

12-1-2015

# Time-Resolved, Near Atomic Resolution Structural Studies at the Free Electron Laser

Jason James Tenboer

*University of Wisconsin-Milwaukee*

Follow this and additional works at: <https://dc.uwm.edu/etd>

 Part of the [Biophysics Commons](#), and the [Physics Commons](#)

---

## Recommended Citation

Tenboer, Jason James, "Time-Resolved, Near Atomic Resolution Structural Studies at the Free Electron Laser" (2015). *Theses and Dissertations*. 1086.

<https://dc.uwm.edu/etd/1086>

This Dissertation is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact [open-access@uwm.edu](mailto:open-access@uwm.edu).

TIME-RESOLVED, NEAR ATOMIC RESOLUTION STRUCTURAL STUDIES AT THE  
FREE ELECTRON LASER

by

Jason Tenboer

A Dissertation Submitted in  
Partial Fulfillment of the  
Requirements for the Degree of

Doctor of Philosophy

in Physics

at

The University of Wisconsin-Milwaukee

December 2015

## ABSTRACT

### TIME-RESOLVED, NEAR ATOMIC RESOLUTION STRUCTURAL STUDIES AT THE FREE ELECTRON LASER

by

Jason Tenboer

The University of Wisconsin-Milwaukee, 2015  
Under the Supervision of Professor Marius Schmidt

*Time-resolved serial femtosecond crystallography (TR-SFX) employs X-ray free electron lasers (XFELs) to provide X-ray pulses of femtosecond (fs) duration with  $10^{12}$  photons per pulse. These XFELs are more than a billion times more brilliant than 3<sup>rd</sup> generation synchrotron X-ray sources. For structure determination, protein crystals on the micrometer length scale (microcrystals) are injected into the X-ray beam and the resulting diffraction patterns are recorded on fast-readout pixel detectors. Although these intense pulses deposit enough energy to ultimately destroy the protein, the processes that lead to diffraction occur before the crystal is destroyed. This so-called diffraction-before-destruction principle overcomes radiation damage, which is one of the challenges that time-resolved crystallographers face at synchrotron X-ray sources. Most importantly, since each diffraction image is obtained from a fresh crystal, reversible and non-reversible reactions may be studied since both are now placed on equal footing. This is not currently possible at synchrotrons. Therefore, XFELs may provide a path forward to study reactions catalyzed by enzymes.*

A TR-SFX experiment requires enormous effort and success hinges upon thorough preparation: a sufficient quantity of purified protein must be produced for the study; techniques

for creating microcrystals need to be developed; these samples should then be tested with a *gas dynamic virtual nozzle* (GDVN) and initial studies must be performed to characterize these crystals. Since only 15% of all XFEL experiment proposals are ultimately accepted, previous results that strongly support such proposals significantly improve the chances for obtaining beamtime. I have, therefore, constructed three instruments: 1) a micro-focus X-ray diffraction beamline, 2) a near ultraviolet / visual wavelength fast microspectrophotometer and 3) a GDVN fabrication and testing facility. These machines supply the crucial initial information that is needed, not only for creating engaging XFEL beamtime proposals, but also for preparing for these experiments once beamtime has been awarded.

With an initial experiment performed at the *Linac Coherent Light Source* (LCLS) we demonstrated for the first time that time-resolved near atomic resolution serial crystallography was possible at an X-ray FEL. This study laid the groundwork for observing the uncharacterized structures of the *trans-cis* isomerization of the *photoactive yellow protein* (PYP) photocycle on the fs timescale. Continuing on this work, we have now determined these previously unknown structures with another experiment at the LCLS. This successful fs time-resolved experiment demonstrates the full capability and vision of XFELs with respect to photoactive proteins.

In addition to studying both reversible and irreversible photo-initiated reactions, XFELs offer the unique opportunity to explore irreversible enzymatic reactions by the mix-and-inject technique. In this method, microcrystals are mixed with a substrate and the following reaction is probed by the fs X-ray pulses in a time-resolved fashion. An interesting candidate for the mix-and-inject method is *cytochrome c nitrite reductase* (ccNiR). This protein uses a 6 electron reduction of nitrite to produce ammonia, which is one of the key reduction processes in the nitrogen cycle. High quality, large single crystals and microcrystals of ccNiR have been produced. This work is being

done in collaboration with the Pacheco group in the Chemistry Department at the University of Wisconsin-Milwaukee. We have obtained a 1.65 Å native structure and a 2.59 Å nitrite-bound structure of ccNiR. These early studies will provide the foundation for a future time-resolved mix-and-inject XFEL proposal to study this protein.

## TABLE OF CONTENTS

<b>ABSTRACT</b> .....	<b>ii</b>
<b>LIST OF FIGURES</b> .....	<b>viii</b>
<b>LIST OF TABLES</b> .....	<b>x</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>xi</b>
<b>1 Introduction</b> .....	<b>1</b>
1.1 A (very) Brief History of Crystallography .....	2
1.2 3 <sup>rd</sup> Generation Synchrotron X-ray Sources .....	4
1.3 XFELs and TR-SFX .....	4
1.4 Gas Dynamic Virtual Nozzles (GDVNs) .....	6
1.5 Photoactive Yellow Protein (PYP) .....	7
1.6 Cytochrome c Nitrite Reductase (ccNiR).....	9
1.7 $\alpha$ -phycoerythrocyanin ( $\alpha$ -PEC) .....	11
1.8 Protein Microcrystals.....	12
1.9 Schmidt Lab (SLAB) Beamlines 1, 2 and 3 (SBL1, SBL2 and SBL3).....	13
<b>2 Materials and Methods</b> .....	<b>14</b>
2.1 Monochromatic X-ray Crystallography.....	14
2.1.1 Scattering from an Atom .....	14
2.1.2 Scattering from a Molecule .....	16
2.1.3 Scattering from a Crystal.....	16
2.1.4 The Reciprocal Lattice .....	18
2.1.5 Mosaicity and Structural Heterogeneity.....	21
2.1.6 Polarization Factor, Lorentz Factor and the Integrated Scattered Intensity .....	23
2.1.7 Single Crystal Sample Preparation.....	24
2.1.8 Single Crystal Sample Mounting .....	25
2.1.9 Single Crystal Data Collection.....	27
2.1.10 Single Crystal Data Reduction .....	27
2.1.11 Assessment of Data Quality, Structural Refinement and Analysis.....	30

2.2 Time-Resolved Serial Femtosecond Crystallography (TR-SFX).....	32
2.2.1 TR-SFX: Microcrystal Samples .....	32
2.2.2 TR-SFX: T <sub>0</sub> Determination and the Timing Tool .....	35
2.2.3 TR-SFX: Data Collection.....	40
2.2.4 TR-SFX: Data Reduction .....	42
2.2.5 Singular Value Decomposition (SVD) of Time-Dependent DED Maps .....	45
2.2.6 Posterior Analysis and Extrapolated Electron Density Maps .....	46
2.2.7 TR-SFX: Extent of Reaction Initiation .....	47
2.2.8 TR-SFX: Assessment of Data Quality .....	47
2.3 SBL1, SBL2, SBL3, 3 <sup>rd</sup> Generation Synchrotrons and XFELs.....	48
2.3.1 Schmidt Beamline 1 (SBL1) .....	48
2.3.2 Schmidt Beamline 2 (SBL2) .....	54
2.3.3 Schmidt Beamline 3 (SBL3) and Gas Dynamic Virtual Nozzles (GDVNs) .....	64
2.3.4 APS, BioCARS 14-ID-B and 14-BM-C Beamlines .....	65
2.3.5 X-ray Free Electron Lasers (XFELs) and the Coherent X-ray Imaging (CXI) Beamline .....	67
2.4 Dynamics in Biological Macromolecules with Nanosecond Time Resolution.....	68
2.5 Dynamics in Biological Macromolecules with Femtosecond Time Resolution .....	72
2.6 Structure Determination of the Native Form of ccNiR .....	73
2.7 Structure Determination of the Nitrite-bound Form of ccNiR .....	74
<b>3 Results .....</b>	<b>75</b>
3.1 PYP Structural Dynamics with Nanosecond Time Resolution .....	75
3.2 PYP Structural Dynamics with Femtosecond Time Resolution.....	81
3.3 Structure Determination of the Native Form of ccNiR to 1.65 Å .....	85
3.4 Structure Determination of the Nitrite-bound Form of ccNiR to 2.59 Å .....	88
<b>4 Discussion.....</b>	<b>92</b>
4.1 LCLS and APS Scientific Output.....	92
4.2 Reduction of Protein Consumption and Expansion of Structural Biology with SFX93	
4.3 Photoactive Proteins: <i>cis-trans</i> Isomerizations and E-Z Isomerizations .....	97
4.4 Enzymes: Structure and Applications.....	105

4.5 Conclusion and Outlook .....	109
<b>5 References.....</b>	<b>111</b>
<b>Curriculum Vitae .....</b>	<b>126</b>



## LIST OF FIGURES

Figure 1.1 Gas dynamic virtual nozzle (GDVN) schematic.....	6
Figure 1.2 Photoactive yellow protein (PYP).....	8
Figure 1.3 Cytochrome c nitrite reductase (ccNiR).....	10
Figure 1.4 $\alpha$ -phycoerythrocyanin ( $\alpha$ -PEC).....	11
Figure 2.1 Scattering of X-rays by an atom and atomic form factors.....	15
Figure 2.2 A 2-dimensional crystal lattice.....	18
Figure 2.3 The reciprocal lattice and the Ewald Sphere.....	19
Figure 2.4 Mosaicity of a crystal.....	21
Figure 2.5 Large single crystal sample mounting methods.....	26
Figure 2.6 Microcrystal formation.....	33
Figure 2.7 The timing tool.....	40
Figure 2.8 Visual representation of an SVD analysis.....	45
Figure 2.9 SBL1.....	50
Figure 2.10 SBL1 graphical user interfaces (GUIs).....	52
Figure 2.11 Data collection and reduction tools.....	53
Figure 2.12 SBL2.....	56
Figure 2.13 SBL2 GUI of everything GUI (GOEGUI).....	58
Figure 2.14 SBL2 sample mounting methods.....	62

Figure 2.15 Gas Dynamic Virtual Nozzle (GDVN) Construction.....	65
Figure 2.16 TR-SFX nanosecond PYP experiment laser schematic.....	70
Figure 3.1 Simulation of ns laser penetration depth for PYP.....	78
Figure 3.2 Comparison of DED maps from TR-SFX and TR-LX .....	79
Figure 3.3 Stereo representation of the extrapolated map at the 1 $\mu$ s time-delay.....	80
Figure 3.4 The <i>trans-cis</i> isomerization of the PYP chromophore.....	84
Figure 3.5 Structure determination of the native form of ccNiR.....	87
Figure 3.6 Structure determination of the nitrite-bound form of ccNiR.....	90
Figure 4.1 Comparison of the scientific output of the LCLS and APS.....	92
Figure 4.2 The Lipidic cubic phase (LCP) injector.....	95
Figure 4.3 Comparison of PAS domains.....	98
Figure 4.4 E-Z isomerization of the phycoviolobin (PVB) chromophore of $\alpha$ -PEC.....	99
Figure 4.5 The concept of the QM/MM method.....	102
Figure 4.6 Fluorescence spectroscopy of PYP.....	104
Figure 4.7 A rectangular crystal of dimensions 2a, 2b and 2c.....	106
Figure 4.8 Mixing injectors.....	107
Figure 4.9 A comparison of the heme arrangement of HAO, OTR and ccNiR.....	108

## LIST OF TABLES

Table 1.1 The 20 amino acids.....	1
Table 1.2 Protein Data Bank (PDB) holdings.....	2
Table 1.3 Comparison of the LCLS and APS X-rays.....	5
Table 2.1 TR-SFX experiment personnel.....	41
Table 2.2 Primary components of SBL1.....	49
Table 2.3 Primary components of SBL2.....	55
Table 2.4 SBL2 timing signal connections.....	60
Table 2.5 DG645 channel delays for SBL2.....	61
Table 2.6 Primary components of SBL3.....	64
Table 3.1 Result summary from the TR-SFX ns PYP experiment.....	77
Table 3.2 Result summary from the TR-SFX fs PYP experiment.....	82
Table 3.3 Result summary from the structure determination of the native form of ccNiR.....	86
Table 3.4 Result summary from the structure determination of the nitrite-bound form of ccNiR...88	
Table 4.1 Diffusion times, $\tau_D$ , for various crystal sizes.....	107

## ACKNOWLEDGEMENTS

First, I wish to express my deepest respect and gratitude for my advisor, Professor Marius Schmidt. His passion and dedication for research excellence have driven and guided me throughout my undergraduate and graduate degrees. I have appreciated both his honesty and his generosity when mistakes were made, as this has helped to transform these mistakes into learning opportunities. His unwavering support for me has fostered not only a strong professional relationship, but also a friendship I consider myself extremely privileged to enjoy.

I would like to thank my committee members, Professor Valerica Raicu, Professor Dilano Saldin, Professor Peter Schwander and Professor Andy Pacheco for their helpful comments and encouragement not only in the preparation of this dissertation but also during their classes and in our Biophysics colloquium discussions.

I wish to thank my collaborators: Dr. Keith Moffat, Dr. Vukica Srajer and Dr. Robert Henning at BioCARS for their top-notch support, dedication and guidance with our numerous synchrotron experiments; Professor Andy Pacheco in the UWM Chemistry Department for our discussions and collaboration on the ccNiR experiments and to all the members of the STC for their insight and support that has enabled the unbelievable success of our LCLS experiments and subsequent publications.

To my colleagues in the ORM Schmidt Lab (SLAB), I am grateful for your friendship, humor, patience and assistance. Our lively discussions never allowed the lab to feel stale, and our group is extremely strong because of the knowledge and talents you provide. I would like to thank Jennifer Scales for her hard work purifying PYP, without which the LCLS beamtimes would not have been possible. To Glenn Niles and Tyler Norwood, I want you to know that I am nothing but optimistic about your futures, and I encourage you to tackle the tough problems. I would like to

express my gratefulness to Dr. Namrta Purwar for her support during our early experiments at SLAB and for our continuing friendship. All the best.

Last, I dedicate this degree to my family for their unwavering love and faith in my ability to realize this goal. For my mother and father, I wish to express my admiration, respect and awe at the sacrifices you have made and the support you have given my family and myself to make this achievement possible. For my loving wife Jody and our two wonderful children Juliana and Gabriella, I want you to know how grateful I am to have such an understanding, supportive and inspiring foundation to come home to each evening. This accomplishment belongs to all of us.

# 1 Introduction

What are proteins? Perhaps the most basic definition is that proteins are biological macromolecules constructed from unbranched polypeptide chains. The sequence of the polypeptide chain is called the *primary structure* of the protein which is formed from the 20 amino acids listed in Tab. 1.1 (Rupp 2009). After synthesis, the protein undergoes a folding process (Karplus and Weaver 1976; Kim and Baldwin 1982; Walters 2011) which forms the various *secondary structures*, such as  $\alpha$ -helices,  $\beta$ -strands and  $\beta$ -sheets. A particular arrangement of these secondary structures defines the *tertiary structure* which then combines with other tertiary structures and possibly with multiple ligands to form the *quaternary structure*, which is typically the biologically active form of the protein.

