display Rmeas vs I c). after the French-Wilson treatment in TRUNCATE. The figure is reprinted from **paper II**

The resulting electron density after model building and refinement show detailed features, for instance the aromatic rings in the summed case (figure 4.7a), and contours of the amino acids in the non-summed case (figure 4.7b). The average maximum-likelihood coordinate error, related to the refined average B-factors, were 0.18 Å^2 and 0.31 Å^2, for both models respectively.

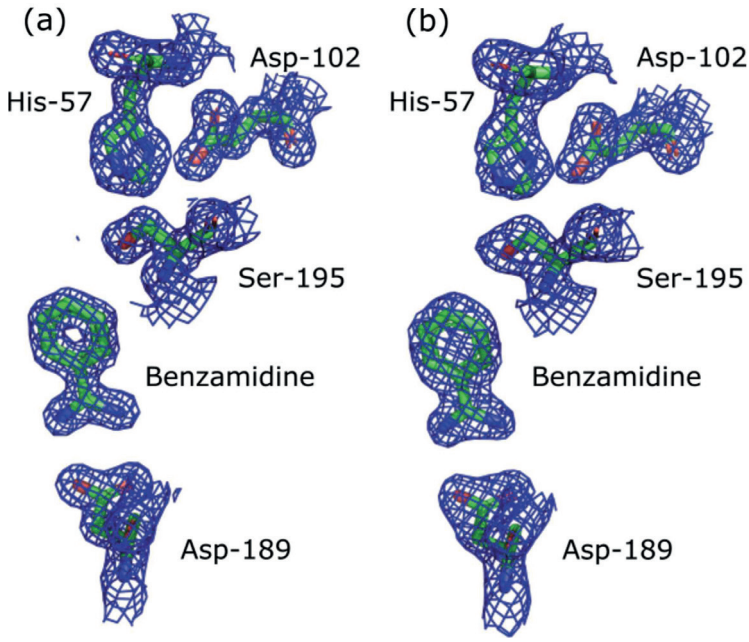


Figure 4.7: A figure over the $2mF_o$-DF_c electron density of the catalytic triad and the benzamidine atom in the summed structure (a), and the non-summed structure (b). The contour level of electron density was set as 1.5 σ. Missing observed electron density was not added as calculated density. The figure is reprinted from **paper II**

In general, the results from the data collection, the processing statistics and the resulting model all indicate a working and well diffracting model system, with no apparent issues with treating the streaked diffraction profiles, and ToT mode of the detector. However, bovine trypsin is one of the best diffracting proteins and only the best diffraction statistics was acquired from the summed images. Therefore,

collecting multiple images is most likely a requirement for weaker diffracting protein systems. The requirement will be less strict when the beamline is upgraded with higher repetition rates. However, the maximum achievable data transfer rate will ultimately limit the amount of frames which can be collected.

4.3 Picosecond out-of-equilibrium structural dynamics in photosyntetic reaction center, induced by optical excitation

In **paper III**, the extremely high multiplicity were utilised in a in a optical pump- X-ray probe SFX experiment of photosynthetic reaction center crystals from *Blastochloris viridis*, to show an induced structural dynamics, with statistical trends from bootstrapped C_α atom coordinate displacement. The atom coordinate displacement data is supported by an in depth structural and statistical analysis of electron difference maps, time-resolved vibrational spectroscopy and from a hybrid quantum mechanics–molecular mechanics modelling.

Booststrapping is a computational resampling technique, introduced by Efron [77, 78, 79], to estimate the unknown joint probability distribution (F^n, a n-fold product of F) of a random sample $\chi_n = \{X_1, ..., X_n\}$, where X_i is a set of identical and identically distributed random variables), and the (point) estimator $\hat{\theta}_n$ and its associated statistical parameter $\theta = \theta(F)$ of the marginal distribution F [79]. If $\chi_n^* = \{X_1^*, ..., X_n^*\}$ describe the bootstrapped sample, obtained from the estimated distribution $\tilde{F}_n$, the parameter θ_n^* expresses the bootstrapped point estimator and $\theta = \theta(\tilde{F}_n)$ in this case. Two centralised random variables can be defined from the expression of the estimators: $\hat{\theta}_n - \theta$ and $\theta_n^* - \theta$, of which their probabilities are evaluated by the sampling distributions: G_n and $\hat{G}_n$, respectively. As an example, if the sampling distribution of the sample mean is estimated, $\hat{\theta}_n - \theta$ and $\theta_n^* - \theta$ can be described as: $T_n = \sqrt{n}(\hat{X}_n - \mu)$ and $T_{n,n}^* = \sqrt{n}(\hat{X}_n^* - \hat{X}_n)$, for the two centralised random variables, respectively. The expression for the maximum error

of the distributions is given by: $\Delta_n = \sup_x |\hat{G}_n(x) - G_n(x)| = \sup_x |P_*(T^*_{n,n}(x) \leq x) - P(T_n) \leq x)| \rightarrow O(1)$, as $n \rightarrow \infty$, and where sup denote the supremum norm of the difference, over the variable x. From the last expression, it shows that the bootstrapping method is a valid approximation for the sampling distribution G [79].

In a structural biology context, using the bootstrapping technique have shown relevance in a large variety of research, such as single particle reconstruction [80], 3D analysis of electron diffraction maps [81] and sequence comparison [82]. With the recent emerge of structures from ultrafast X-ray diffraction techniques, resampling-based analysis of structural dynamics, have gained importance ([83, 84], **paper III**, **paper IV**).

4.3.1 Data processing

The diffraction intensities from the X-ray diffraction experiments, were processed in Cheetah and in the CrystFEL software suit. During the partialator step of the CrystFEL data reduction, the diffraction intensities were assigned a specific label, corresponding to the excitation state of the experiment (non-excited, or excited with a specific time delay), and separated, using a custom shell script. The separated data were either directly used in further data reduction, model-building and refinement, to produce the electron difference maps, or bootstrapped, to produce the models for the atom displacement analysis. Each of the resampled models of the dark (non-excited) and excited structures were produced from 100 cycles of rigid body and isotropic refinement, respectively, using REFMAC5. A starting model to the refinement was produced with molecular replacement, using the Bragg reflections of one of the resampled datasets and the PDB entry 5NJ4 as the search model.

4.3.2 Resampling as a data analysis tool

In total, the intensities for each unique miller index in the asymmetric units were bootstrapped 100 times. From the resulting lists of mean intensities and corresponding $\sigma(I)$s (each unique miller index, with an associated mean intensity and $\sigma(I)$ was used only once), 100 resulting datasets of each photo-excited state (including the dark state) were produced. In turn, the 100 datasets resulted in 100 protein models after the subsequent refinement in REFMAC5. General trends of C_α Euclidian distances of individual atoms between the excited structures and the dark structures were visualised by first producing a 100×100 distance matrix for each individual C_α atom. The calculation was repeated for the pairs of dark structures. The data in the matrices were used as a basis for calculating an average of the ratio of the Euclidian distances in the excited states to the dark states, over the distances of the atoms between the dark states only.

In this research, the bootstrapping methodology demonstrated a powerful method to obtain statistically relevant information, using model parameters of individual atoms, as a complement to electron difference maps. The bootstrapping is powerful in the sense that it doesn't require any information of the underlying probability distribution of the resampled parameter [77]. In addition, resampling, or the act of investigating statistical questions from a repeated randomisation of data, demonstrate a large versatility in how the statistical questions is answered. For instance, jackknife resampling could be used to estimate a point estimate by reiterative summing over an existing sample, and excluding a sample point each time, according to the expression: $\tilde{\theta}_i = n\hat{\theta}_n - (n-1)\frac{1}{n}\sum_{n=1}^{\infty}\hat{\theta}(-i)$, where (-i) denote the exclusion of observation i [85], or a resampling without replacement methodology were utilised in **paper IV** to better express the statistical trends from a large sample of dependent X-ray diffraction images.

4.3.3 Model analysis

The photosynthetic reaction is an integral membrane protein, existing in light-harvesting organisms such as plants, algae, cyanobacteria and photosynthetic bacteria. Is is involved in transferring charge over the cell membrane in the complex photosystem II in bacteria, and over the thylakoid membrane of chloroplasts in the complexes photosystem I and photosystem II, in oxygen producing organisms [86, 87].

Time-resolved optical spectroscopy [86], cyclic voltammetry [88, 89], and Spectro-electrochemical measurements [89] have revealed a charge transfer kinetics from the the special pair of bacteriochlorophylls, travelling primarily along the L-subunit of transmembrane helices through the bacteriochlorophyll and bacteriopheophytin molecules, to the menaquinone active site. These results are supported by structure determination [90]. Wöhri et al. were able to relate the structural dynamics of the dark state to a 3 ms excited state with electron difference maps near the special pair, time-resolved optical spectroscopy, and proposed a model, with support from MD simulations [91]. However, the results in **paper III** show a more detailed analysis with SFX data measured at 1 ps, 5 ps, 20 ps, 300 ps and 8 µs, after the excitation event.

The analysis of the average C_α displacement ratios (4.8) reveal a structural difference, originating close to the special pair and follow the protein model along the transmembrane helices toward the menaquinone and ubiquinone.

The analysis demonstrate the importance of the resampling technique, in this case, as a complement to other experimental measures. Diederichs et al. discuss and show the importance of a high multiplicity in anomalous difference data, to improve data reduction statistics and to obtain better structural models at higher resolution after refinement [92]. Similarly, the resampling of the X-ray diffraction data show how relevant structural information is obtained with a large number of data points.

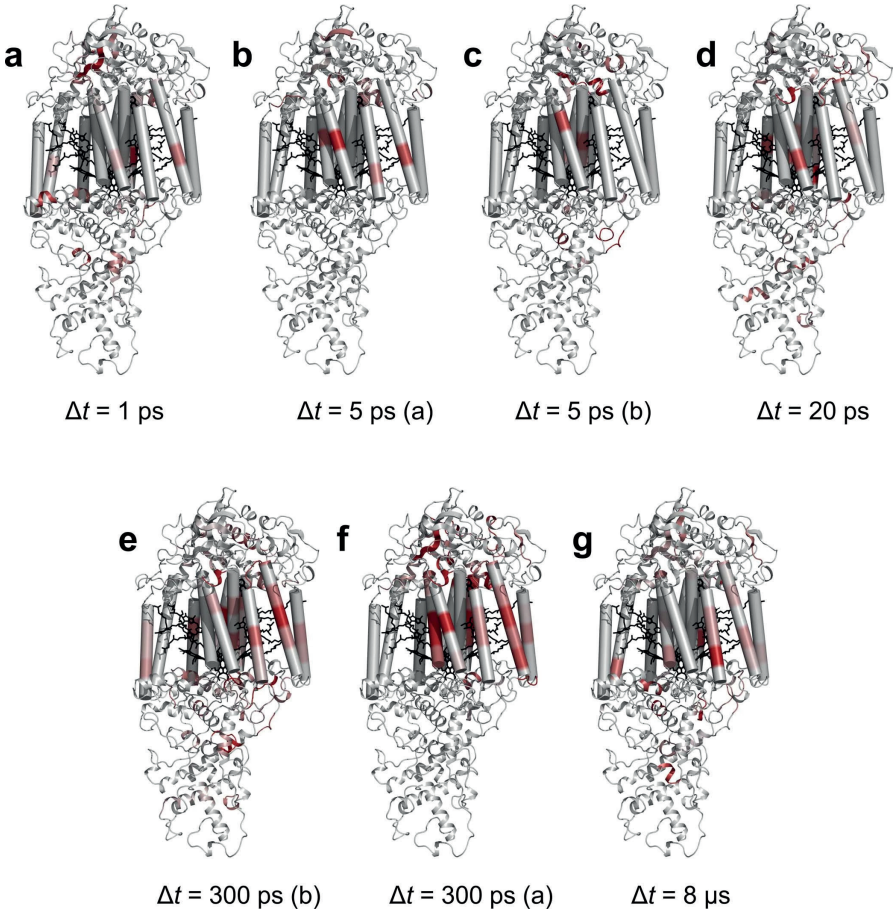


Figure 4.8: Structures, obtained from the different resampled datasets, with the bootstrapped C_α distances mapped. The red colour depict the (correlation corrected) average value of the ratio of the light over dark distances. The colour range is defined as: < 80 % of the maximum error-weighted mean ratio, to red as ≥ 95 % of the maximum error-weighted mean ratio. The figure is reprinted from **paper III**

4.4 Picosecond out-of-equilibrium structural dynamics in bovine trypsin, induced by a strong terahertz electromagnetic field

In **paper IV**, the structural dynamics in bovine trypsin were measured at ps timescales at the FemtoMAX beamline, during irradiation of a strong (1.2×10^2 V/μm) THz electromagnetic field. This research demonstrate the first steps to elucidate the temporal details of the structural perturbations, induced by THz irradiation.

In addition, the research presents the results of the first time-resolved MX experiment at the FemtoMAX beamline. It exemplifies the non-destructive nature of an extensive (18 h) low X-ray flux pump-probe experiment. With overall satisfactory data reduction statistics, no significant radiation damage were observed on X-ray diffraction intensities, .

Finally, the research applies and expands on the resampling strategy, developed in **Paper III**. It focuses on how model parameters of individual atoms can independently describe structural dynamics from external perturbations, with support from molecular dynamics simulations and theoretical predictions.