Absent from this description, however, is an explanation as to the function of proteins. Proteins

**Table 1.1.** The 20 amino acids.

Amino Acid	Code	Type
Alanine	ALA	hydrophobic
Glycine	GLY	hydrophobic
Isoleucine	ILE	hydrophobic
Leucine	LEU	hydrophobic
Methionine	MET	hydrophobic/sulfur
Phenylalanine	PHE	hydrophobic
Proline	PRO	hydrophobic
Valine	VAL	hydrophobic
Aspartate	ASP	acidic
Glutamate	GLU	acidic
Arginine	ARG	basic
Histidine	HIS	basic
Lysine	LYS	basic
Serine	SER	polar hydroxyl
Threonine	THR	polar hydroxyl
Tyrosine	TYR	polar hydroxyl
Asparagine	ASN	polar amino
Glutamine	GLN	polar amino
Tryptophan	TRP	polar amino
Cysteine	CYS	sulfur

generally fall into two categories: functional and structural (Raicu and Popescu 2008). Structural proteins such as collagen found in connective tissues (van der Rest and Garrone 1991), keratin found in hair, nails, and the outer layer of skin (Bragulla and Homberger 2009) and elastin found in skin and other connective tissue (Cruise 1957) supply support and protection. Functional proteins are generally involved with enzymatic or photoactive reactions, however they are also engaged in the signaling and cross-membrane transport processes of cells (Chrispeels et al. 1999;

Lodish et al. 2000). Life depends on the specific functions of proteins (Alberts et al. 2002). To determine these functions is the mission of a time-resolved crystallographer. In order to thoroughly understand how proteins work, we must first determine their 3-dimensional molecular structures and then observe how these structures change to accomplish their function. This dissertation aims to contribute to an understanding of protein structure and function by employing the methods of *time-resolved Laue crystallography* (TR-LX), *time-resolved serial femtosecond crystallography* (TR-SFX) and *time-resolved absorption spectroscopy* (TR-AS).

### 1.1 A (very) Brief History of Crystallography

For the discovery of X-ray radiation in 1895 (Röntgen 1898) Wilhelm Conrad Röntgen was awarded the Nobel Prize in Physics in 1901. X-rays were shown in 1912 to be diffracted by crystals by Max Von Laue (Laue et al. 1913), followed by the father and son contributions of William Henry Bragg and William Lawrence Bragg in 1912-1914 which demonstrated an effective technique for structure determination (Bragg 1912a; Bragg 1912b; Bragg 1913; Bragg 1914). The first 3-dimensional static protein structures were determined by John C. Kendrew on sperm whale myoglobin in 1958 (Kendrew 1958; Kendrew et al. 1958) followed by Max Perutz a few years later on hemoglobin (Perutz

et al. 1960). Seven years later in 1965 the first structure of an enzyme was published by Blake and Phillips (Blake et al. 1965; Phillips 1967), and the structure of one of the

**Table 1.2.** Protein Data Bank (PDB) holdings (information obtained from the PDB 10/20/2015).

Method	Proteins	Nucleic Acid (NA)	Protein/NA Complexes	Other	Total
X-ray	94,424	1,683	4,715	4	100,826
Electron Microscopy	636	29	213	0	878
Nuclear Magnetic Resonance (NMR)	9,782	1,131	227	8	11,148
Hybrid	81	3	2	1	87
other	168	4	6	13	191
<b>Total</b>	<b>105,091</b>	<b>2,850</b>	<b>5,163</b>	<b>26</b>	<b>113,130</b>

largest complexes, the 30S ribosomal subunit, was solved in 2000 (Wimberly et al. 2000; Ban et al. 2000; Schluenzen et al. 2000).

The *Research Collaboratory for Structural Bioinformatics* (RCSB) *Protein Data Bank* (PDB) (Berman et al. 2002) is an online resource managed by multiple institutions that is used to store molecular structures. Tab. 1.2 lists details on the structures stored in the PDB at the time of this writing. While depositions from techniques such as *Nuclear Magnetic Resonance* (NMR) spectroscopy and *Electron Microscopy* (EM) account for ~11% of the total submissions, those solved using X-ray methods account for almost 89%. X-ray crystallography has proven to be a powerful technique for determining 3-dimensional structures at atomic resolution. X-ray radiation is particularly well-suited to produce atomic resolution images since their wavelength range extends over the atomic and chemical bond length scale.

Although static structures may provide limited insight into protein function they offer no information on protein dynamics (Han and Han 2014). To fully understand how proteins function one must investigate structural changes as they happen. Monochromatic synchrotron X-ray sources require exposure times of hundreds of microseconds to seconds in duration to create a suitable diffraction image, enabling access to only those intermediates that occur on these longer timescales. The interesting reaction intermediates accumulate on much faster timescales. Beginning in the late 1980's and continuing into the 21<sup>st</sup> century, synchrotrons have developed the technique of TR-LX (Ren and Moffat 1994; Ren et al. 1999). Dr. Keith Moffat and collaborators at the *Advanced Photon Source* (APS), BioCARS 14-ID-B beamline (Graber et al. 2011) developed TR-LX to probe intermediate structures of proteins using polychromatic (Laue) X-rays from the *picosecond* (ps) timescale onward (Moffat 1989). This technique has propelled crystallography to new frontiers since the structures of reaction intermediates and the kinetics can



be simultaneously extracted from the same data (Schmidt et al. 2003; Schmidt et al. 2004b; Schmidt et al. 2005; Schmidt 2008; Schmidt et al. 2010; Schmidt et al. 2013). During an enzymatic or photoactive reaction several intermediates may form and decay (Purwar 2013) and the capability to record a complete time-series of these fast processes enables the dynamics of these intermediates to be determined.

## 1.2 3<sup>rd</sup> Generation Synchrotron X-ray Sources

Until the first hard *X-ray Free Electron Laser* (XFEL) became operational in 2009, 3<sup>rd</sup> generation synchrotron facilities were the sources of the brightest X-rays available. The APS, located at Argonne National Laboratory in Argonne, Illinois is a premier example of this technology. Instead of using bending magnets to create monochromatic X-rays, an insertion device such as a wiggler or undulator (Winick et al. 1981; Ōnuki and Elleaume 2003) may be employed to create intense so-called “pink” beams which span a wider bandwidth of wavelengths. The BioCARS 14-ID-B beamline located at the APS uses undulators to generate polychromatic X-rays with a  $\Delta E/E = 5 \times 10^{-2}$  in the energy range from 7-19 keV. This beamline can produce a flux of  $5 \times 10^{10}$  photons / pulse at 12 keV into a spot size of  $15 \times 20 \mu\text{m}^2$ . Each X-ray pulse is 100 ps in duration. The results on *cytochrome c nitrite reductase* (ccNiR) discussed in this dissertation have been obtained from the BioCARS 14-ID-B (undulator, polychromatic) and 14-BM-C (bending magnet, monochromatic) beamlines.

## 1.3 XFELs and TR-SFX

The world’s first hard XFEL source, the *Linac Coherent Light Source* (LCLS) (Emma; Arthur et al. 1998; Wu and Emma 2006), is located at the *Stanford Linear Accelerator Center* (SLAC) *National Accelerator Laboratory* in Menlo Park, California. LCLS became operational in 2009 and this machine produces X-ray pulses of tens of *femtosecond* (fs) duration with  $10^{12}$  photons per

pulse. These XFELs are more than a billion times more brilliant than 3<sup>rd</sup> generation synchrotron X-ray sources, opening the door to study irreversible enzymatic reactions, an experiment not currently possible at synchrotrons (Schmidt 2013; Wang et al. 2014). Tab. 1.3 compares the capabilities of the *Coherent X-ray Imaging* (CXI) instrument (Boutet and Williams 2010) located at the LCLS (Arthur et al. 1998) with the APS, BioCARS 14-ID-B beamline (Graber et al. 2011). While the time-averaged flux per second of the two sources is comparable, the ratio of the peak brilliances is 10<sup>13</sup>.

The fs pulses of an XFEL are so intense they destroy a protein crystal after only a single shot. Fortunately, the processes that lead to the recording of a diffraction pattern occur before the sample is destroyed. This *diffraction-before-destruction* principle (Neutze et al. 2000) is the cornerstone of a TR-SFX experiment. Since each diffraction image is serially obtained from a fresh crystal,

reversible and irreversible processes may be studied in the same fashion. Additionally, the brilliant pulses allow crystal sizes to be reduced to the micrometer length scale (microcrystals) and even smaller (Kupitz et al. 2014a). The use of microcrystals facilitates a greater range of proteins to be examined since microcrystals are often easier to

**Table 1.3.** Comparison of the LCLS and APS X-rays (modified from a talk given by Marius Schmidt).

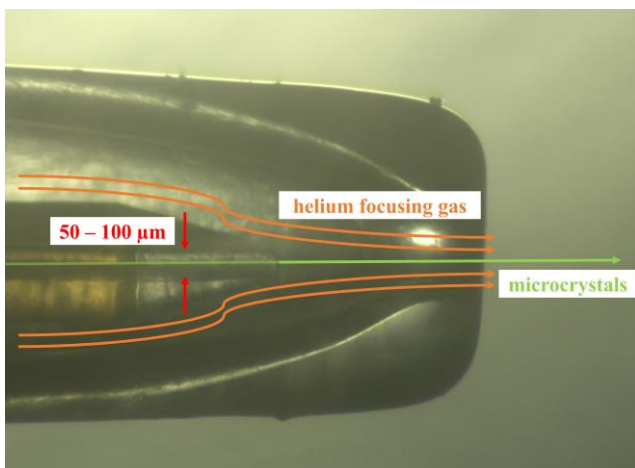
	XFEL (CXI)	Synchrotron (BioCARS 14-ID-B)
<b>X-ray pulse duration</b>	40 fs	100 ps
<b># photons/pulse</b>	10 <sup>12</sup>	5 × 10 <sup>10</sup>
<b>Bandwidth (BW)</b>	ΔE/E = 0.1%	ΔE/E = 5%
<b>Beam size [μm<sup>2</sup>]</b>	1 × 1	15 × 20
<b>Flux/pulse [photons/μm<sup>2</sup>]</b>	10 <sup>12</sup>	1.7 × 10 <sup>8</sup>
<b>Flux/pulse [photons/(μm<sup>2</sup> 0.1% BW)]</b>	10 <sup>12</sup>	3.3 × 10 <sup>6</sup>
<b>Average Flux/second [photons/(sec μm<sup>2</sup> 0.1% BW)]</b>	10 <sup>14</sup>	2 × 10 <sup>13</sup>
<b>Peak brilliance [Photons/(sec mm<sup>2</sup> mrad<sup>2</sup> 0.1% BW)]</b>	10 <sup>33</sup>	10 <sup>20</sup>

grow than large single crystals (Boutet et al. 2012; Kupitz et al. 2014b). In addition to photoactive

proteins, enzymatic reactions may also be potentially probed due to the decrease in substrate diffusion times for these small crystals (Schmidt 2013). These so-called *mix-and-inject* experiments would involve mixing microcrystals with a substrate and probing the subsequent enzymatic reaction in a time-resolved experiment at an XFEL. One of the primary missions of an XFEL is to routinely record a complete time-series at atomic resolution and physiological temperature of these non-reversible fast enzymatic reactions. This capability would catalyze the rapid advancement of structure-based drug designs (Blundell 1996).

#### 1.4 Gas Dynamic Virtual Nozzles (GDVNs)

Although the incredible brilliance of XFELs provide new avenues for exploring biological systems, they also introduce significant challenges including sample mounting / sample delivery. XFEL radiation is transported under vacuum and, when focused, these X-rays destroy nearly everything in their path. Standard crystallographic techniques used at synchrotrons employ a



**Figure 1.1.** Gas dynamic virtual nozzle (GDVN) schematic. The inner capillary of inner diameter (ID)  $\text{\O}50\text{-}100\ \mu\text{m}$  carries the pressurized microcrystal mixture. Surrounding the inner capillary is an outer capillary of  $\text{\O}765\ \mu\text{m}$  ID made from borate silica glass in which helium gas flows. This gas sheath acts as a dynamic lens, focusing the microcrystal stream into a laminar flow  $2\text{-}5\ \mu\text{m}$  in diameter.

variety of methods for sample mounting, however all of these designs employ fixed targets (Hunter et al. 2014; Feld et al. 2015) at atmospheric pressure. Since crystals are destroyed after a single XFEL shot, any fixed-style mount is therefore required to accommodate a large number of crystals.

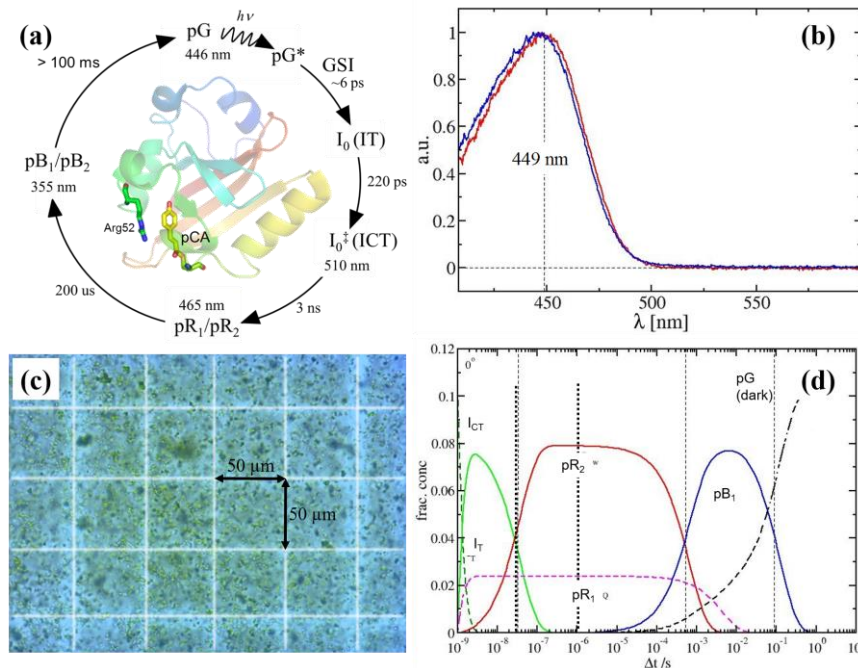
The mounting method must also be able to withstand the intense beam while keeping background scattering at a minimum.

Additionally, the mounting of new samples should not break or contaminate the sample chamber vacuum.

Techniques have been developed, including liquid jets (Sierra 2010) and aerosol sprays (Sierra et al. 2012) to deliver fully solvated microcrystals into the X-rays without compromising vacuum. Pioneering work on laminar flow *gas dynamic virtual nozzles* (GDVN) in 1998 by Dr. Alfonso Gañán-Calvo and collaborators (Gañán-Calvo 1998) laid the foundation for subsequent development and implementation of this method at the LCLS a decade later by a team led by Drs. John Spence, Bruce Doak and Uwe Weierstall at Arizona State University (Doak 2008; Doak et al. 2012; Weierstall et al. 2012; Weierstall 2014; Weierstall et al. 2014). Although fixed target sample mounts have been implemented at an XFEL (Hunter et al. 2014), the GDVN injector shown in Fig. 1.1 is so far the most frequently used method for TR-SFX (Chapman et al. 2011; Arnlund et al. 2014; Kupitz et al. 2014a). GDVNs offer numerous benefits over fixed target mounts, including ease of sample delivery, maintenance of chamber vacuum during sample and nozzle exchanges, accommodation for an essentially infinite number of microcrystals and very little contribution to background scattering. There are, however, other difficulties with this method including clogging of nozzles and the need for a large quantity of purified protein required for an experiment. Development will certainly continue towards improving the designs for both GDVNs and aerosol sprays, primarily to increase reliability, stability and to significantly reduce protein consumption.