4.4.1 Data collection, resampling and processing

X-ray crystallographic data were collected under a commissioning beamtime, with the same experimental parameters as in **Paper II**. In this experiment, groups of 100 still X-ray diffraction images were collected at an angle of 0.1° per group, over a total of approximately 130°, with a fixed time delay between the pump and the probe pulse (between 35-50 ps over three individual datasets).

The high redundancy per each group of images provided ideal conditions for image processing with resampling. To preserve the correlated errors of adjacent

frames, 50 pairs of odd and even frames per rotation step were resampled without replacement. Half of the data was randomly removed per sampling to increase the randomisation. To increase the signal, a final photon count per rotation step was obtained by summing the photon count in each pixel over the sampled 25 frames. The resampled datasets were individually processed in XDS, XSCALE, XDSCONV and REFMAC5, using a constant starting model and no prior phase information in the refinement.

A limitation of the resampling strategy involves the fact that differences of model parameters, originating from the odd or the even dataset, could include averaging over its inverse difference (for instance: an average of the difference from frame 3 - frame 2 could include frame 2 - frame 3 as well). Therefore, a future experiment requires a modulation in the resampling strategy.

One proposed strategy involves separating the data in two equally large sets and independently resample the sets. If differences are specifically calculated from example set 1 - set 2, the problem of overlapping frames are avoided.

To preserve the time correlation between each frame, a variety of the block bootstrapping strategy, such as stationary block boostrapping could be utilised [79]. The advantages would be a faster data collection since not a complete rotation would be required. In addition, a higher degree of randomisation of the diffraction images could be achieved, resulting in the possibility of a larger production of resampled datasets, with the requirements of a higher redundancy of diffraction images per set.

Similar to **paper I**, the odd and even data separation didn't provide any noticeable differences in the subsampling of the rocking curves. Thus, non-THz related differences in model parameters are more likely related to subtle differences in experimental parameters, for instance an uneven x-ray flux per frame, local variations in temperature (however unlikely, given the fast time scale of the pump pulse with respect to the probe), or slight crystal-to-crystal heterogeneity due to, for instance, radiation damage, a slight crystal slippage or an uneven sampling of photon count per frame. Strong differences in the said parameters would most likely generate

evident differences in the data reduction and refinement statistics. The more subtle differences would most likely be included within the error bars (standard deviation) of each respective average.

4.4.2 Model analysis

The resampling methodology provided an suitable platform to statistically examine the refined model parameters, and extract strong differences of the THz-irradiated data, compared to the reference. In this study, we focused the analysis on the phonon contribution to individual isotropic B-factors, displacements of atomic positions and their associated Euclidean distances. The experimental results were supported with theoretical calculations and MD simulations of displacement correlations (not described in this thesis).

The insufficient resampling strategy resulted in an artificial reduction of the error bars in data which were only dependent on one parameter (ΔB for instance). Therefore, averages of ΔB from non-irradiated datasets were removed from the analysis in the paper.

B-factor analysis

Average isotropic B-factor differences, calculated between the THz-irradiated and non-irradiated frames and normalised over their standard deviations, were investigated for individual atoms. A mapping of the normalised differences of C_α atoms to a standard protein model, revealed discrete sequences of amino acids (labelled in the figure) with either large positive or negative B-factors differences, suggesting a local rearrangement of ordered and disordered amino acid residues upon THz irradiation. These discrete areas were retained in two other crystals (data not shown).

The fact that the order or disorder is mostly concentrated to the two β-barrel

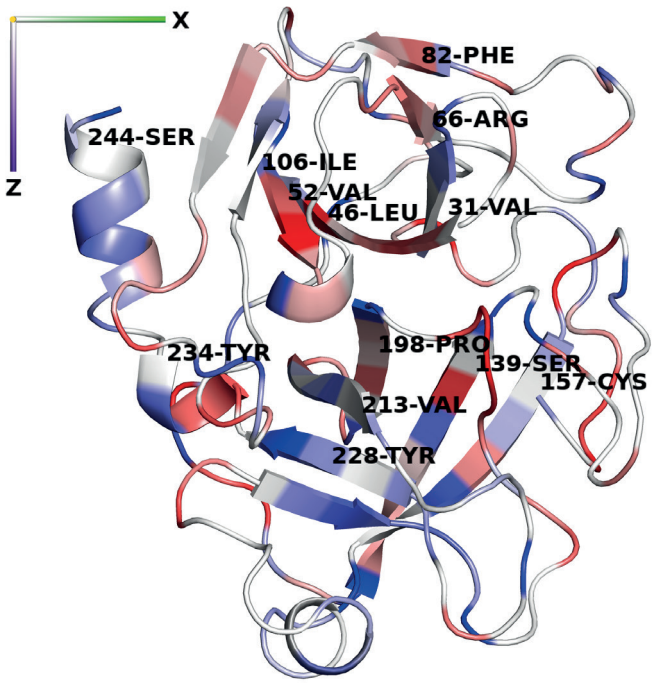


Figure 4.9: The average of difference B factors, over their corresponding standard deviations, mapped to the bovine trypsin model. The protein is coloured as $\langle \Delta B \rangle \leq -2\sigma$ (blue), $\langle \Delta B \rangle \approx 0$ (white) and $\langle \Delta B \rangle \geq 2\sigma$ (red), $\sigma = \mathrm{SD}$. The figure is reprinted from **paper IV**

regions suggests a concerted motion of the barrels. Low frequency THz vibrations in beta barrels of trypsin-like structures have been observed with Fourier Transform Infrared Spectroscopy (FTIR) [93], and a Gaussian network of normal modes have demonstrated correlated movement in the barrels [94].

Displacements and their associated Euclidean distances

Ratios of average Euclidian distances of individual C_αs, between the pairs of THz-irradiated (even) datasets and non-irradiated (odd) datasets, over a difference between non-irradiated datasets, mapped to a bovine trypsin model, revealed large displacement distances in the same local groups av amino acids (or C_α atoms in directly adjacent amino acids) which exhibited large difference B-factor ratios (4.10) (Gly 62 being an exception).

In addition, average displacements of the C_αs between the pairs of even-odd datasets vary discontinuously (cyan vectors), but suggests an directional preference, either in the x-y plane, or in the z-direction of the crystallographic axii.

The fact that approximately the same local groups of residues are affected in figure 4.9 and figure 4.10 highlights the importance of the residues upon the THz-pumping. The estimated atom coordinate error $\approx 45 - 150 \times \delta r$ for individual atoms [95], but the directional preference is still prevalent.

Further analysis by Principal Component Ananlysis (PCA) of the average displacement vectors for C_α, C_β and carbonyl oxygen atoms, in two different types of structures: flexible loops, or more ordered α-helices and β-sheets, exemplify the directional preference (figure 4.11).