### **1.5 Photoactive Yellow Protein (PYP)**

*Photoactive yellow protein* (PYP) is a small water-soluble light-reactive protein isolated from the bacterium *Halorhodospira halophila*, originally known as *Ectothiorhodospira halophila* (Meyer 1985), so named for its distinctive yellow color and photoactivity. This protein is suspected



**Figure 1.2.** Photoactive Yellow Protein (PYP). (a) The PYP photocycle extends from fs to seconds and contains multiple intermediates including  $I_T$ ,  $I_{CT}$ ,  $pR_1$ ,  $pR_2$ ,  $pB_1$  and  $pB_2$ . (b) Dark (no laser excitation) absorption spectrum from a crushed single crystal of PYP is displayed in blue, showing that the ground state absorbs maximally at a wavelength of 449 nm. Another dark absorption spectrum is shown in red, however, this spectrum has been obtained after a series of laser excitations, showing that no permanent photo-bleaching has occurred. (c) PYP microcrystals. Each crystal is  $\sim 2 \times 2 \times 5 \mu\text{m}^3$ . The faint white lines form a grid of  $50 \times 50 \mu\text{m}^2$  boxes. (d) Intermediate concentrations plotted as a function of time. Each time-delay may contain a mixture of multiple intermediates. A complete time-series is required to un-mix these mixtures. The dashed vertical lines denote the time-delays probed during the nanosecond PYP experiment. Fig. 1.2a and Fig. 1.2b modified from Purwar et al. 2013 and Fig. 1.2d obtained from Tenboer et al. 2014.

to play an integral role in the negative phototactic effect observed in free swimming prokaryotes (Sprengrer et al. 1993).

PYP has a molecular weight of 14 kDa and contains a *para*-hydroxycinnamic, or *para*-coumaric acid

(pCA) chromophore embedded into the protein (Meyer et al. 1987). The chromophore is covalently bound to Cys69 and hydrogen bonded to multiple residues, including

Tyr42, Glu46 and the peptide bond nitrogen of Cys69 (Van Beeumen et al. 1993; Kort et al. 1996).

Upon absorption of a blue light photon (449 nm) PYP enters a reversible photocycle, shown in Fig. 1.2a (Hellingwerf et al. 2002; Tripathi et al. 2012; Schmidt et al. 2013), with numerous intermediates ranging from the fs to seconds timescale. The *trans* to *cis* isomerization of the C2=C3 double bond after photon absorption is the signature characterization of the PYP

photocycle (Schotte et al. 2012; Jung et al. 2013). This isomerization occurs on ultrafast timescales and highly strains the chromophore, eventually leading to the red-shifted (pR<sub>1</sub> and pR<sub>2</sub>) and blue-shifted (pB<sub>1</sub> and pB<sub>2</sub>) intermediates (Kim et al. 2012; Schotte et al. 2012; Jung et al. 2013). The structures of these intermediates are known to high resolution from TR-LX data. In this dissertation we investigate the PYP photocycle on ultrafast times.

## 1.6 Cytochrome c Nitrite Reductase (ccNiR)

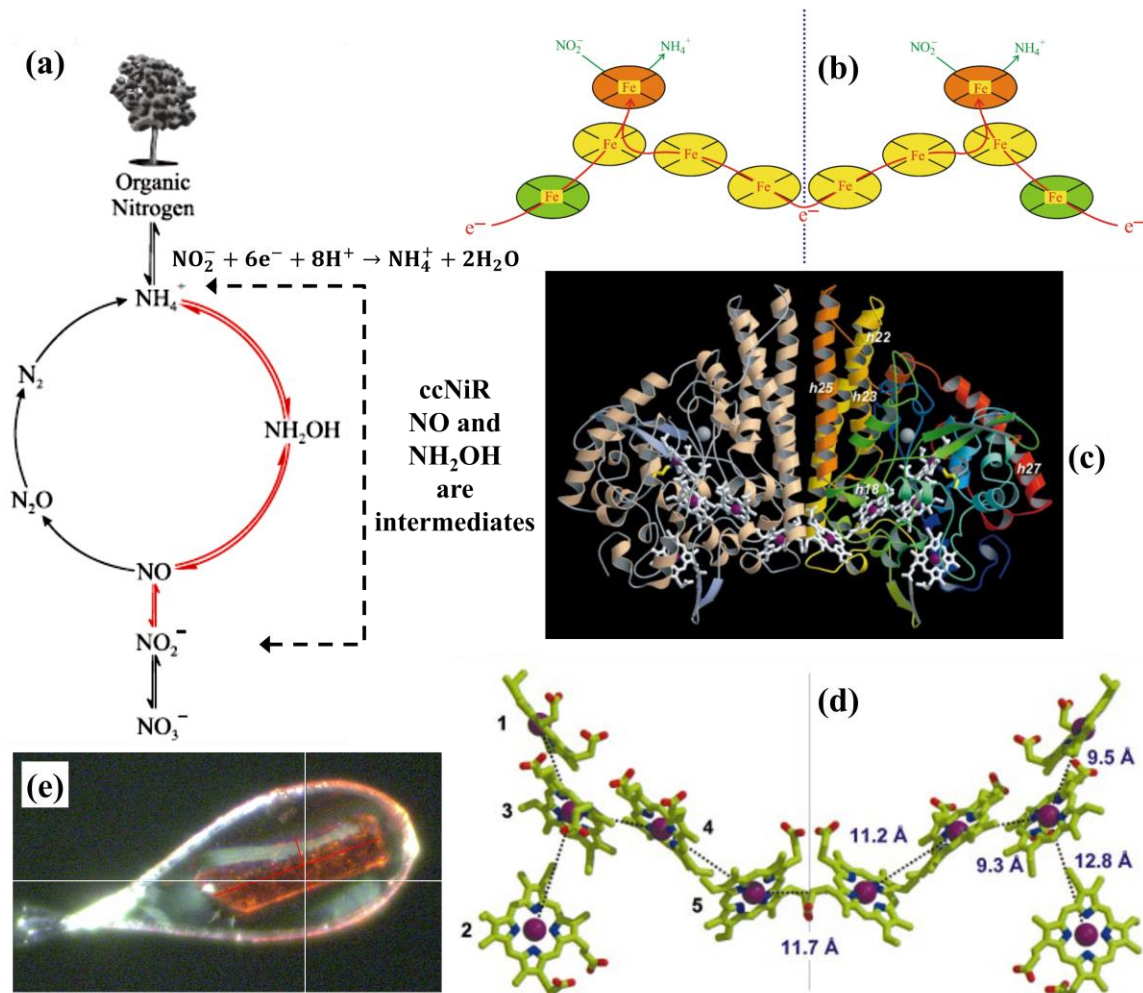
A promising candidate for a *mix-and-inject* investigation is *cytochrome c nitrite reductase* (ccNiR) (Einsle et al. 1999). This enzyme has a species-dependent molecular weight ranging from 52-65 kDa (Youngblut et al. 2014) and catalyzes a six electron reduction of nitrite to ammonia (Einsle, O., Stach, P., Messerschmidt, A., Simon, J., Kroger, A., Huber, R., Kroneck et al. 2000). This reduction is part of the nitrogen cycle, shown in Fig. 1.3a and discussed below. The results presented in this dissertation are obtained from the bacteria *Shewanella oneidensis*, originally isolated from Lake Oneida, NY (Venkateswaran et al. 1999), which utilizes ccNiR as the sole mechanism in this reduction process for its respiration (Haddock and Jones 1977; Cole and Brown 1980; Richardson 2000; Youngblut 2013).

CcNiR is a homodimeric protein that contains five hemes per monomer, and the active site is a five-coordinate c-type heme, as shown in Fig. 1.3b-Fig. 1.3d (Youngblut et al. 2012). Heme 2, shown in Fig. 1.3d (and in the corresponding position in green in Fig. 1.3b), is exposed to the protein exterior and possible electron pathways are noted by the red arrows in Fig. 1.3b leading to the active sites at heme 1 and heme 6 (Einsle et al. 2002; Judd et al. 2012). A similar arrangement is also present in the second subunit with hemes 6-10. Results will be presented for the structure of the native form of ccNiR to 1.65 Å resolution and also for a 2.59 Å resolution nitrite-bound

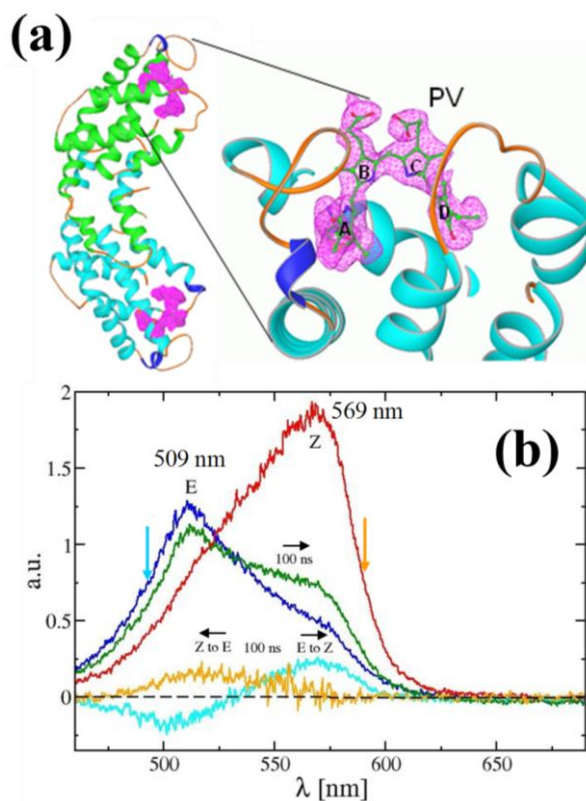


structure recorded using a home-built micro-focus X-ray diffraction beamline, the *Schmidt Beamline 1* (SBL1).

Nitrogen is required by all living organisms and is contained primarily in the atmosphere ( $N_2$ ), soil ( $NH_4^+$  and  $NO_3^-$ ) and oceans (Galloway et al. 2008). Plants cannot use atmospheric nitrogen directly but instead must absorb the  $NH_4^+$  and  $NO_3^-$  ions from the soil (Alexander 1977) (see Fig.



**Figure 1.3.** Cytochrome *c* nitrite reductase (ccNiR). (a) The nitrogen cycle. The 6 electron reduction process of nitrite to ammonia is shown by the red lines on the right side going from bottom to top. (b) A schematic representation of the heme arrangement. The orange ellipses represent the active site hemes. The vertical line denotes the dimer interface. (c) A front view of the overall structure of ccNiR. (d) The heme arrangement is shown in the same orientation as that in panels (b) and (c). The active sites are found at hemes 1 & 6, located at the top left and its mirror counterpart in the second monomer at top right, respectively. (e) A ccNiR crystal during data collection on SBL1. Fig. 1.3a from Kostera et al. 2010, Fig. 1.3b from Youngblut et al. 2012, Fig. 1.3c and Fig. 1.3d from Einsle et al. 1999.



**Figure 1.4.** (a) The dimeric structure of  $\alpha$ -PEC is shown in blue and green in its  $\alpha_E$ -PEC form (PDB entry 2J96). The  $\alpha_Z$ -PEC form, PDB entry 2C7L, is not shown. An electron density for the phycoviolobilin (PVB) chromophore with pyrrole rings is shown. (b) Absorption spectra from  $\alpha$ -PEC obtained using the Schmidt Beamline 2 (SBL2). Dark absorption spectra for the  $\alpha_E$ -PEC state (blue) and  $\alpha_Z$ -PEC state (red) of  $\alpha$ -PEC in solution and transient (difference) spectra after a short (5 ns) laser pulse. All spectra are obtained from 60 accumulations of 50 ns exposures. Green spectrum: 100 ns after the  $\alpha_E$ -PEC species is excited by a laser pulse at 490 nm (cyan arrow). Cyan spectrum: 100 ns difference spectrum (green minus blue). Orange spectrum: 100 ns difference spectrum after the  $\alpha_Z$ -PEC species is excited by a laser pulse at 590 nm (orange arrow, spectral change not shown). Direction of transition is shown by black arrows. Fig. 1.4a and Fig. 1.4b modified from Purwar et al. 2013.

1.3a). The nitrogen cycle contains two parts: aerobic reductive (Fig. 1.3a, black arrow) and anaerobic reductive and oxidative pathways (Fig. 1.3a, red arrows) (Schmidt et al. 2004a; Kostera et al. 2010). CcNiR metabolizes nitrite through respiration and uses the extracted energy for the formation of Adenosine triphosphate (ATP) with no free NO or  $\text{NH}_2\text{OH}$  released during the reduction, as shown by the upward vertical red arrow and dashed black arrows in Fig. 1.3a. These intermediates remain bound to the active site heme, as proposed by density functional theory (DFT) calculations (Einsle et al. 2002; Bykov and Neese 2011; Bykov and Neese 2012). CcNiR is therefore an interesting candidate for a *mix-and-inject* X-ray FEL experiment that may reveal significant insights for regulation of the nitrogen cycle.

### 1.7 $\alpha$ -phycoerythrocyanin ( $\alpha$ -PEC)

The protein  $\alpha$ -phycoerythrocyanin ( $\alpha$ -PEC) is used to demonstrate the capabilities of the Schmidt Beamline 2 (SBL2) to record spectra from photo-reactive proteins which do not enter a



reversible photocycle.  $\alpha$ -PEC is a model photo-reactive switch that can be reversibly changed between the two stable forms,  $\alpha_E$ -PEC (Fig. 1.4a) and  $\alpha_Z$ -PEC. These forms have distinct spectra that can be prepared using light of different wavelengths (Fig. 1.4b). Unlike PYP, however, this reaction will not return unaided to the initial state after photo-excitation by a laser. To prepare the original species the proper wavelength of light must be repeatedly exposed to the sample following each laser pulse. The method for preparing these forms will be discussed later in detail (see §2.3.2).

$\alpha$ -PEC is isolated from *phycoerythrocyanin*, which is an antenna protein found in some cyanobacterial phycobilisomes (Sidler 1994). The PEC of *Mastigocladus laminosus* has been notably characterized (Duerring et al. 1990; Schmidt et al. 2006; Schmidt et al. 2007) and consists of a trimer of  $\alpha\beta$ -dimers. However, only the  $\alpha$ -subunit is photoactive (Hong et al. 1993). The  $\alpha_E$ -PEC form is prepared by light of wavelength 580 nm while the  $\alpha_Z$ -PEC form by blue light of 490 nm. The structures of both forms have been well characterized by X-ray crystallography (Duerring et al. 1990; Schmidt et al. 2006; Schmidt et al. 2007).  $\alpha$ -PEC is an interesting model for study due to the similarities with phycobilisome structures in other proteins. An XFEL is the ideal machine to probe these reaction intermediates on the ultrafast timescales.

## 1.8 Protein Microcrystals

Since the incredible brilliance of an XFEL destroys protein crystals after only a single shot, large quantities of microcrystals are currently needed for each experiment. The small sizes of microcrystals also provide the benefits of decreased diffusion times and larger levels of photo-reaction initiation. There are many proteins that prove challenging to grow large enough crystals suitable for a TR-LX experiment. CcNiR is one such example, however beautiful microcrystals are easily produced. Additionally, probing the structures of the intermediates in this irreversible enzymatic reduction of nitrite to ammonia is not currently possible at a synchrotron. An XFEL is

the only machine capable of performing this type of experiment and therefore the technique of microcrystal creation is an integral component of these 4<sup>th</sup> generation light sources.

### **1.9 Schmidt Lab (SLAB) Beamlines 1, 2 and 3 (SBL1, SBL2 and SBL3)**

A TR-SFX experiment requires enormous effort and success hinges upon thorough preparation: techniques for creating microcrystals must be developed, these samples should then be tested with a GDVN and initial studies must be performed to characterize these crystals. Since only 15% of all XFEL experiment proposals are ultimately accepted, previous results that strongly support such proposals significantly improve the chances for obtaining beamtime. I have, therefore, constructed three instruments for the Schmidt Lab (SLAB): 1) the *Schmidt Beamline 1* (SBL1) is a micro-focus X-ray diffraction beamline, 2) the *Schmidt Beamline 2* (SBL2) is a near ultraviolet / visual wavelength fast microspectrophotometer (Purwar et al. 2013) and 3) the *Schmidt Beamline 3* (SBL3) is a GDVN fabrication and testing facility. These machines supply the crucial initial information that is needed, not only for creating engaging XFEL beamtime proposals, but also in preparing for these experiments once beamtime has been awarded.

## 2 Materials and Methods

### 2.1 Monochromatic X-ray Crystallography

#### 2.1.1 Scattering from an Atom

When X-rays impinge on an atom they are scattered by the electrons. Since the proton is orders of magnitude heavier than an electron the nucleus is considered stationary during the scattering process. The probability of finding an electron at position  $\vec{r}$  is given by the electron density,  $\rho(\vec{r})$ . The electron density surrounding the nucleus scatters the X-rays within an infinitesimal volume element,  $dV$ . Scattering and then interference results from each partial wave. If just two volume elements are considered, as shown in Fig. 2.1a, then the path difference of the incoming waves from direction  $\hat{s}_0$  into the scattered direction  $\hat{s}_1$  is

$$\Delta p = \vec{r} \cdot \hat{s}_0 - \vec{r} \cdot \hat{s}_1 = (\hat{s}_1 - \hat{s}_0) \cdot \vec{r} = \vec{S} \cdot \vec{r} \quad (2.1)$$

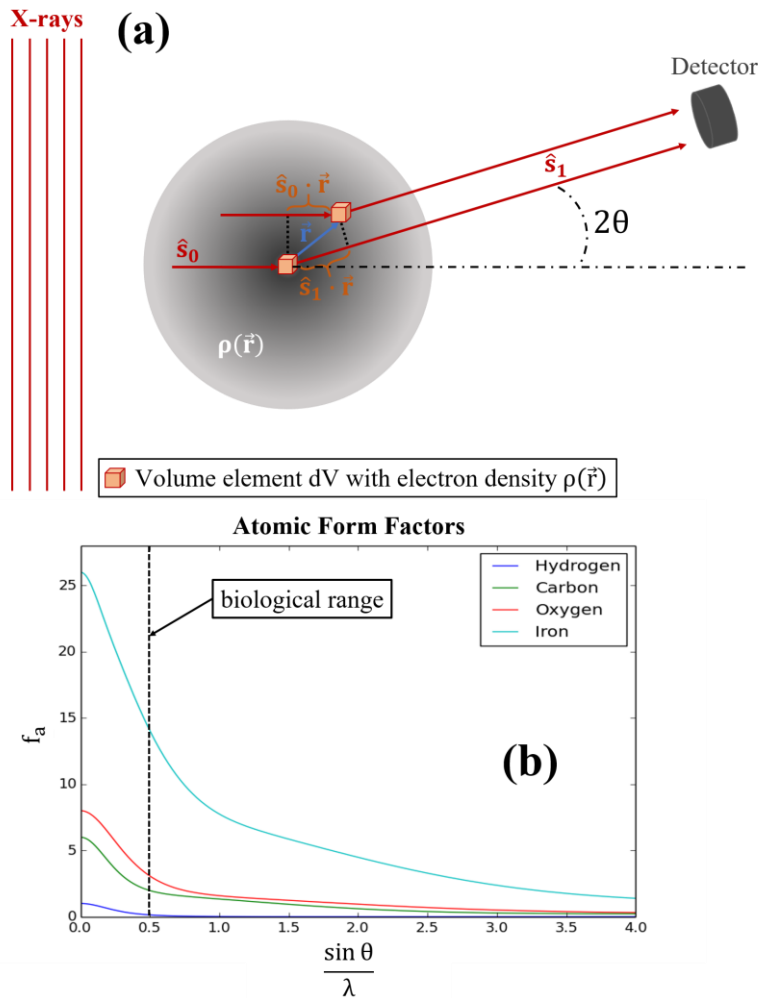
and the phase difference is

$$\Delta\varphi = \frac{2\pi\Delta p}{\lambda} = 2\pi\vec{H} \cdot \vec{r}, \quad \text{with } \vec{H} = \frac{\vec{S}}{\lambda} \quad (2.2)$$

The strength with which the electrons scatter the X-rays is proportional to the total number of electrons of the atom. Since the electron density is spherically symmetric in the first order approximation, the total partial wave emanating from the atom into the *scattering angle*  $2\theta$  (Fig. 2.1a) is obtained by integration over the electron density of the entire atom,

$$f_a = \sqrt{\sigma_e} \int_{V_{\text{atom}}} \rho(\vec{r}) e^{2\pi i \vec{H} \cdot \vec{r}} d\vec{r} \quad (2.3)$$

$f_a$  in Eqn. 2.3 is called the *atomic form factor*,  $\sqrt{\sigma_e}$  is the *Thomson scattering length* and  $\vec{H}$  is the



**Figure 2.1.** (a) Scattering of X-rays by an atom. Incoming X-ray plane waves with direction  $\hat{s}_0$  are scattered by an electron density,  $\rho(\vec{r})$ . Two infinitesimal volume elements,  $dV$ , are separated by a vector  $\vec{r}$ . These volume elements scatter the X-rays and the partial waves then interfere into the scattering direction  $\hat{s}_1$ . The resulting intensity is recorded on a detector. (b) Atomic form factors plotted for Hydrogen, Carbon, Oxygen and Iron. The dashed vertical line denotes the biological range.

scattering vector. Eqn. 2.3 is also recognizable as the Fourier equation, therefore the *atomic form factor* is the Fourier transform of the electron density of the atom. Fig. 2.1b contains plots of the *atomic form factors* for Hydrogen, Carbon, Oxygen and Iron using tables available from the International Tables of Crystallography (Prince 2006). The form factor equals the atomic number,  $Z$ , in the forward scattering direction.

### 2.1.2 Scattering from a Molecule

Molecules are constructed from an assembly of atoms, each scattering X-rays into the *scattering angle*  $2\theta$  according to their *atomic form factors*. By extension, the *structure factor of the molecule* may be calculated as

$$\mathbf{F}^M(\vec{H}) = \sum_{j=1}^N f_j e^{2\pi i \vec{H} \cdot \vec{R}_j} \quad (2.4)$$

where the sum is now taken over all atoms,  $N$ , in the molecule at positions  $\vec{R}_j$  with the  $j^{\text{th}}$  atom scattering with its *atomic form factor*,  $f_j$ . Importantly, the *structure factor of the molecule* is a complex number having both a magnitude and phase.

### 2.1.3 Scattering from a Crystal

Protein crystals are periodic structures with internal symmetry. If we divide the crystal into small subunits and place within these volumes a protein molecule, the result is called a *unit cell*. The entire crystal is then constructed from a finite periodic array of these *unit cells*. The *structure factor of the crystal* is given by

$$\mathbf{F}^C(\vec{H}) = \sum_{m=1}^M f_m e^{2\pi i \vec{H} \cdot \vec{R}_m} \quad (2.5)$$

where the sum is over all atoms,  $M$ , in the crystal. Since the crystal is periodic we can rewrite the vector in the exponential in two parts,

$$\mathbf{F}^C(\vec{H}) = \sum_{n=1}^N \sum_{j=1}^{\text{atoms}} f_j e^{2\pi i \vec{H} \cdot (\vec{R}_j + \vec{R}_{UC})} \quad (2.6)$$

where  $\vec{\mathbf{R}}_j$  is a vector to an atom within an arbitrary *unit cell* and  $\vec{\mathbf{R}}_{\text{UC}}$  is a vector to the origin of this particular *unit cell* from the origin of the coordinate system, as shown in Fig. 2.2. With  $\vec{\mathbf{R}}_m = \vec{\mathbf{R}}_j + \vec{\mathbf{R}}_{\text{UC}}$  we can separate this double sum into a product of two sums,

$$\mathbf{F}^c(\vec{\mathbf{H}}) = \sum_{n=1}^N e^{2\pi i \vec{\mathbf{H}} \cdot \vec{\mathbf{R}}_{\text{UC}}} \sum_{j=1}^{\text{atoms}} f_j e^{2\pi i \vec{\mathbf{H}} \cdot \vec{\mathbf{R}}_j} \quad (2.7)$$

The first sum is called the *lattice factor*,  $\mathbf{G}$ , and the second sum is recognizable as the *structure factor of the molecule*, now summed over all atoms in the *unit cell*. Since  $\vec{\mathbf{R}}_{\text{UC}}$  is a linear combination of the *unit cell vectors*  $\vec{\mathbf{a}}$ ,  $\vec{\mathbf{b}}$  and  $\vec{\mathbf{c}}$ , the *lattice factor* can be rewritten as a triple sum

where  $\mathbf{u}$ ,  $\mathbf{v}$  and  $\mathbf{w}$  are integers and  $\mathbf{U} \cdot \mathbf{V} \cdot \mathbf{W} = \mathbf{N}$ , where  $\mathbf{N}$  is the number of *unit cells* in the crystal,

$$\begin{aligned} \mathbf{G} &= \sum_{n=1}^N e^{2\pi i \vec{\mathbf{H}} \cdot \vec{\mathbf{R}}_{\text{UC}}} = \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} \sum_{w=0}^{W-1} e^{2\pi i \vec{\mathbf{H}} \cdot (u\vec{\mathbf{a}} + v\vec{\mathbf{b}} + w\vec{\mathbf{c}})} \\ &= \sum_{u=0}^{U-1} e^{2\pi i \vec{\mathbf{H}} u \cdot \vec{\mathbf{a}}} \sum_{v=0}^{V-1} e^{2\pi i \vec{\mathbf{H}} v \cdot \vec{\mathbf{b}}} \sum_{w=0}^{W-1} e^{2\pi i \vec{\mathbf{H}} w \cdot \vec{\mathbf{c}}} \end{aligned} \quad (2.8)$$

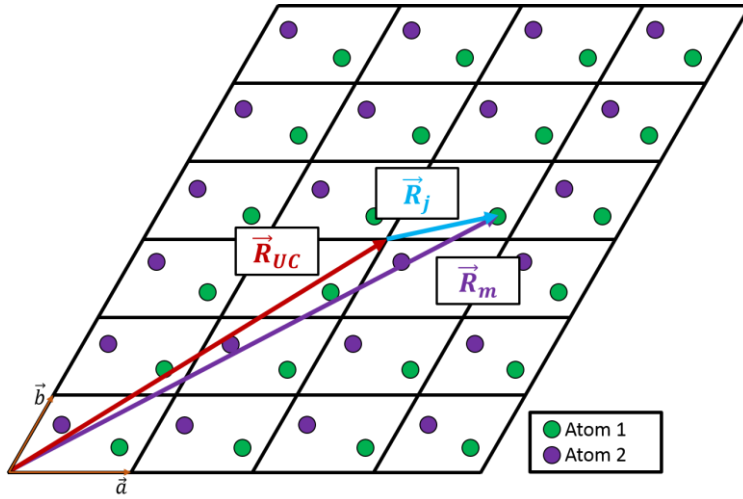
(U along  $\vec{\mathbf{a}}$ , V along  $\vec{\mathbf{b}}$  and W along  $\vec{\mathbf{c}}$ ). Each of these sums are generally zero. However, when

$$\vec{\mathbf{H}} \cdot \vec{\mathbf{a}} = h \quad (2.9a)$$

$$\vec{\mathbf{H}} \cdot \vec{\mathbf{b}} = k \quad (2.9b)$$

$$\vec{\mathbf{H}} \cdot \vec{\mathbf{c}} = l \quad (2.9c)$$

with  $\mathbf{h}$ ,  $\mathbf{k}$  and  $\mathbf{l}$  integers, these sums will be different from zero. Eqns. 2.9 are known as the *Laue conditions*.



**Figure 2.2.** A 2-dimensional crystal lattice. Each unit cell contains one molecule that is constructed from just two atoms. The vector  $\vec{R}_m$  can be separated into the two vectors  $\vec{R}_{UC}$  and  $\vec{R}_j$ . The unit cell vectors,  $\vec{a}$  and  $\vec{b}$ , are noted in the lower left.

## 2.1.4 The Reciprocal Lattice

From the *real space crystal lattice* we may also construct the *reciprocal lattice* in such a way that the *Laue conditions* will be automatically fulfilled. If a 3-dimensional *real space crystal lattice* is considered with unit cell vectors  $\vec{a}$ ,  $\vec{b}$  and  $\vec{c}$ , the reciprocal lattice vectors  $\vec{a}^*$ ,  $\vec{b}^*$  and  $\vec{c}^*$  may

be calculated as

$$\vec{a}^* = \frac{\vec{b} \times \vec{c}}{\vec{a} \cdot (\vec{b} \times \vec{c})} \quad (2.10a)$$

$$\vec{b}^* = \frac{\vec{c} \times \vec{a}}{\vec{a} \cdot (\vec{b} \times \vec{c})} \quad (2.10b)$$

$$\vec{c}^* = \frac{\vec{a} \times \vec{b}}{\vec{a} \cdot (\vec{b} \times \vec{c})} \quad (2.10c)$$

The scattering vector,  $\vec{H}$ , may now be expressed as a linear combination of the *reciprocal lattice vectors* with the *Laue integers* as the constants,

$$\vec{H}_{hkl} = h\vec{a}^* + k\vec{b}^* + l\vec{c}^* \quad (2.11)$$

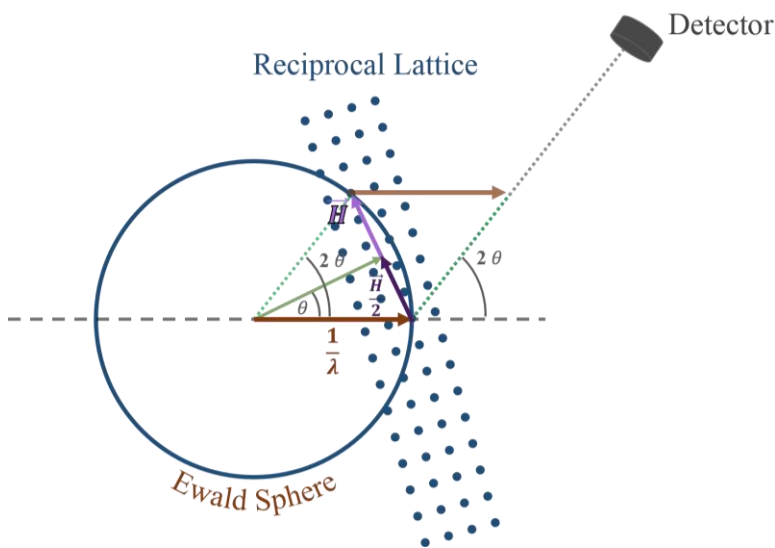
Since  $\vec{a}^*$ ,  $\vec{b}^*$  and  $\vec{c}^*$  are generally not mutually orthogonal,

$$|\vec{H}|_{hkl} = \sqrt{(h\vec{a}^*)^2 + (hk\vec{a}^* \cdot \vec{b}^*) + (hl\vec{a}^* \cdot \vec{c}^*) + (kh\vec{b}^* \cdot \vec{a}^*) + (k\vec{b}^*)^2 + (kl\vec{b}^* \cdot \vec{c}^*) + (lh\vec{c}^* \cdot \vec{a}^*) + (lk\vec{c}^* \cdot \vec{b}^*) + (l\vec{c}^*)^2} \quad (2.12a)$$

One can also show that

$$|\vec{H}|_{hkl} = \frac{1}{d_{hkl}} \quad (2.12b)$$

where  $d_{hkl}$  is the spacing between *reciprocal lattice planes* with *Miller Indices* (hkl). Readily apparent is the observation that for higher-order reflections (i.e.,  $\vec{H}_{hkl} = nh\vec{a}^* + nk\vec{b}^* + nl\vec{c}^*$ , n an integer),  $|\vec{H}|_{hkl}$  is longer and  $d_{hkl}$  is correspondingly shorter. Eqns. 2.10-2.12 may be described schematically as shown in Fig. 2.3 by a *reciprocal lattice* construction.