The PCA is represented by the eigenvalue decomposition of the precision matrix $\boldsymbol{S}^{-1}$ (the inverse of the covariance matrices $\boldsymbol{S}$ of the displacements) in the basis of a given position vector ($\boldsymbol{u}$), centred around the sample means of the vectors ($\boldsymbol{\mu}$),

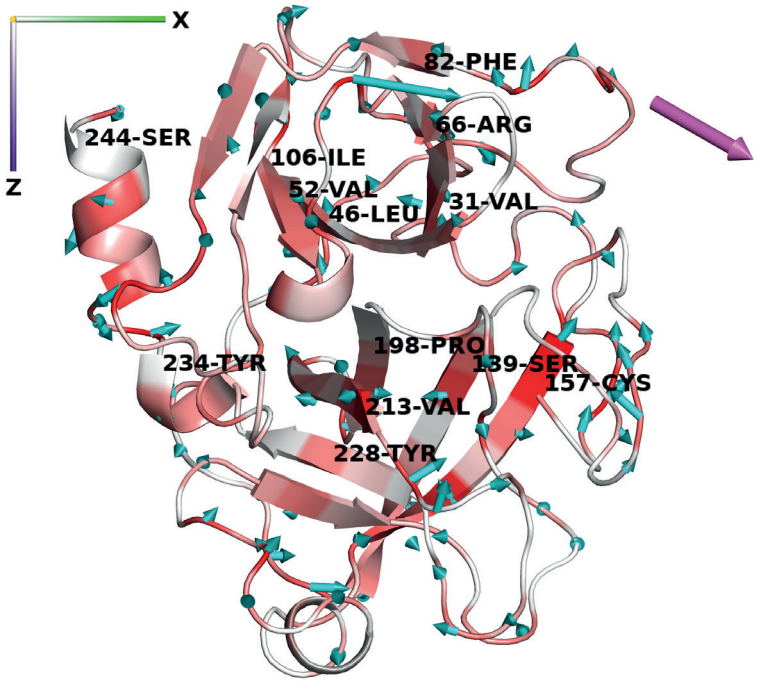


Figure 4.10: The mapped resampled average distance ratios and their corresponding average displacement from the FemtoMAX experiment. The color range is defined as: ≥ 2 (red) to $1 \leq$ (white) The figure is reprinted from **paper IV**.

where (x1-μ1,y1-μ1,z1-μ1), (x2-μ2,y2-μ2,z2-μ2) etc.) as:

$$\frac{1}{2}(\boldsymbol{u}-\boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\boldsymbol{u}-\boldsymbol{\mu}) > 0 \tag{4.2}$$

obtained from the expression of the multivariate normal distribution.

The calculations show the first principal component of C_α and C_β is approximately colinear with the crystallographic x-axis, and the first principal component of the carbonyl O is approximately colinear with the z-axis, respectively.

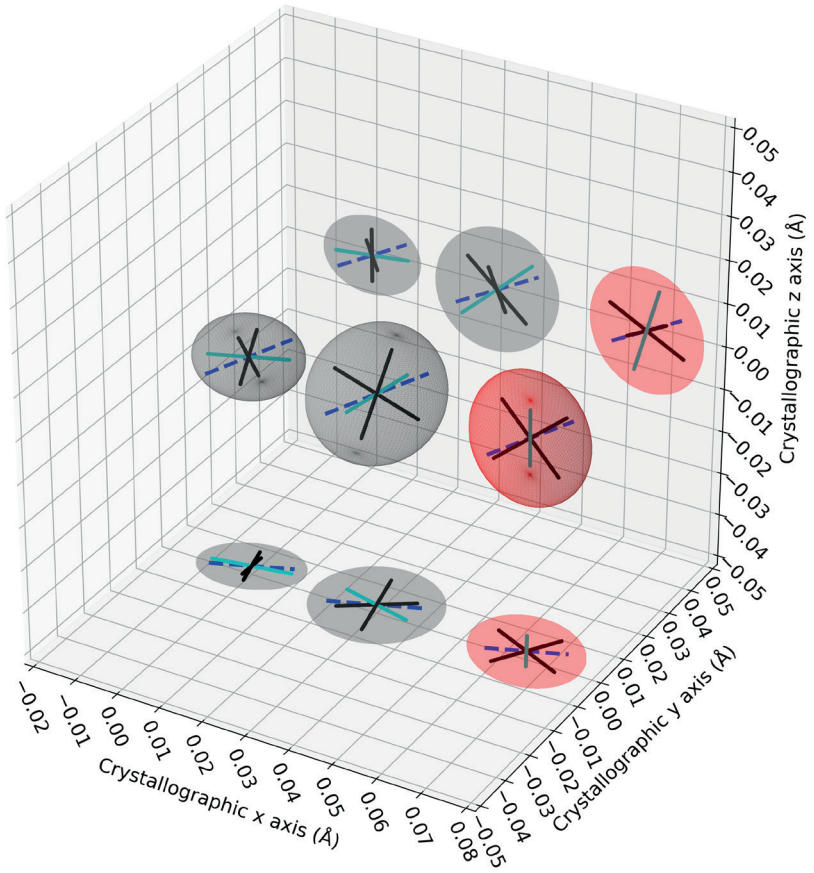


Figure 4.11: The graphical representation of the PCA components of the individual average displacements in C_α, C_β, and carbonyl O atoms (C_α - leftmost to carbonyl O - rightmost). The eigenvectors of the precision matrices and the square root of the eigenvalues describe the direction and radius of each ellipsoid's semi-axii. The first principal component vector (the longest axis of the ellipsoids) is highlighted with the cyan colour. The dashed line depicts the polarisation direction of the THz field. The figure is reprinted from **paper IV**

Chapter 5

Conclusions and outlook

Immediately upon the discovery of the first protein structure, myoglobin, the need for a dynamical model to describe the entry of its bound oxygen became apparent [96]. With that mindset, structural biologists have actively developed protein models and research tools to derive structural dynamics. With this thesis, a contribution to that research field is made.

In **paper I** and **paper IV**, this thesis demonstrated on different time scales, differences in individual atomic parameters, which can be attributed to a coherent state of non-thermal phonon vibrations, upon THz irradiation. The research regarding vibrational states is important for the understanding of vibrational-induced or vibrational-assisted biochemical reactions. Could for instance a reaction rate be manipulated by inducing the vibrations? Such results are predicted by quantum mechanics/molecular mechanics (QM/MM) modelling in purine nucleoside phosphorylase [97]. These types of questions could be tested with for instance bovine trypsin and a substrate in a THz pump-optical probe spectroscopy set-up. Additionally, by investigating the clustering of atoms, based on similarities of the the components of the anisotropic displacement parameter tensor, **paper I**

demonstrated a method to extract structural information which could be connected to the function of the protein. Similar research is demonstrated by Wenzhe et al. with a Gaussian network model [94]. The clustering data in **paper I** provided the necessary background for a machine learning algorithm, thus demonstrating the potential use in structural prediction. This type of research is important in for instance detecting protein amyloid fibril formation in neurodegenerative diseases, such as Alzheimer's disease or Parkinson's disease [98].