**Figure 2.3.** The Ewald Sphere has a radius of  $1/\lambda$ . The Laue conditions state that only those scattering vectors which lie on the Ewald Sphere will lead to Bragg spots at the detector in the far field. The scattering angle is  $2\theta$ . If we consider the vector  $\vec{H}/2$ , then Bragg's equation may be derived.

An assumption is made that the X-ray energy is conserved (i.e., the interaction is elastic), implying that the wavelength does not change during the scattering process, hence  $\lambda_{in} = \lambda_{out}$ . When radiation of wavelength  $\lambda$  scatters from the *reciprocal lattice* of a crystal, then the *Laue conditions* imply that only those scattering vectors



$\vec{H}_{hkl}$  which lie on the so-called *Ewald Sphere* will produce a scattering pattern at the detector in the far field. The radius of the *Ewald Sphere* is  $1/\lambda$  and the *scattering angle* is  $2\theta$ . If we now consider  $\vec{H}_{hkl}/2$ , then Bragg's equation can be derived by trigonometry,

$$\sin \theta = \frac{\frac{|\vec{H}|_{hkl}}{2}}{\frac{1}{\lambda}} = \frac{\lambda |\vec{H}|_{hkl}}{2} \quad (2.13)$$

or,

$$\frac{2 \sin \theta}{|\vec{H}|_{hkl}} = \frac{2 \sin \theta}{\frac{1}{d_{hkl}}} = 2d_{hkl} \sin \theta = \lambda \quad (2.14)$$

The *Laue conditions* combined with the *elastic scattering condition* therefore leads to Bragg's equation,

$$\frac{2 \sin \theta}{n\lambda} = \frac{1}{d} \quad (2.15)$$

where  $\mathbf{d}$  is the spacing between reflection planes and  $\mathbf{n}$ , an integer. Fig. 2.3 also provides a straightforward means to determine the *physical resolution limit*. In Eqn. 2.15, we see that the minimum resolution is obtained when  $\sin \theta$  is a maximum (e.g.,  $\theta = \pi/2$ ) so that

$$d_{\min} = \frac{\lambda}{2} \quad (2.16)$$

This implies that any *reciprocal lattice points* which lie further away than  $2/\lambda$  from the origin of the *reciprocal lattice* will not cut through the *Ewald Sphere* and will never produce Bragg spots. In practice, however, the *physical resolution limit* is not realistically achievable since detectors usually do not cover  $4\pi$ . The *experimental resolution limit* is then set by the minimum sample-to-detector distance, the wavelength of the X-ray radiation and the size of the detector. A specific

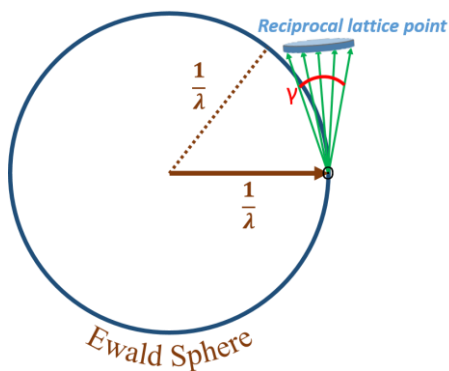
example of this constraint will be discussed later for SBL1. The general equation for determining the *experimental resolution* is

$$d = \frac{\lambda}{2 \sin\left(\frac{1}{2} \tan^{-1}\left(\frac{r}{L}\right)\right)} \quad (2.17)$$

where  $\lambda$  is the X-ray wavelength,  $r$  is the radius of the detector and  $L$  is the sample-to-detector distance.

### 2.1.5 Mosaicity and Structural Heterogeneity

Protein crystals are not perfect. During crystallization the molecules do not align themselves in an ideal periodic arrangement throughout the crystal but instead may form numerous mosaic blocks (Rupp 2009). A crystal is therefore not a single object but rather is a collection of these mosaic blocks, each with slightly different orientations. The *mosaicity* characterizes this long range disorder in the crystal. The effect of *mosaicity* is an angular broadening of the *reciprocal lattice points*, as seen in Fig. 2.4. This broadening leads to Bragg reflections that are observed through an angle of rotation,  $\gamma$ . Consequently, if the crystal is not rotated during an X-ray exposure then only a partial Bragg reflection will be observed. If the crystal is rotated through an angle of



**Figure 2.4.** A reciprocal lattice point has been broadened by the mosaicity,  $\gamma$ , of the crystal.

rotation,  $\phi$ , less than the *mosaicity* (i.e.,  $\phi < \gamma$ ), then a partial reflection will again be observed. However, the next diffraction images will record the remainder of the Bragg reflection and crystallographic software would properly integrate the full reflection intensities corresponding to the same *reciprocal lattice points*. Improper freezing during sample mounting can also increase the *mosaicity*.

Other factors contribute to a broadening or reduction of the measured Bragg intensities. For example, the beam divergence leads to a broadening of Bragg reflections. For lab X-ray sources the bandwidth is reasonably small but the beam divergence may be too large to resolve reflections for those protein crystals with large *unit cell* dimensions. An ideal crystal has *unit cells* in which the molecules not only have the same position but also the same orientation throughout the crystal. In reality, however, molecules are flexible constructs which leads to structural heterogeneity where atoms can be slightly shifted in each *unit cell* from their equilibrium positions. These types of disorder are called *static disorder* since the motion occurs on timescales slower than the time resolution of the experiment. *Dynamic disorder* describe effects that occur on timescales much faster, such as vibrations of the atoms about their equilibrium positions. These vibrations lead to a broadening of the observed electron density. Additionally, parts of the molecules may vary transiently between slightly different conformations.

All of the disorders described above, both static and dynamic, contribute to the Debye-Waller Factor (DWF) (Housley and Hess 1966)

$$\text{DWF} = \exp \left[ -8\pi^2 \langle x^2 \rangle \frac{\sin^2 \theta}{\lambda^2} \right] \quad (2.18)$$

where  $\langle x^2 \rangle$  is the mean square displacement,  $\theta$  is one half of the *scattering angle*,  $\lambda$  is the wavelength of the X-ray radiation and  $8\pi^2 \langle x^2 \rangle$  is the B-factor. One can see that the DWF is less than one except in the forward scattering direction when  $\theta$  is zero. Therefore, the scattered intensity is reduced for larger values of  $\theta$ . Because the DWF depends on  $\langle x^2 \rangle$  there is no explicit distinction between *static disorder* and *dynamic disorder*.

### 2.1.6 Polarization Factor, Lorentz Factor and the Integrated Scattered Intensity

Since X-rays are scattered by the electrons in an atom and the crystal is rotated during an exposure, this leads to correction factors that must be properly accounted for when integrating the reflection intensities. In addition to taking care that all partial reflections corresponding to a particular *reciprocal lattice point* are integrated as a single reflection, any data reduction software must also include corrections due to the *polarization factor* and *Lorentz factor*. When X-rays are diffracted the amplitude of the scattered radiation depends on the angle between the initial polarization and the scattered polarization directions as  $\cos 2\theta$ , where  $2\theta$  is the *scattering angle*. There are 3 cases to consider: (i) the incoming X-ray polarization is perpendicular to the polarization of the scattered radiation; (ii) the incoming X-ray polarization is coplanar with the polarization of the scattered radiation; (iii) the incoming X-rays are completely unpolarized and the amplitude of the scattered X-rays is a mixture of cases (i) & (ii). Since the intensity of the radiation goes as the square of the amplitude, the *polarization factor* is

$$\frac{\cos^2(2\theta) + 1}{2} \quad (2.19)$$

The *Lorentz factor* depends in part on the intersection of the diffraction volume of a reflection with the *Ewald Sphere*, which goes as  $1/\lambda^2$ . The time a *reciprocal lattice point* takes to cut through the *Ewald Sphere* is

$$\frac{\lambda}{\omega \sin(2\theta)} \quad (2.20)$$

and the *Lorentz factor* becomes

$$\frac{\lambda^3}{\omega \sin(2\theta)} \quad (2.21)$$

Combining all of these correction terms the intensity scattered by a crystal from incoming X-rays of wavelength  $\lambda$  and intensity  $I_0$  into the *scattering angle*  $2\theta$  is found to be

$$I_{h,k,l} = \frac{I_0 \sigma_e N \lambda^3 \cos^2(2\theta) + 1}{\omega \cdot 2 \sin(2\theta)} |F^M(\vec{H})|^2 \quad (2.22)$$

where  $\sigma_e$  is the Thomson scattering cross section and  $|F^M(\vec{H})|^2$  is the amplitude of the *structure factor of the molecule squared* (Rupp 2009).

### 2.1.7 Single Crystal Sample Preparation

For structure determination protein crystals must be produced. The protein is grown using a suitable expression system and then harvested from these bacteria. To select (purify) the desired protein, chromatographic methods are employed. A detailed description of these methods is beyond the scope of this dissertation and will not be discussed. However, the general process for producing large single crystals of PYP and ccNiR is now presented.

PYP is overexpressed and purified as reported elsewhere (Borgstahl et al. 1995). Large single crystals of PYP are grown using Na-malonate as a precipitant. Hanging drops are formed by mixing 5  $\mu$ L of 35 mg/mL PYP with 2.7 mol/L Na-malonate (pH=7). Crystals which are large enough for a single crystal monochromatic X-ray diffraction experiment normally grow within a week.

The overexpression and purification of ccNiR is described elsewhere (Youngblut et al. 2012). To create suitable large single crystals of the native ccNiR form, 50 mmol/L of *triethanolamine* (TEA) is used as buffer with 18% *polyethylene glycol* (PEG) 4k (pH=8.25). Crystal drops are mixed using a 1  $\mu$ L:1  $\mu$ L ratio with the purified protein and the plates are stored at 4 °C. Hanging drops are used and crystals typically form after 7 days.

Protein overexpression and purification for the nitrite-bound form of ccNiR is identical to that described above for the native form. Large single crystals are grown in a 50 mmol/L buffer of TEA (pH=8.25) with 18% PEG 4k and 20 mmol/L Na-nitrite. The solution is mixed in a 1  $\mu$ L:1  $\mu$ L ratio with the purified ccNiR. The hanging drop method is used and large single crystals usually develop in 7-14 days.

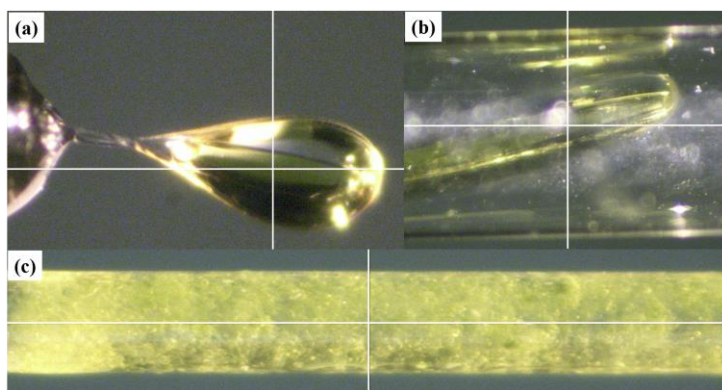
### 2.1.8 Single Crystal Sample Mounting

Once high quality crystals of large enough size have been produced, the 3-dimensional structure may be determined by single crystal monochromatic X-ray diffraction methods. General techniques for mounting samples for these experiments are presented here. The SBL1 and BioCARS 14-ID-B beamline can accommodate sample mounting methods that use either a standard brass specimen pin (Hampton Research, HR4-661) or a magnetic base. The brass specimen pin has the dimensions 12.8 mm length, 3.2 mm *outer diameter* (OD) and 1.4 mm *inner diameter* (ID). The magnetic base has a magnetic platform with an OD of 9.7 mm and a mounting post of 12.8 mm length and 3.2 mm OD which mounts to the goniometer in the same manner as the brass specimen pin.

Single crystals may be mounted in cryoloops (Fig. 2.5a) or in a glass capillary (Fig.2.5b) attached to a brass specimen pin. Powder samples may also be used when mounted in glass capillaries, as seen in Fig. 2.5c. To mount a sample in a cryoloop, a crystal is first harvested from the crystallization tray and placed into a ~2-4  $\mu$ l drop of an appropriate cryobuffer on a cover slide. These cryobuffers will vary, but for PYP a mixture of 10% glycerol, 2.8 mol/L Na-malonate and 20% sucralose would be used. For collecting images of the native form of ccNiR the cryobuffer consists of a mixture of 20% PEG 8k, 10% Glycerol and 50 mmol/L TEA. These cryobuffers are used to reduce ice formation and protect the crystal during the freezing process as the samples are

normally maintained at 100 K by a nitrogen cryostream. A cryoloop is used to collect the crystal from the cryobuffer and is then mounted immediately and swiftly onto the goniometer using the magnetic base so that the crystal is flash frozen by the cryostream.

Glass capillary mounted crystals are typically used to collect room temperature data. An Ø1 mm OD borate silica glass capillary is first mounted inside an Ø1.4 mm ID brass specimen pin and glued in place with epoxy. Once the epoxy has cured the capillary is cleaved at both ends using a tweezer. One end of the capillary is clipped flush with the brass specimen pin. To determine where to cleave the other end the capillary is mounted onto the goniometer and the approximate location of the X-rays on the capillary is indicated using a fine tip marker. The capillary may now be trimmed with the tweezers a few centimeters past this mark. A crystal is then taken from the crystallization tray and placed into a ~2-4 µl drop of mother liquor. A pipette tip is connected with a short piece of rubber tubing to the brass specimen pin. The crystal is then transferred out of the drop into the capillary using the pipette and carefully positioned using a glass fiber at the pre-marked location that will place it into the X-ray beam once mounted onto the goniometer. Using finely cut strips of coffee filter paper the excess liquid is wicked away, leaving only a thin layer



**Figure 2.5.** SBL1 sample mounting methods. (a) PYP crystal mounted in a cryoloop. (b) PYP crystal mounted in a capillary. (c) PYP microcrystal powder sample mounted in a capillary. The white crosshairs in each panel denotes the X-ray interaction region.

surrounding the crystal. A small amount of mother liquor is placed at the end of the capillary to prevent drying of the sample before both ends are sealed using epoxy. Powder samples are created in a similar manner as the glass capillary method, except instead of a large

single crystal a dense mixture of small microcrystals is transferred into the capillary and the ends are then sealed as before with epoxy.

### 2.1.9 Single Crystal Data Collection

A general data collection strategy for obtaining 2-dimensional diffraction images from large single crystals at cryogenic temperatures using monochromatic X-ray diffraction (e.g., the BioCARS 14-BM-C instrument and SBL1 beamline) will be described here. Diffraction images are collected using the *rotation method* (Bragg and Bragg 1949). The sample is aligned into the X-ray beam at the rotation axis and during the X-ray exposure the sample is continuously rotated through an angle  $\phi$ , typically  $0.1^\circ$ - $0.3^\circ$ . The diffraction pattern is read out and images are collected until the *reciprocal space* has been sufficiently sampled. The total angular rotation needed depends on the symmetry of the crystal. For a crystal with no symmetry (except Friedel symmetry), a total of  $180^\circ$  rotation would be the minimum required to completely sample the *reciprocal space* once. If each of the images spanned a rotation width of  $\phi = 0.2^\circ$ , for instance, then a total of 900 images would need to be recorded.