With **paper II**, this thesis provided and discussed the results of the first time resolved MX experiment at the FemtoMAX beamline. These results are a milestone in the development of ultrafast X-ray sources at large-scale facilities. With the experimental difficulties handled, the results show high resolution X-ray diffraction data, with no apparent effects of radiation damage or data processing issues. With the planned upgrade of a 50-fold increase in the repetition rate, a significantly higher amount of data can be collected. However, the higher repetition rate induces a larger demand on technical aspects of the data collection software, such as converting the raw data into processable file formats, and assigning metadata information at a sufficient rate. A balance could be achieved by for instance collecting over a smaller total rotation (without too much sacrifice of the completeness), simultaneously as collection over a higher amount of frames.

Finally, with **paper III** and **paper IV**, this thesis bridged a new way to use classical statistics with computational methods, in the context of MX data analysis. With the resampling strategies, statistically significant differences of pumped to non-pumped protein models were associated with individual atoms, without the requirement of modelling any unknown atom specific probability distributions. Further research need include alternative resampling strategies. For instance, a block bootstrapping strategy, described by Lahiri et al. [79], could provide a suitable alternative to preserve the dependency of the data, while keeping the randomisation per dataset high. To evaluate the results from the different resampling strategies, a standard method is assess the mean squared error between the different resampling strategies [79].

Acknowledgements

First and formost, I would like to thank my supervisor **Gergely**. You accepted me as a PhD student in 2016 and since then I've never had any dull moment during my work. You have always been an inspiration for me and someone to aspire to when it comes down how to develop a critical thinking, and evolve as a scientist. Over the years, we have discussed so many things. For instance: science, life, teaching at the university, politics, the climate change, economy, technology, food, beer, strange behaviour in animals ... You say you are a man of few words, but I can tell you, find the right subject and we could go on for hours!

Second, I would also like to thank **Richard**, my co-supervisor. You are also an amazing scientist and I really appreciate your passion, your drive and your way of summarising research projects in order to make them understandable for everyone. I will forever remember our first interview. You affected me so strongly back then that I almost forgot my second interview with Gergely!

Third, I would like to thank **Leif** for being my examiner and for the time during the KEM011-teaching.

My research projects would have been impossible without the contribution from two sources: our group's **research partners** and **my colleagues**.

Thank you **Juan**, **Helena** and **Jan** at the Terahertz and Millimetre Wave Laboratory at Chalmers, for constantly borrowing your THz-equipment and for your

endless support with our THz-sources. Thank you **Stefano** and **Ran** for the nice discussions during the THz projects, for the theoretical calculations of protein vibrations and for the MD and NMA simulations.

Thank you all new and old colleagues at the Lundberg lab who made the stay enjoyable during coffee breaks or lunch breaks, and provided overall very nice, helpful and enjoyable work environment!

To the old and new Katona group: Thank you **Maja**, **Laras**, **Majo** and **Rhawnie**. I will always remember the adventures we shared during beamtimes and conferences, and fika times.

Thank you **Rob Dods**, **Adams** and **Petra Båth** for all the help during the reaction center project.

Thank you **Cecilia**, **Jonatan** and **Tinna** for sharing the office with me and enduring the long discussions I had with Gergely, and for letting me play death metal without any (major) complaints.

Thank you **Florian** for the Magic gaming sessions and our totally not serious discussions about games and music. The bromance is indeed real!

Thank you **Greger** for the running sessions when you are running in front of me, constantly talking, while I'm panting behind. Your motivation inspired me to run my first half-marathon!

Thank you my lovely family for staying by me and supporting me no matter what. Especially thank you **Anna**, **Freja** and **the little one**. You three are the love of my life and my apples in my eyes.

Bibliography

[1] Antosiewicz J, Shugar D. UV–Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 1: basic principles and properties of tyrosine chromophore. Biophysical Reviews. 2016 05;8.

[2] Prasad S, Mandal I, Singh S, Paul A, Mandal B, Venkatramani R, et al. Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. Chem Sci. 2017;8:5416-33. Available from: http://dx.doi.org/10.1039/C7SC00880E.

[3] Leermakers PA, Vesley GF. Organic photochemistry and the excited state. Journal of Chemical Education. 1964;41(10):535. Available from: https://doi.org/10.1021/ed041p535.

[4] Ragnoni E, Catalini S, Becucci M, Lapini A, Foggi P. Linear and Non-Linear Middle Infrared Spectra of Penicillin G in the CO Stretching Mode Region. Symmetry. 2021;13(1). Available from: https://www.mdpi.com/2073-8994/13/1/106.

[5] Alexiev U, Farrens D. Fluorescence spectroscopy of rhodopsins: Insights and approaches. Biochimica et biophysica acta. 2013 10;1837.

[6] Ozakai Y. Infrared Spectroscopy-Mid-infrared, Near-infrared, and Far-infrared/Terahertz Spectroscopy. Analytical Sciences. 2021;37(9):1193-212.

[7] In: Lakowicz JR, editor. Protein Fluorescence. Boston, MA: Springer US; 2006. p. 529-75. Available from: https://doi.org/10.1007/978-0-387-46312-4_16.

[8] Xu Y, Havenith M. Perspective: Watching low-frequency vibrations of water in biomolecular recognition by THz spectroscopy. The Journal of Chemical Physics. 2015;143(17):170901. Available from: https://doi.org/10.1063/1.4934504.

[9] Turton DA, Senn HM, Harwood T, Lapthorn AJ, Ellis EM, Wynne K. Terahertz underdamped vibrational motion governs protein-ligand binding in solution. Nature communications. 2014;5(1):1-6.

[10] Yu B, Zeng F, Yang Y, Xing Q, Chechin A, Xin X, et al. Torsional vibrational modes of tryptophan studied by terahertz time-domain spectroscopy. Biophysical Journal. 2004;86(3):1649-54.

[11] Austin RH, Roberson MW, Mansky P. Far-infrared perturbation of reaction rates in myoglobin at low temperatures. Physical review letters. 1989;62(16):1912.

[12] Yang X, Zhao X, Yang K, Liu YP, Liu Y, Fu W, et al. Biomedical Applications of Terahertz Spectroscopy and Imaging. Trends in Biotechnology. 2016 05;34.