### 2.1.10 Single Crystal Data Reduction

Once a complete monochromatic dataset has been recorded the data must be reduced and an interpretable *electron density map* (EDM) produced. EDMs delineate the *real space* average positions of the electrons that created the diffraction patterns. EDMs will be discussed in more detail below. For the purposes of this section a structural model of the protein in question is assumed to be known. Methods to determine unknown phases are beyond the scope of this dissertation but can be determined by *multiple isomorphous replacement* (MIR) (Hendrickson 1985), *multiple anomalous dispersion* (MAD) (Hendrickson 1997) and *molecular replacement* (MR) (Mccoy et al. 2007; Mccoy 2007). Only the method of MR will be discussed. All of the



programs described here in *‘Italics’* are part of the *CCP4 Crystallographic Software Suite* (Bailey 1994; Winn et al. 2011; Krissinel 2015).

The diffraction images are first loaded into *‘iMosflm’* to perform spot finding, indexing, cell refinement and integration of reflection intensities. During spot finding reflections are located in the image(s) that satisfy user-adjustable parameters. These spots are then indexed by assigning *Miller Indices* (hkl) (see §2.1.4 and Eqn. 2.11) to the reflections, which define a set of planes orthogonal to the *scattering vector*  $\vec{H}_{hkl}$ . Since subsequent images are related to previous ones simply by a rotation of the angle  $\phi$ , manual indexing of every diffraction image is not necessary. Although one image is sufficient for determining an initial crystal orientation, in practice *‘iMosflm’* typically needs a small subset of the dataset (<10 exposures) for accurate indexing. A collection of indexing solutions is listed by *‘iMosflm’* from which the user selects the best orientation, spacegroup and initial *unit cell parameters*. The *mosaicity* is also estimated during indexing and is further refined as data reduction proceeds through the integration of reflection intensities.

Next, cell refinement is performed by comparing the predicted reflections with the observed reflections and minimizing the differences by adjusting the *unit cell parameters* and detector parameters. This geometric refinement process uses the *root mean squared deviations* (RMSD) and number of overlapped reflections as a basis for comparison between the predicted reflections of the refined parameters and the observed intensities until minimization is achieved. Finally, *‘iMosflm’* integrates the reflection intensities for each (hkl) and produces a single output file.

Once the integrated reflection intensities have been determined, *‘Scala’* (Evans 2006) merges the reflections and places them on the same scale and then *‘Truncate’* determines the *structure factor amplitudes*. The program *‘Freerflag’* is used to select a small subset (normally ~5%) of the *structure factor amplitudes* which will not be used during refinement of the structural model. This

5% portion of the data are later compared to all other reflections to assess if over-fitting of the data has occurred (Brunger 1992; Brunger 1993). Further discussion concerning refinement will be described in detail in §2.1.11.

If the observed *unit cell parameters* differ significantly from the known structural model then a MR strategy may be employed. The *structure factor amplitudes* determined by ‘*Truncate*’ and the known model are used to phase a new model with the program ‘*Phaser*’. The first step in MR is to determine the orientation of the known structural model in the crystal. Next, a translation determines the position of the model within the *unit cell*. If the MR is successful, ‘*Phaser*’ will output a new structural model of the protein with improved phases.

To generate an EDM the *structure factors* are needed, which requires both amplitudes and phases. To do this, the observed *structure factor amplitudes*,  $|F_{\text{obs}}(\text{hkl})|$ , are brought to an absolute scale by scaling them with the calculated *structure factor amplitudes* of the model,  $|F_{\text{calc}}(\text{hkl})|$  (Schmidt et al. 2010). The program ‘*SFALL*’ is used to determine *structure factors* from the *structure factor amplitudes* provided by ‘*Truncate*’ and the phases of the structural model,  $\phi(\text{hkl})$ , obtained from the PDB (or from the MR solution output by ‘*Phaser*’). Once the *structure factors* have been determined an EDM may be produced. Since the *structure factor* is the Fourier transform of the electron density distribution of the crystal the EDM may be calculated by the equation

$$\rho(\vec{R}) = \frac{1}{V} \sum_{\text{hkl}} |F_{\text{obs}}(\text{hkl})| e^{i\phi(\text{hkl})} e^{-2\pi i(\vec{H} \cdot \vec{R})} \quad (2.23)$$

where  $V$  is the volume of the *unit cell*. The program ‘*FFT*’ is used to calculate EDMs. Once an EDM has been determined a molecular visualization program such as *PyMOL* (DeLano and

Schrödinger), 'Coot' (Emsley et al. 2010) or *Chimera* (Pettersen et al. 2004) can display the map along with the structural model.

### 2.1.11 Assessment of Data Quality, Structural Refinement and Analysis

Some measures are generated by the program 'Scala' which guide the assessment of data quality, completeness, resolution ranges and internal consistency. One of these measures is the quantity  $R_{\text{merge}}$  (Evans 2006),

$$R_{\text{merge}} = \sum_{\mathbf{h}} \sum_m \frac{|I_{\mathbf{h}m} - \langle I_{\mathbf{h}} \rangle|}{\langle I_{\mathbf{h}} \rangle} \quad (2.24)$$

where  $I_{\mathbf{h}m}$  is  $m^{\text{th}}$  observation of the reflection  $\mathbf{h}$  and  $\langle I_{\mathbf{h}} \rangle$  are the average integrated intensities for all observed reflections of *Miller Indices* (hkl). Another coefficient,  $R_{\text{meas}}$ , may be calculated by the equation

$$R_{\text{meas}} = \sum_{n_{\mathbf{h}}} \left( \frac{n_{\mathbf{h}}}{n_{\mathbf{h}} - 1} \right)^{\frac{1}{2}} \frac{\sum_m |I_{\mathbf{h}m} - \langle I_{\mathbf{h}} \rangle|}{\sum_{\mathbf{h}} \sum_m \langle I_{\mathbf{h}} \rangle} \quad (2.25)$$

where  $n_{\mathbf{h}}$  are the number of observations of reflection  $\mathbf{h}$ . The values of  $R_{\text{merge}}$  and  $R_{\text{meas}}$  should be <10% overall and <50% in the high resolution shell. Other data quality measures include the completeness and multiplicity of the reflections. These quantities describe the amount of reciprocal space that has been sampled by the data and the average number of identical (or symmetry equivalent) reflections that have been observed, respectively. For high quality data the completeness should be >95% with a multiplicity > 3.

The program 'Refmac' (Murshudov et al. 1997; Vagin et al. 2004) is normally used for structural refinement. The two most frequently employed refinement modes are the rigid-body and restrained refinement. In a rigid-body refinement the entire molecule is moved as one unit (the

individual atoms are not moved independently) and maximally overlapped with the observed structure factors. During a restrained refinement the structural model is fitted to the observed structure factors while keeping the geometry of the model within an acceptable parameter range defined by empirical factors called the restraints.

A measure of the quality of the refined model is given by comparing the values of the correlation coefficients  $R_{\text{cryst}}$  and  $R_{\text{free}}$ . Both of these quantities are calculated using the same equation. However,  $R_{\text{free}}$  uses only the 5% subset of the reflections that were flagged earlier by 'Freerflag' which were not used for refinement.  $R_{\text{cryst}}$  is calculated from the other 95% of the reflections which were used during refinement.  $R_{\text{cryst}}$  and  $R_{\text{free}}$  may be calculated by the equation,

$$R_{\text{cryst}} = R_{\text{free}} = \sum_{\mathbf{h}} \frac{||F_{\mathbf{h}}| - |F_{\text{calc}}||}{\sum_{\mathbf{h}} |F_{\mathbf{h}}|} \quad (2.26)$$

where  $|F_{\mathbf{h}}|$  are the observed *structure factor amplitudes* of reflection  $\mathbf{h}$  and  $|F_{\text{calc}}|$  are the calculated *structure factor amplitudes* from the refined model.

Refinement typically begins by manually removing the waters from the known structural model obtained from the PDB. A rigid-body refinement is performed until the values of  $R_{\text{cryst}}$  and  $R_{\text{free}}$  do not change appreciably from cycle to cycle. A restrained refinement is then run until the values of  $R_{\text{cryst}}$  and  $R_{\text{free}}$  converge and again remain largely unchanged. A water search is performed using 'Coot' on the so-called *difference electron density* (DED) map produced by 'Refmac' and these molecules are then added to the refined model. A final restrained refinement is run until either  $R_{\text{cryst}}$  and  $R_{\text{free}}$  converge or once more do not vary significantly.

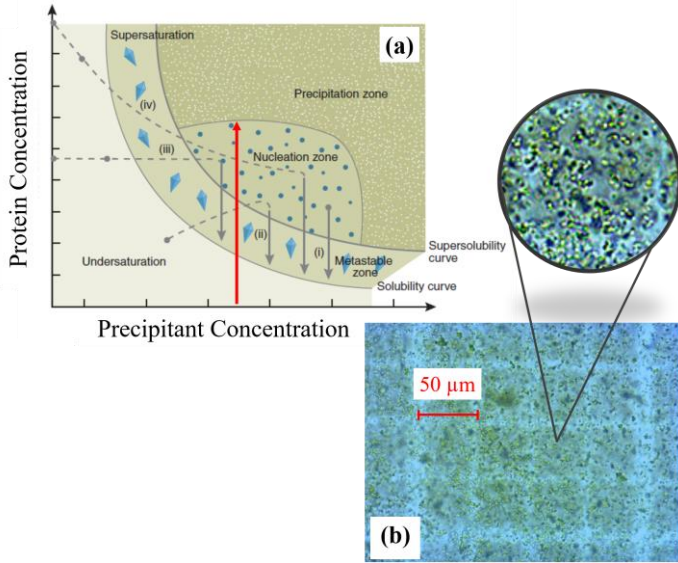
To determine if the refinement is successful, 'Coot' is used to check the fit of the refined structural model with the observed electron density. Upon inspection, if there are residues found

to be misaligned with the EDM then these may be reoriented using tools provided in 'Coot'. If any manual model corrections are made then a restrained refinement should again be performed. Once refinement is complete, the final values of  $R_{\text{cryst}}$  and  $R_{\text{free}}$  are compared to assess whether  $|R_{\text{cryst}} - R_{\text{free}}| \leq 0.07$ , which will be the case for excellent to average quality data. 'Coot' also provides the capability to measure atomic distances and angles within the model.

## 2.2 Time-Resolved Serial Femtosecond Crystallography (TR-SFX)

### 2.2.1 TR-SFX: Microcrystal Samples

For SFX experiments, microcrystals of sufficient quality and quantity must be produced. The phase diagram shown in Fig. 2.6a outlines the theory for the formation of protein crystals (Bergfors 1999; McPherson 1944- 1999). These phase zones include the undersaturation zone, supersaturation zone, precipitation zone, nucleation zone and metastable zone (Chayen and Saridakis 2008). Although computer modeling and high-throughput robotic screening facilities may accelerate a mapping of the phase diagram (Rupp and Wang 2004), trial-and-error remains an effective method for determining the optimum conditions for crystal growth. The red line in Fig. 2.6a demonstrates one possible path for microcrystal creation. In this method the precipitant concentration is held fixed while the protein concentration is quickly increased into the boundary between the nucleation and precipitation zones. This process leads to the creation of a large number of crystals. These crystals then grow until an equilibrium is achieved in the metastable zone. The size of the subsequent crystals can therefore be controlled by limiting the volume of the available protein surrounding these sites. The mechanism for this control is the adjustment of the density of nucleation sites. Although this technique is generally applicable to almost any protein, trial-and-error is required to determine the optimum pH, precipitant and protein concentrations.



**Figure 2.6.** Microcrystal formation. (a) The Protein crystallization phase diagram guides the selection of protein and precipitant concentrations. The red line represents a possible trajectory for microcrystal formation. (b) PYP microcrystals of  $\sim 2 \times 2 \times 10 \mu\text{m}^3$  are shown here in a Neubauer counting chamber, along with a close-up detail. The faint white squares are  $50 \times 50 \mu\text{m}^2$  in size and are used to estimate the crystal size and the total number of microcrystals per unit volume. A microcrystal concentration of  $10^{10}$ - $10^{11}$  microcrystals/mL is usually sufficient for an XFEL experiment. Fig. 2.6a from Chayen et al. 2008.

A protocol for growing PYP microcrystals is described here. This method was used for both the nanosecond (ns) and fs PYP experiments. Although this procedure is optimized for PYP the general technique is applicable for other proteins, including ccNiR and  $\alpha$ -PEC. The protein volumes used in this protocol are intended for large scale production of microcrystals during an XFEL experiment. However, these volumes are scalable to practically any desired size.

$\sim 15$  mL total volume of purified PYP protein (30-40 mg/mL) is combined from multiple Eppendorf tubes into one 15 mL centrifuge tube. The initial concentration of the combined PYP ( $C_{\text{PYP},i}$ ) is determined by diluting a  $\sim 10 \mu\text{L}$  volume of the solution by an appropriate factor (e.g.,  $D_{\text{fac}} = 10$ ) and measuring the absorption at the absorption maximum wavelength for PYP in solution of 446 nm ( $A_{446}$ ) with a Thermo Scientific NanoDrop 2000c spectrometer.  $C_{\text{PYP},i}$  is calculated using the equation

$$C_{\text{PYP},i} = A_{446} \times \frac{14,700 \text{ [g/mol]}}{45,500 \text{ [cm}^2/\text{mmol]}} \times D_{\text{fac}} \times \frac{1}{L} \quad (2.27)$$

where  $L$  is the path length of the spectrometer (1 mm for the NanoDrop). Using standard conversion values,  $C_{PYP,i}$  is expressed in units of [mg/mL]. The desired final PYP concentration is  $C_{PYP,f} = 135$  mg/mL. The concentration factor ( $C_{fac}$ ) is determined by the equation

$$C_{fac} = \frac{C_{PYP,f}}{C_{PYP,i}} \quad (2.28)$$

The ~15 mL of PYP is distributed equally into 4 centrifuge concentrators and the final protein volume is determined by taking the initial volume in each concentrator and dividing by  $C_{fac}$ . The PYP solution is then centrifuged at 8000 g at 22 °C for ~ 11 minutes until the final volume,  $V_{PYP,f} = V_{PYP,i}/C_{fac}$ , is obtained. Combining the PYP from all spin filters into one 15 mL centrifuge tube, 100  $\mu$ L of the Na-citrate run through from one of the spin filters is used to rinse all of the membranes (each of the spin filters is rinsed with this same 100  $\mu$ L Na-citrate run through). The 100  $\mu$ L PYP rinse solution is added to the 15 mL centrifuge tube to harvest all possible PYP. The final concentration of the combined PYP solution ( $C_{PYP,f}$ ) is determined by diluting a ~10  $\mu$ L volume of the the solution by an appropriate factor (e.g.,  $D_{fac} = 100$ ) and again measuring the absorption at 446 nm ( $A_{446}$ ) with a spectrometer.  $C_{PYP,f}$  is calculated in the same manner as  $C_{PYP,i}$  using Eqn. 2.27 with the new values for  $A_{446}$  and  $D_{fac}$ . If  $C_{PYP,f}$  is too large then the solution is diluted with the Na-citrate buffer run through. Conversely, if  $C_{PYP,f}$  is too small then centrifugation is continued until the target  $C_{PYP,f}$  is obtained.