[13] Lacidogna G, Piana G, Bassani A, Carpinteri A. Raman spectroscopy of Na/K-ATPase with special focus on low-frequency vibrations. Vibrational Spectroscopy. 2017;92:298-301. Available from: https://www.sciencedirect.com/science/article/pii/S0924203117301996.

[14] Dielmann-Gessner J, Grossman M, Conti Nibali V, Born B, Solomonov I, Fields GB, et al. Enzymatic turnover of macromolecules generates long-lasting protein–water-coupled motions beyond reaction steady state. Proceedings of the National Academy of Sciences. 2014;111(50):17857-62. Available from: https://www.pnas.org/content/111/50/17857.

[15] Niessen K, Xu M, Paciaroni A, Orecchini A, Snell E, Markelz A. Moving in the Right Direction: Protein Vibrations Steering Function. Biophysical Journal. 2017 03;112:933-42.

[16] Yamazaki T, Van Driessche AE, Kimura Y. High mobility of lattice molecules and defects during the early stage of protein crystallization. Soft matter. 2020;16(8):1955-60.

[17] Frontzek (neé Svanidze) AV, Paccou L, Guinet Y, Hédoux A. Study of the phase transition in lysozyme crystals by Raman spectroscopy. Biochimica et Biophysica Acta (BBA) - General Subjects. 2016;1860(2):412-23. Available from: https://www.sciencedirect.com/science/article/pii/S0304416515002925.

[18] Nanev CN. Advancements (and challenges) in the study of protein crystal nucleation and growth; thermodynamic and kinetic explanations and comparison with small-molecule crystallization. Progress in Crystal Growth and Characterization of Materials. 2020;66(2):100484. Available from: https://www.sciencedirect.com/science/article/pii/S0960897420300115.

[19] Zhou RB, Cao HL, Zhang CY, Yin DC. A review on recent advances for nucleants and nucleation in protein crystallization. CrystEngComm. 2017;19:1143-55. Available from: http://dx.doi.org/10.1039/C6CE02562E.

[20] Lutsko JF. How crystals form: A theory of nucleation pathways. Science Advances. 2019;5(4):eaav7399. Available from: https://www.science.org/doi/abs/10.1126/sciadv.aav7399.

[21] Beck C, Grimaldo M, Roosen-Runge F, Maier R, Matsarskaia O, Braun M, et al. Following Protein Dynamics in Real Time during Crystallization. Crystal Growth & Design. 2019;19(12):7036-45. Available from: https://doi.org/10.1021/acs.cgd.9b00858.

[22] Lovelace J, Borgstahl G. Characterizing pathological imperfections in macromolecular crystals: lattice disorders and modulations. Crystallography Reviews. 2019 12;26:1-48.

[23] Fenwick RB, van den Bedem H, Fraser JS, Wright PE. Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. Proceedings of the National Academy of Sciences. 2014;111(4):E445-54. Available from: https://www.pnas.org/content/111/4/E445.

[24] Nave C, Sutton G, Evans G, Owen R, Rau C, Robinson I, et al. Imperfection and radiation damage in protein crystals studied with coherent radiation. Journal of Synchrotron Radiation. 2016 Jan;23(1):228-37. Available from: https://doi.org/10.1107/S1600577515019700.

[25] Trueblood KN, Bürgi HB, Burzlaff H, Dunitz JD, Gramaccioli CM, Schulz HH, et al. Atomic Dispacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature. Acta Crystallographica Section A. 1996 Sep;52(5):770-81. Available from: https://doi.org/10.1107/S0108767396005697.

[26] Na H, Hinsen K, Song G. The amounts of thermal vibrations and static disorder in protein X-ray crystallographic B-factors. Proteins: Structure, Function, and Bioinformatics. 2021;89(11):1442-57. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.26165.

[27] Rupp B. Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology. New York: Garland Science, Taylor & Francis Group, LLC; 2010.

[28] Meisburger SP, Thomas WC, Watkins MB, Ando N. X-ray scattering studies of protein structural dynamics. Chemical reviews. 2017;117(12):7615-72.

[29] Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. *REFMAC*5 for the refinement of macromolecular crystal structures. Acta Crystallographica Section D. 2011 Apr;67(4):355-67. Available from: https://doi.org/10.1107/S0907444911001314.

[30] Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. Acta Crystallographica Section

D. 2019 Oct;75(10):861-77. Available from: https://doi.org/10.1107/S2059798319011471.

[31] Winn M, Ballard C, Cowtan K, Dodson E, Emsley P, Evans P, et al. Overview of the CCP4 suite and current development. Acta crystallographica Section D, Biological crystallography. 2011 04;67:235-42.

[32] Murshudov G, Skubak P, Lebedev A, Pannu N, Steiner RA, Nicholls R, et al. REFMAC5 For the refinement of macromolecular crystal structures. Acta crystallographica Section D, Biological crystallography. 2011 04;67:355-67.

[33] Zaccai G. How Soft Is a Protein? A Protein Dynamics Force Constant Measured by Neutron Scattering. Science. 2000;288(5471):1604-7. Available from: https://www.science.org/doi/abs/10.1126/science.288.5471.1604.

[34] Pearce N, Gros P. A method for intuitively extracting macromolecular dynamics from structural disorder. Nature Communications. 2021 09;12:5493.

[35] Meisburger SP, Case DA, Ando N. Diffuse X-ray scattering from correlated motions in a protein crystal. Nature Communications. 2020;11(1):1271.

[36] Weissenborn R, Diederichs K, Welte W, Maret G, Gisler T. Non-thermal microwave effects on protein dynamics? An X-ray diffraction study on tetragonal lysozyme crystals. Acta Crystallographica Section D. 2005 Feb;61(2):163-72. Available from: https://doi.org/10.1107/S0907444904030902.

[37] Lundholm IV, Rodilla H, Wahlgren WY, Duelli A, Bourenkov G, Vukusic J, et al. Terahertz radiation induces non-thermal structural changes associated with Fröhlich condensation in a protein crystal. Structural Dynamics. 2015;2(5):054702. Available from: https://doi.org/10.1063/1.4931825.

[38] Enquist H, Jurgilaitis A, Jarnac A, Bengtsson ÅUJ, Burza M, Curbis F, et al. FemtoMAX - an X-ray beamline for structural dynamics at the short-pulse facility of MAX IV. Journal of Synchrotron Radiation. 2018 Mar;25(2):570-9. Available from: https://doi.org/10.1107/S1600577517017660.

[39] McPherson A. Introduction to protein crystallization. Methods. 2004;34(3):254-65.

[40] Russo Krauss I, Merlino A, Vergara A, Sica F. An Overview of Biological Macromolecule Crystallization. International Journal of Molecular Sciences. 2013 05;14:11643-91.