The concentrated PYP is divided into small centrifuge concentrators in  $V_{PYP} = 100$   $\mu$ L each and a volume of Na-malonate (pH=7) is added into each concentrator given by the formula

$$V_{malonate} = V_{PYP} \times \frac{C_{malonate,desired}}{(C_{malonate,stock} - C_{malonate,desired})} \quad (2.29)$$

where  $C_{\text{malonate,desired}}$  is the desired Na-malonate concentration (3.2 mol/L for this protocol) and  $C_{\text{malonate,stock}}$  is the stock Na-malonate concentration (4.45 mol/L for this protocol). All spin concentrators are now vortexed until everything is thoroughly mixed (the solution should be entirely cloudy). The spin filters are kept at room temperature for one hour after adding the Na-malonate. The samples are then centrifuged at 10,000 g for 2 hours at 22 °C. Next, the spin filters are vortexed vigorously and stored at room temperature for 4 hours. The microcrystal mixture from all of the concentrators is then distributed equally into Eppendorf tubes and the concentrator membranes are rinsed with 100  $\mu\text{L}$  of 3 mol/L Na-malonate to harvest all of the sample. The Eppendorf tubes are then spun down at 20,000 g for 10 minutes (the PYP microcrystals float to the top and the solution below should be nearly clear). Using a needle and 1 mL syringe (or a pipette with a gel loading tip) as much as possible of the clear solution below is transferred out and 800  $\mu\text{L}$  of 3 mol/L Na-malonate is added into each Eppendorf tube and vortexed thoroughly. The samples are then filtered with a 20  $\mu\text{m}$  stainless steel filter into a 15 mL centrifuge tube and vortexed one last time. Immediately before loading the samples into a reservoir for the experiment the samples are vortexed and filtered through a 10  $\mu\text{m}$  stainless steel frit.

### 2.2.2 TR-SFX: $T_0$ Determination and the Timing Tool

The quantity  $T_0$  designates the temporal overlap between the arrival time of the X-ray and laser pulses. The determination of  $T_0$  for a ns XFEL experiment is straightforward and may be accomplished by using a photo-diode placed at the sample position to record the arrival of the laser pulses using an oscilloscope. These traces are then compared to the traces from the LCLS timing signals. A delay called the *ns offset time* (discussed in detail below) is varied until temporal overlap of the two signals is achieved at the sample position. This particular *ns offset time* now defines  $T_0$  to ns accuracy. This timing method was used for the ns PYP experiment.



When experiments at the LCLS probe time-delays at the fs timescale then an instrument called the *timing tool* (Bionta et al. 2014; Hartmann et al. 2014) is used to accurately determine the delay between X-ray and laser pulses with single-digit fs precision for every snapshot. For TR-SFX experiments on photo-active proteins at ultrafast timescales the LCLS normally uses a *titanium sapphire* (Ti:sapph) Coherent Legend CPA system (Coherent Inc.) to generate 40 fs laser pulses at an 800 nm wavelength and 120 Hz. These pulses are then split by a beam splitter so that a portion of the pulse train is sent to the *timing tool* and the majority to a HE-TOPAS-Prime (Light Conversion Ltd.) parametric amplifier (if a wavelength other than 800 nm is desired). The 40 fs Ti:sapph laser pulses (800 nm wavelength) that are sent to the *timing tool* are then chirped with glass for a final pulse duration of 1.5 ps. This method produces a stretched pulse where longer wavelengths arrive before shorter wavelengths. These pulses are then directed to a motorized delay stage before being sent through the 2  $\mu\text{m}$  thick silicon nitride ( $\text{Si}_3\text{N}_4$ ) target at the *timing tool*. The resulting spectrums are recorded at 120 Hz using an Acton SP2500 (Princeton Instruments) spectrometer.

Every X-ray pulse is also passed through the  $\text{Si}_3\text{N}_4$  target which causes the reflectivity of the material to change and the chirped laser pulses are more heavily absorbed. This increased absorption results in a dip of the spectrum which corresponds to a time-delay between X-rays and laser due to the chirped nature of the laser pulse. Using this timing information each diffraction pattern may be sorted into arbitrary *timing bins*, enabling a flexible data collection strategy where multiple time-delays may be collected from a single time-delay. The shot-to-shot jitter of the X-ray arrival time is 280 fs (Glowonia et al. 2010) and so for a particular time-delay,  $T_p$ , data may be conveniently separated into *timing bins* from  $T_p - 280$  fs to  $T_p + 280$  fs. The shortest width of a *timing bin* therefore only depends on how many images are obtained at a particular time-delay.

Further details concerning the *timing tool* will be presented below during the discussion of how  $T_0$  is determined for fs experiments.

The determination of  $T_0$  for a fs experiment is a much more involved process than that needed for a ns experiment. To completely describe the timing method three terms must first be explained: (i) the *ns offset delay* adjusts the relative time-delay between X-rays and laser pulses; (ii) the *ns target time* sets the desired time-delay between the laser and X-ray pulses; (iii) the *timing tool delay stage* is a motorized set of mirrors used to match the time-delay set by the *ns target time* to the *timing tool* so that the spectroscopic traces of the laser pulses through the timing target can be captured on the spectrometer.

The procedure for determining  $T_0$  and calibrating the *timing tool* is described here. A fast photo-diode is mounted outside of the evacuated chamber and pointed inward toward a YAG target that is placed at the sample location. The *ns target time* is set to 0.0 ns and the X-ray focal spot is defocused so that the spot diameter is slightly larger than the laser spot size ( $\text{\O}90 \mu\text{m}$  for the fs PYP experiment). The laser mirror is adjusted so that the X-ray and laser pulses are spatially overlapped on the YAG target. Photo-diode traces are examined with only the X-rays present at the YAG target and the amplitude of the curve is noted. Laser pulses are then started and the *ns offset time* is slowly adjusted until the amplitude of the diode trace experiences a minimum. For this *ns offset time* the X-ray and laser pulses are temporally overlapped at the YAG target and  $T_0$  has been determined to better than 130 fs. This *ns offset time* now defines the experimental time  $T_0$ .

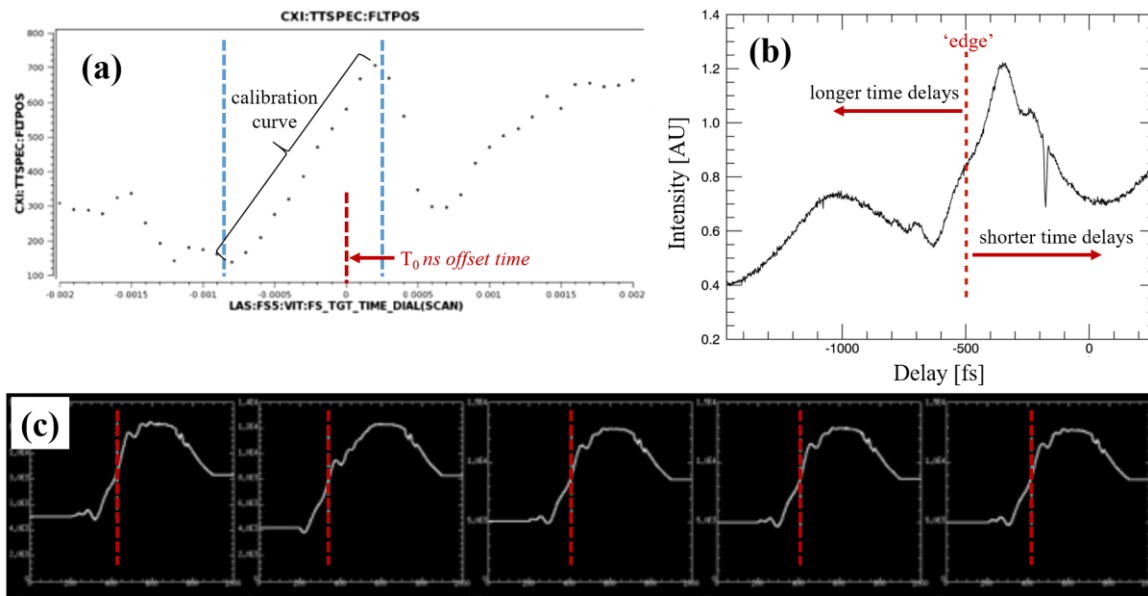
Keeping both the *ns offset time* and *ns target time* fixed, the *timing tool delay stage* is moved until the dip in the laser spectrum recorded by the spectrometer is positioned, on average, at pixel number 600. The current position of the *timing tool delay stage* is now correlated with the *ns offset*

time,  $T_0$ , determined previously by the photo-diode. Keeping the *timing tool delay stage* and *ns target time* fixed the *ns offset time* is now varied in fine time steps through a total of 1.5 ps, which is the width of the chirped laser pulse. At each new *ns offset time* the X-ray / laser delay is calculated by determining the pixel number of the midpoint of the rising edge of the spectrum, as shown in Fig. 2.7b. This pixel number is called the *edge pixel*. Approximately 500 of these spectrums are used and the average of the calculated *edge pixel* values for these images is determined for each *ns offset time*. These *edge pixels* now map directly to a particular *ns offset time* and, therefore, to a time-delay between the X-ray and laser pulses. The resulting *calibration curve* is shown in Fig. 2.7a. The *edge pixel* calculation for the fs PYP experiment was performed by Dr. Anton Barty by first subtracting the spectrum of the chirped laser pulse with no X-rays present from the ‘X-ray + laser’ spectrum. This difference spectrum was then convoluted with a digital filter signal provided by CXI beamline scientist Dr. Jason Koglin. The pixel number corresponding to the largest value of the resulting convoluted signal was then selected as the *edge pixel*. This method has also been described elsewhere (Barends et al. 2015).

Before data collection begins the *ns offset time* is returned to the  $T_0$  value determined by the photo-diode and the *ns target time* is changed to the desired experimental time-delay. The *timing tool delay stage* is also moved by  $\Delta d [\mu\text{m}] = \Delta t \frac{0.3 \mu\text{m}/\text{fs}}{2}$  so that the spectrum will be placed within the timing window of the spectrometer. If the chirped laser pulse is not properly recorded then no timing information will be available for that image. Each snapshot now includes the raw trace from the *timing tool*. From these spectra the *edge pixel* is calculated (Fig. 2.7c) and the *calibration curve* (Fig. 2.7a) is used to correlate this pixel number with an actual time-delay between the X-ray and laser pulses for every image. For the fs PYP experiment, I plotted each calibration curve and calculated a pixel gradient value as  $P_g = \Delta\text{time}/\Delta\text{pixel}$ . The pixel gradients were typically  $\sim 1.78$

fs/pixel. To determine the time-delay for each image the nominal pixel number of 600 was subtracted from the *edge pixel*. The result was then multiplied by the pixel gradient and added to the current *ns target time* to determine the actual time-delay between laser and X-ray for every image.

During a shift there are two different effects present in the X-ray beam that affect timing. The first is the shot-to-shot jitter of 280 fs discussed previously which actually works favorably for creating time-series covering an almost 600 fs total time span. The other effect is the long term drift of  $T_0$  that happens over the course of hours. This long term drift negatively impacts the recording of timing spectra from the *timing tool* since spectra may shift outside of the timing window once the relative delay between the laser and X-ray pulses becomes too large. The beamline scientist monitors the traces on the *timing tool* and when these begin to drift outside of the timing window they must be brought back so that the pixel number of the midpoint of the rising edge is placed once again at pixel number 600. This may be accomplished in one of three ways: (i) adjust the *ns offset time*; (ii) adjust the *ns target time*; (iii) move the *timing tool delay stage*. The first two of these methods does not affect time binning of the data and the *timing tool* pixel number may be used without any modification with one caveat: if a new time-delay is desired then the method described above must first be repeated to re-establish  $T_0$  and calibrate the *timing tool*. The preferred method to account for drift is to adjust the *ns target time*. The third method is not desirable since no correction is made to adjust for the drift. Unfortunately, all three of these methods were used during the fs PYP beamtime to adjust for the drift in  $T_0$ . The only solution was to go through the electronic log of the experiment for each data collection run so that the timing information from the *timing tool* could be interpreted appropriately.



**Figure 2.7.** Timing tool traces and analysis. (a) The ns offset time was varied in fine time steps during the timing tool calibration, scanning the chirped laser pulse. At each ns offset time a pixel number corresponding to the ‘edge’ of the resulting spectrum was calculated and 500 images were recorded to obtain an averaged edge pixel value. This averaged edge pixel value corresponds to a particular ns offset time. The result of this calibration is called the calibration curve. The linear portion delimited by the vertical dashed blue lines of this curve is the time width of the timing tool and is  $\sim 1.2$  ps. A time gradient is calculated as  $\Delta\text{time}/\Delta\text{pixel}$  and is usually 1.77 fs/pixel. The vertical dashed red line shows the location of the ns offset time corresponding to  $T_0$ . (b) A typical timing tool spectrum from a diffraction image is shown here and the vertical dashed red line denotes the calculated edge pixel, representing the relative time-delay between laser and X-ray. The value of zero on the horizontal axis is not necessarily  $T_0$  since the timing tool delay stage is moved when a ns target time is selected. (c) Edge pixel calculation plots are shown here from snapshots determined to be hits from the experiment provided by Anton Barty. The vertical dashed red lines again denote the calculated edge pixels. These calculated edge pixels represent the actual time-delay for each image.

### 2.2.3 TR-SFX: Data Collection

To record TR-SFX diffraction data with an XFEL, suitable microcrystals must first be produced (see §2.2.1). Initial tests should be performed on these microcrystals to determine the optimal parameters for GDVN construction and operation. At a synchrotron an entire time-series can be obtained quite easily by a single person. This is not currently possible at an XFEL. Tab. 2.1 lists the 15-24 people required to be present at all times during a 12-24 hour shift for a TR-SFX experiment. The *principal investigators (PIs)* manage the team and guide the experiment.

**Table 2.1.** TR-SFX experiment personnel.

Task	# of People	Role
Principle Investigator	1—2	manages team / guides experiment
Beamline Scientist	2—3	controls XFEL equipment
Session Logger	2—3	records detailed notes
GDVN Operation	3—4	run the injector
Sample Preparation	2—3	make microcrystals
Sample Loading	2—3	load samples into hydraulic reservoirs
Data Reduction	3—6	run hit-finding/index images/produce electron density maps

*Beamline scientists* are responsible for controlling the XFEL equipment. The *session logger* keeps accurate and detailed notes with a spreadsheet for each shift in addition to posting important comments and information on the experiment electronic log provided by SLAC. A *GDVN operation* team builds nozzles,

monitors nozzle jetting, controls sample flow rates and gas pressures and exchanges broken nozzles. Microcrystal samples are produced by the *sample preparation* group and the samples are loaded into the hydraulic reservoirs by those in charge of *sample loading*. The *data reduction* team runs hit-finding, refines experimental geometry, generates data statistics to guide the data collection strategy and produces EDMs. The atmosphere can be hectic, especially at the beginning of each shift when the experiment is starting and data collection strategies are being devised. Unforeseen issues also arise frequently throughout the shift, including XFEL beam stability and down time, nozzle clogging and breakage and beam sharing with other experiments. For time-resolved experiments on photo-active proteins the laser system (pump) must be installed and aligned onto the X-ray / sample interaction region. The microcrystals are loaded into hydraulic sample reservoirs and the pressure is supplied by either gas cylinders or a chromatography solvent delivery pump to drive the sample flow. Once a stable sample stream has been obtained and aligned into the X-rays data collection may proceed. The X-ray (probe) repetition rate of the LCLS is 120

Hz and images are collected on the *CSPAD* detector at this speed. The delay between the pump and probe is varied until a complete time-series has been obtained.

#### 2.2.4 TR-SFX: Data Reduction

An experiment at the LCLS generates large amounts of data. Potentially 5 million diffraction images can be collected during a single 12 hour shift. However, due to the nature of a TR-SFX experiment only a small fraction of these images actually contain a diffraction pattern from a crystal hit and even fewer of these hits can be indexed by the crystallographic software. The majority of images (>80%) are blank. With so much data the challenge to determine hits from no-hits is of primary importance. A software package named *Cheetah* (Barty et al. 2014) can quickly and accurately determine crystal hits from this large pool of data. This hit-finding is a crucial step in the pre-processing of the data.