[41] Junius N, Vahdatahar E, Oksanen E, Ferrer JL, Budayova-Spano M. Optimization of crystallization of biological macromolecules using dialysis combined with temperature control. Journal of Applied Crystallography. 2020 Jun;53(3):686-98. Available from: https://doi.org/10.1107/S1600576720003209.

[42] Chayen N, Saridakis E. Protein crystallization: From purified protein to diffraction-quality crystal. Nature methods. 2008 03;5:147-53.

[43] Beale JH, Bolton R, Marshall SA, Beale EV, Carr SB, Ebrahim A, et al. Successful sample preparation for serial crystallography experiments. Journal of Applied Crystallography. 2019 Dec;52(6):1385-96. Available from: https://doi.org/10.1107/S1600576719013517.

[44] Ahlberg Gagnér V, Jensen M, Katona G. Estimating the probability of coincidental similarity between atomic displacement parameters with machine learning. Machine Learning: Science and Technology. 2021;2.

[45] White TA, Barty A, Stellato F, Holton JM, Kirian RA, Zatsepin NA, et al. Crystallographic data processing for free-electron laser sources. Acta Crystallographica Section D. 2013 Jul;69(7):1231-40. Available from: https://doi.org/10.1107/S0907444913013620.

[46] White TA. Processing serial crystallography data with *CrystFEL*: a step-by-step guide. Acta Crystallographica Section D. 2019 Feb;75(2):219-33. Available from: https://doi.org/10.1107/S205979831801238X.

[47] Mehrabi P, Bücker R, Bourenkov G, Ginn HM, von Stetten D, Müller-Werkmeister HM, et al. Serial femtosecond and serial synchrotron crystallography can yield data of equivalent quality: A systematic comparison. Science Advances. 2021;7(12):eabf1380. Available from: https://www.science.org/doi/abs/10.1126/sciadv.abf1380.

[48] Shimizu N, Hirata K, Hasegawa K, Ueno G, Yamamoto M. Dose dependence of radiation damage for protein crystals studied at various X-ray energies. Journal of Synchrotron Radiation. 2007 Jan;14(1):4-10. Available from: `https://doi.org/10.1107/S0909049506049296`.

[49] Hofmann P. 4. Thermal Properties of the Lattice. In: Solid State Physics: An Introduction. 2nd ed. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2015. p. 48.

[50] Comsol. COMSOL Multiphysics®, ver. 5.6 [computer software]. Stockholm: Comsol AB; 2020. [cited 190122]. Available from: `www.comsol.com`.

[51] Sarkar R. Native flexibility of structurally homologous proteins: insights from anisotropic network model. BMC Biophysics. 2017;10.

[52] Merritt EA. Expanding the model: anisotropic displacement parameters in protein structure refinement. Acta Crystallographica Section D. 1999 Jun;55(6):1109-17. Available from: `https://doi.org/10.1107/S0907444999003789`.

[53] Hekstra D, White K, Socolich Mea. Electric-field-stimulated protein mechanics. Nature. 2016;540(7633):400–405.

[54] Wang X, Ekström JC, Bengtsson AUJ, Jarnac A, Jurgilaitis A, Pham VT, et al. Role of Thermal Equilibrium Dynamics in Atomic Motion during Nonthermal Laser-Induced Melting. Phys Rev Lett. 2020 Mar;124:105701. Available from: `https://link.aps.org/doi/10.1103/PhysRevLett.124.105701`.

[55] Jarnac A, Wang X, Bengtsson ÅUJ, Ekström JC, Enquist H, Jurgilaitis A, et al. Communication: Demonstration of a 20 ps X-ray switch based on a photoacoustic transducer. Structural Dynamics. 2017;4(5):051102. Available from: `https://doi.org/10.1063/1.4993730`.

[56] Saaring J, Feldbach E, Nagirnyi V, Omelkov S, Vanetsev A, Kirm M. Ultrafast Radiative Relaxation Processes in Multication Cross-Luminescence Materials. IEEE Transactions on Nuclear Science. 2020;67(6):1009-13.

[57] Rack A, Weitkamp T, Riotte M, Grigoriev D, Rack T, Helfen L, et al. Comparative study of multilayers used in monochromators for synchrotron-based coherent hard X-ray imaging. Journal of Synchrotron Radiation. 2010 Jul;17(4):496-510. Available from: https://doi.org/10.1107/S0909049510011623.

[58] Bury CS, Brooks-Bartlett JC, Walsh SP, Garman EF. Estimate your dose: RADDOSE-3D. Protein Science. 2018;27(1):217-28. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/pro.3302.

[59] Atakisi H, Conger L, Moreau DW, Thorne RE. Resolution and dose dependence of radiation damage in biomolecular systems. IUCrJ. 2019 Nov;6(6):1040-53. Available from: https://doi.org/10.1107/S2052252519008777.

[60] Dectris/detecting the future. PILATUS3 X for Synchrotrons [Internet]. Baden-Dättwil: Dectris AG; [cited 190122]. Available from: https://www.dectris.com/detectors/x-ray-detectors/pilatus3/pilatus3-for-synchrotrons/pilatus3-x/.

[61] Dectris/detecting the future. EIGER2 X/XE for Synchrotrons [Internet]. Baden-Dättwil: Dectris AG; [cited 190122]. Available from: https://www.dectris.com/detectors/x-ray-detectors/eiger2/eiger2-for-synchrotrons/eiger2-x/.

[62] Silly MG, Ferté T, Tordeux MA, Pierucci D, Beaulieu N, Chauvet C, et al. Pump−probe experiments at the TEMPO beamline using the low-α operation mode of Synchrotron SOLEIL. Journal of Synchrotron Radiation. 2017 Jul;24(4):886-97. Available from: https://doi.org/10.1107/S1600577517007913.

[63] Stepanov AG, Hauri CP. Short X-ray pulses from third-generation light sources. Journal of Synchrotron Radiation. 2016 Jan;23(1):141-51. Available from: https://doi.org/10.1107/S1600577515019281.

[64] Couprie ME, Filhol JM. X radiation sources based on accelerators. Comptes Rendus Physique. 2008;9(5):487-506. Synchrotron x-rays and condensed matter. Available from: https://www.sciencedirect.com/science/article/pii/S1631070508000571.

[65] Di Mitri S. One way only to synchrotron light sources upgrade? Journal of Synchrotron Radiation. 2018 Sep;25(5):1323-34. Available from: https://doi.org/10.1107/S160057751800810X.

[66] Bergamaschi A, Mozzanica A, Schmitt B. XFEL detectors. Nature Reviews Physics. 2020;2(7):335–336.

[67] Förster A, Brandstetter S, Schulze-Briese C. Transforming X-ray detection with hybrid photon counting detectors. Phil Trans R Soc A. 2019;377:20180241.

[68] Kabsch W. *XDS*. Acta Crystallographica Section D. 2010 Feb;66(2):125-32. Available from: https://doi.org/10.1107/S0907444909047337.

[69] Emsley P, Cowtan K. *Coot*: model-building tools for molecular graphics. Acta Crystallographica Section D. 2004 Dec;60(12 Part 1):2126-32. Available from: https://doi.org/10.1107/S0907444904019158.

[70] Deacon A, Appleby T, Bilderback D, Ealick S, Fontes E, Thiel D. Protein crystallography using a multilayer monochromator. Journal of synchrotron radiation. 1998 06;5:494-6.

[71] Nave C. Matching X-ray beam and detector properties to protein crystals of different perfection. Journal of Synchrotron Radiation. 2014 May;21(3):537-46. Available from: https://doi.org/10.1107/S1600577514003609.

[72] Kazimirov A, Smilgies DM, Shen Q, Xiao X, Hao Q, Fontes E, et al. Multilayer X-ray optics at CHESS. Journal of Synchrotron Radiation. 2006 Mar;13(2):204-10. Available from: https://doi.org/10.1107/S0909049506002846.

[73] Smith E, Helmstaedt H, Flemming R. Survival of Brown Colour in Diamond During Storage in the Subcontinental Lithospheric Mantle. Canadian Mineralogist - CAN MINERALOG. 2010 06;48:571-82.

[74] French S, Wilson K. On the treatment of negative intensity observations. Acta Crystallographica Section A. 1978 Jul;34(4):517-25. Available from: https://doi.org/10.1107/S0567739478001114.

[75] Winn M, Ballard C, Cowtan K, Dodson E, Emsley P, Evans P, et al. Overview of the CCP4 suite and current development. Acta crystallographica Section D, Biological crystallography. 2011 04;67:235-42.

[76] Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallographica Section D. 2013 Jul;69(7):1204-14. Available from: https://doi.org/10.1107/S0907444913000061.

[77] Efron B. Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics. 1979;7(1):1 26. Available from: https://doi.org/10.1214/aos/1176344552.

[78] Kreiss JP, Paparoditis E. Bootstrap methods for dependent data: A review. Journal of the Korean Statistical Society. 2011;40(4):357-78. Available from: https://www.sciencedirect.com/science/article/pii/S1226319211000780.

[79] Lahiri SN. Resampling Methods for Dependent Data: Springer Series in Statistics. New York: Springer-Verlag; 2003.

[80] Penczek PA, Yang C, Frank J, Spahn CMT. Estimation of variance in single-particle reconstruction using the bootstrap technique. Journal of Structural Biology. 2006;154(2):168-83. Available from: https://www.sciencedirect.com/science/article/pii/S1047847706000219.

[81] Cheng A, Yeager M. Bootstrap resampling for voxel-wise variance analysis of three-dimensional density maps derived by image analysis of two-dimensional crystals. Journal of structural biology. 2007;158 1:19-32.

[82] Green RE, Brenner SE. Bootstrapping and normalization for enhanced evaluations of pairwise sequence comparison. Proceedings of the IEEE. 2002;90(12):1834-47.

[83] Grünbein ML, Gorel A, Foucar Lea. Effect of X-ray free-electron laser-induced shockwaves on haemoglobin microcrystals delivered in a liquid jet. Nature Communications. 2021;12(1):1672.

[84] Nass K, Gorel A, Abdullah MMea. Structural dynamics in proteins induced by and probed with X-ray free-electron laser pulses. Nature Communications. 2020;11:1814.

[85] Miller RG. The jackknife-a review. Biometrika. 1974;61:1-15.

[86] Feher G, Allen J, Okamura Mea. Structure and function of bacterial photosynthetic reaction centres. Nature. 1989;339(6220):111–116.

[87] Fukuzumi S, Lee YM, Nam W. Mimicry and functions of photosynthetic reaction centers. Biochemical Society Transactions. 2018 10;46(5):1279-88. Available from: https://doi.org/10.1042/BST20170298.

[88] Trammell SA, Wang L, Zullo JM, Shashidhar R, Lebedev N. Orientated binding of photosynthetic reaction centers on gold using Ni-NTA self-assembled monolayers. Biosensors and Bioelectronics. 2004;19(12):1649-55. Available from: https://www.sciencedirect.com/science/article/pii/S095656630400003X.

[89] Lee SH, Blake I, Larsen A, McDonald J, Ohkubo K, Fukuzumi S, et al. Synthetically tuneable biomimetic artificial photosynthetic reaction centres that closely resemble the natural system in purple bacteria†. Chem Sci. 2016 10;7.

[90] Dods R, Båth P, Arnlund D, Beyerlein K, Nelson G, Liang M, et al. From Macrocrystals to Microcrystals: A Strategy for Membrane Protein Serial Crystallography. Structure. 2017 08;25.

[91] Wöhri AB, Katona G, Johansson LC, Fritz E, Malmerberg E, Andersson M, et al. Light-Induced Structural Changes in a Photosynthetic Reaction Center Caught by Laue Diffraction. Science. 2010;328(5978):630-3. Available from: https://www.science.org/doi/abs/10.1126/science.1186159.

[92] Karplus PA, Diederichs K. Assessing and maximizing data quality in macromolecular crystallography. Current Opinion in Structural Biology. 2015;34:60-8.

Carbohydrate-protein interactions • Biophysical and molecular biological methods. Available from: https://www.sciencedirect.com/science/article/pii/S0959440X15000937.

[93] Balakhnina I, Brandt N, Mankova A, Chikishev AY. The problem of manifestation of tertiary structure in the vibrational spectra of proteins. Vibrational Spectroscopy. 2021 04;114:103250.

[94] Ma W, Tang C, Lai L. Specificity of Trypsin and Chymotrypsin: Loop-Motion-Controlled Dynamic Correlation as a Determinant. Biophysical journal. 2005 09;89:1183-93.

[95] Manickam G, Shankar M, Raju N, Helliwell J, Sekar K. Do we see what we should see? Describing non-covalent interactions in protein structures including precision. IUCrJ. 2014 07;1:74-81.

[96] Kendrew JC, Bodo G, Dintzis HM, Parrish R, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature. 1958;181(4610):662-6.

[97] Schramm VL, Schwartz SD. Promoting vibrations and the function of enzymes. Emerging theoretical and experimental convergence. Biochemistry. 2018;57(24):3299-308.

[98] Pinheiro F, Santos J, Ventura S. AlphaFold and the amyloid landscape. Journal of Molecular Biology. 2021;433(20):167059. From Protein Sequence to Structure at Warp Speed: How Alphafold Impacts Biology. Available from: https://www.sciencedirect.com/science/article/pii/S0022283621002771.