Once images have been selected that contain diffraction patterns data reduction can continue. The software suite *CrystFEL* (White et al. 2012) contains many programs designed to process serial diffraction data. Data reduction proceeds with the program *Indexamajig* which runs spot finding, indexing, cell refinement and integration of reflection intensities. An additional required input for this program is the *geometry file* which contains information about the detector distance and detector pixel locations. *Indexamajig* is essentially the serial crystallographic version of '*iMosflm*'. In fact, the '*mosflm*' indexing algorithm is normally the most commonly used indexing method by *Indexamajig*. The output stream of *Indexamajig* is a text file which contains the results of the indexing and reflection integration from each image selected by *Cheetah* as a crystal hit. This reflection list has no organization and is simply output in the same order in which the images were processed by the software.

If the protein being studied possesses an *indexing ambiguity* then the program *Ambigator* (Brehm and Diederichs 2014) should now be run to resolve this ambiguity. For example, PYP crystals belong to the  $P6_3$  space group and possess hexagonal symmetry so special care must be taken during indexing to ensure that all snapshots are indexed in the same indexing convention. This two-fold ambiguity for the  $P6_3$  space group arises because both  $a^*$  and  $b^*$  have the same length and therefore some images may be indexed as (hkl) while others as (khl). Once each image has been successfully indexed using *Indexamajig* and the indexing ambiguity solved by *Ambigator* the output stream file is fed into *Process\_hkl* to merge the intensities from multiple observations of the same reflections into a single reflection list via Monte Carlo integration (Kirian et al. 2011). The data is now scaled with '*Scala*' and the *structure factor amplitudes* are determined by '*Truncate*'.

For time-resolved experiments, *structure factors* must be determined for the 'light' data (with protein activation) and the 'dark' data (without protein activation). These *structure factor amplitudes* are denoted as  $|F_{\text{obs}}(\text{hkl}, t)|$  and  $|F_{\text{obs}}^{\text{D}}(\text{hkl})|$ , respectively. The observed dark *structure factor amplitudes*,  $|F_{\text{obs}}^{\text{D}}(\text{hkl})|$ , are scaled with the calculated dark *structure factor amplitudes*,  $|F_{\text{calc}}^{\text{D}}(\text{hkl})|$ , to bring them to an absolute scale (Schmidt et al. 2010). The observed time-dependent *structure factor amplitudes*,  $|F_{\text{obs}}(\text{hkl}, t)|$  are then also scaled to the observed dark *structure factor amplitudes*,  $|F_{\text{obs}}^{\text{D}}(\text{hkl})|$ . Once all observed *structure factor amplitudes* have been brought to an absolute scale, the *difference structure factor amplitudes*,  $|\Delta F_{\text{obs}}(\text{hkl}, t)|$ , are calculated as

$$|\Delta F_{\text{obs}}(\text{hkl}, t)| = |F_{\text{obs}}(\text{hkl}, t)| - |F_{\text{obs}}^{\text{D}}(\text{hkl})| \quad (2.30)$$

The *weighted difference structure factor amplitudes*,  $w|\Delta F_{\text{obs}}(\text{hkl}, t)|$ , are determined by multiplying the *difference structure factor amplitudes* with a *weighting factor*  $w$  (Ren et al. 2001),



$$w = \frac{1}{1 + \frac{\sigma^2}{\langle \sigma^2 \rangle} + \frac{|\Delta F_{\text{obs}}(\text{hkl}, t)|^2}{\langle |\Delta F_{\text{obs}}(\text{hkl}, t)|^2 \rangle}} \quad (2.31)$$

where  $\sigma$  is the standard deviation of  $|\Delta F_{\text{obs}}(\text{hkl}, t)|$ . For those *difference structure factor amplitudes* which are large or for those with considerable differences, the *weighting factor* ensures that these are appropriately down-weighted. Also, the *average weighting factor* is used to normalize the  $|\Delta F_{\text{obs}}(\text{hkl}, t)|$  so that they remain on an absolute scale.

The phases from the structural model,  $\phi^{\text{D}}(\text{hkl})$ , are now combined with the *weighted difference structure factor amplitudes*,  $w|\Delta F_{\text{obs}}(\text{hkl}, t)|$ , to determine the *weighted time-dependent difference electron density*,  $\Delta\rho(\mathbf{t})$ ,

$$\Delta\rho(x, y, z, t) = \frac{1}{V} \sum_{\text{hkl}} w|\Delta F_{\text{obs}}(\text{hkl}, t)| e^{i\phi^{\text{D}}(\text{hkl})} e^{-2\pi i(\vec{H} \cdot \vec{R})} \quad (2.32)$$

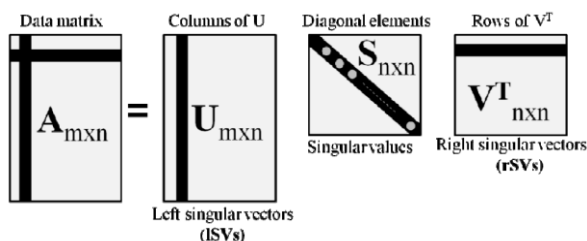
The so-called *difference approximation* states that if the structural differences are small then a (weighted) *time-dependent difference electron density map* (wDED<sub>t</sub>) relates directly to the fractional concentration (or occupancy) of states at that time delay. These DED maps contain positive and negative DED features. Positive DED features are normally pictured in blue or cyan and denote regions where electron densities are measured which are not present in the known structural model. Conversely, negative DED features are typically drawn in red or white and denote regions where the electron densities are not observed but which are present in the known model.

## 2.2.5 Singular Value Decomposition (SVD) of Time-Dependent DED Maps

The *weighted time-dependent difference electron density* maps are analyzed by *singular value decomposition* (SVD) (Henry and Hofrichter 1992) using the program SVD4TX (Schmidt et al. 2003; Zhao and Schmidt 2009). Regions of the  $wDED_t$  maps where large DED features are expected (e.g., around the chromophore) are masked for the analysis. All other regions mainly contribute noise. The matrix  $\mathbf{A}$  is created from the masked data and is decomposed by SVD into matrices of *left singular vectors* (ISVs), *right singular vectors* (rSVs) and *singular values* (see Fig.2.8). This decomposition is written as

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (2.33)$$

The ISVs contain the time-independent structural information. The corresponding rSVs are comprised of the time-dependent structural variations. The number of singular values present in the matrix  $\mathbf{S}$  suggests the total number of significant rSVs. Normally, only a few significant rSVs exist and the rest only contain noise. The first ISV represents an average of the DED maps while the first rSV contains its time-dependence. The second ISV covers the deviation of the DED maps



**Figure 2.8.** Visual representation of an SVD analysis. The data matrix  $\mathbf{A}$  is constructed from the masked region of the weighted time-dependent difference electron density map. The data matrix  $\mathbf{A}$  (dimension  $m \times n$ ) is then decomposed into a matrix  $\mathbf{U}$  (dimension  $m \times n$ ) of left singular vectors, a diagonal matrix  $\mathbf{S}$  (dimension  $n \times n$ ) containing singular values and a matrix  $\mathbf{V}^T$  (dimension  $n \times n$ ) of right singular vectors. Fig. 2.8 from Purwar 2013.

from the average whereas the second rSV describes its time-variation. Importantly, the number of significant vectors is representative of the number of kinetic processes present in the reaction. The SVD analysis for the fs TR-SFX PYP experiment was performed by Prof. Schmidt.

## 2.2.6 Posterior Analysis and Extrapolated Electron Density Maps

Each  $wDED_t$  map in the time-series may be comprised of one (or more) reaction intermediates. If the structures of these intermediates are known, the  $wDED_t$  may be fitted by the calculated time-dependent DED maps from these models. This technique is described elsewhere (Schmidt et al. 2004b).

Time-independent *extrapolated structure factors*,  $F^{ext}$ , are given by

$$F^{ext} = F_{calc}^D + N\Delta F \quad (2.34)$$

where  $F_{calc}^D$  are the calculated *structure factors* of the dark state model,  $N$  is a multiplication factor and  $\Delta F$  is the time-independent *difference structure factors* determined by the model(s) of the reaction intermediates. The goal is to determine a maximum value of  $N$  before large negative features begin to appear in  $F^{ext}$ . Starting with  $N=1$ ,  $F^{ext}$  is calculated and extrapolated EDMs are produced. The *difference electron density* features are integrated and this value is stored.  $N$  is increased and the process repeated until large negative features appear during integration of the *difference electron density* maps. Using the largest acceptable value of  $N$  the extrapolated EDMs are obtained.

### 2.2.7 TR-SFX: Extent of Reaction Initiation

To estimate the extent of reaction initiation, the fractional occupancies of the intermediates which are present for a particular time delay must be resolved. For the purpose of this dissertation, the structures of these intermediates and the dark structure are assumed known (i.e., PDB coordinate files are available). The extent of reaction initiation is estimated by fitting the calculated DED maps from the structural models for the intermediates present at that time delay to the observed DED map. The fitting is accomplished by a minimization process,

$$\min(\Delta\rho^{\text{obs}} - [sf_1\Delta\rho_{I1}^{\text{calc}} + sf_2\Delta\rho_{I2}^{\text{calc}} + \dots])^2 \quad (2.35)$$

where  $\Delta\rho^{\text{obs}}$  is the observed DED map,  $\Delta\rho_{I_n}^{\text{calc}}$  is the calculated DED map for the n-th intermediate  $I_n$  and  $sf_n$  is the n-th scale factor. As discussed in §2.2.4, the DED maps are all on an absolute scale. Therefore, the extent of reaction is the sum of the scale factors  $sf_1 + sf_2 + \dots + sf_n$ . The extent of reaction initiation for the TR-SFX on PYP was performed using a modified version of the program *Getmech* (Schmidt et al. 2004b).

### 2.2.8 TR-SFX: Assessment of Data Quality

The Monte Carlo integration performed by *Process\_hkl* merges many different reflection distributions together and the merged values may not share a high degree of similarity (White et al. 2012). Therefore, a new measure of data quality should be defined. Due to the serial nature of a TR-SFX experiment an additional assessment of data quality is used known as  $R_{\text{split}}$ . The calculation of  $R_{\text{split}}$  is achieved by splitting the reflection list into two interleaved sets,

$$R_{\text{split}} = 2^{-\frac{1}{2}} \frac{\sum |I_{\text{even}} - I_{\text{odd}}|}{\frac{1}{2} \sum (I_{\text{even}} + I_{\text{odd}})} \quad (2.36)$$

where  $I_{\text{even}}$  are the merged reflection intensities from even numbered images and  $I_{\text{odd}}$  are the merged reflection intensities from odd numbered snapshots. This splitting can be accomplished using a shell script and *Process\_hkl*. The resulting merged lists are then compared with the program *Compare\_hkl*. Structure refinement and analysis is identical to that described in §2.1.11.

## 2.3 SBL1, SBL2, SBL3, 3<sup>rd</sup> Generation Synchrotrons and XFELs

### 2.3.1 Schmidt Beamline 1 (SBL1)

SBL1 is a micro-focus X-ray diffraction beamline. X-rays are produced by a copper anode in an evacuated tube and the Cu  $K\alpha$  radiation has a peak energy of 8.04 keV for  $K\alpha_1$  ( $\lambda=1.5418 \text{ \AA}$ ) and a spot size of  $\sim 275 \times 275 \mu\text{m}^2$ , as measured at the sample position. The source can be calibrated by using a PIN-diode. With proper beam alignment an X-ray flux of  $10^8$ - $10^9$  counts per second (CPS) is produced, as calculated by the formula

$$\text{CPS} = \frac{\kappa \times V - V_b \times 10^3}{g} \quad (2.37)$$

where  $\kappa$  is a calibration factor,  $V$  is the voltage reading as measured by the PIN-diode,  $V_b$  is the measured voltage with no X-rays impacting the PIN-diode and  $g$  is a gain factor (Aotore et al. 2007). A Rayonix CCD165 X-ray detector records the diffraction patterns.

SBL1 is unique, meaning that it was not purchased as a turn-key solution from a corporation. Instead, the individual components were acquired separately from several companies and assembled in our lab, according to the designs created by Professor Schmidt and myself. A list of the primary components of SBL1 is outlined in Tab. 2.2 and the machine is shown in Fig. 2.9a along with a schematic in Fig. 2.9c. The maximum resolution for SBL1 is  $d=1.52 \text{ \AA}$ , as calculated by Eqn. 2.17 with the X-ray wavelength of  $\lambda=1.5418 \text{ \AA}$ , the detector radius of  $r=82.5 \text{ mm}$  and the minimum sample-to-detector distance of  $L=45.0 \text{ mm}$ .

**Table 2.2.** Primary components of SBL1.

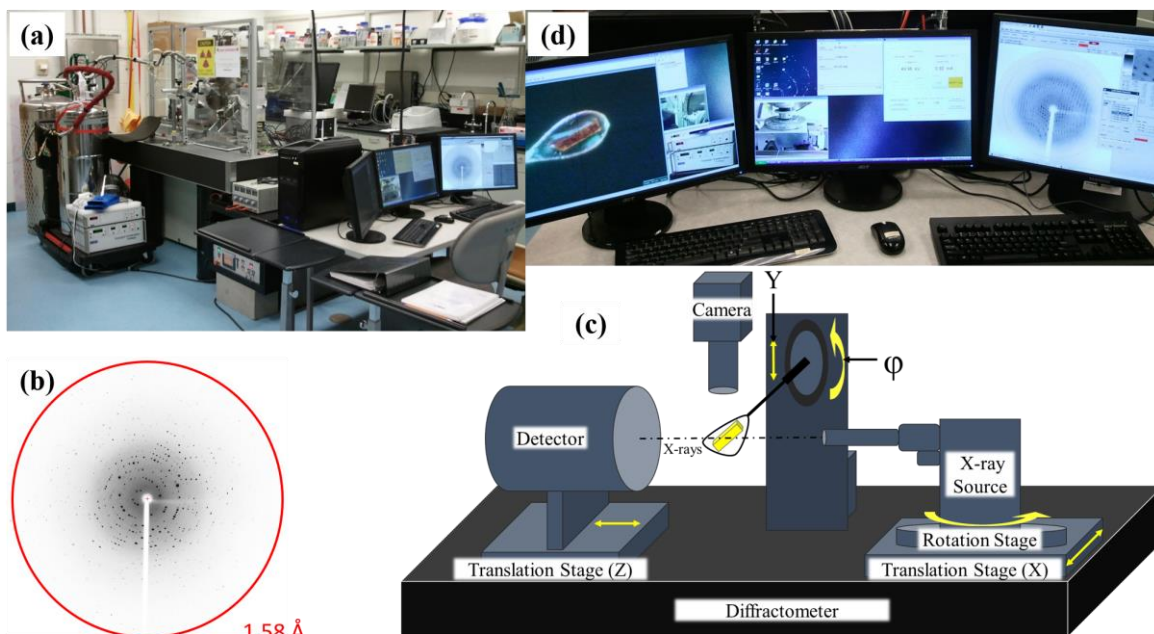
Component	Model #	Manufacturer
GeniX X-ray Source	MP-GeniX-G4	Xenocs; Sassenage, France
Diffractionmeter	Custom	HUBER Diffraction stechnik GmbH & Co. KG; Rimsting, Germany
X-ray Detector	CCD165	Rayonix; Evanston, IL, USA
Optical Table	PTQ12108	Thorlabs Inc.; Newton, NJ, USA
Shutter Controller	SC10	Thorlabs Inc.; Newton, NJ, USA
Optical Shutter	SH05	Thorlabs Inc.; Newton, NJ, USA
Light Source	MI-150	Edmund Optics; Barrington, NJ, USA
Camera	Stingray F-145C	Allied Vision Technologies; Stadroda, Germany

Diffraction images are recorded using the *rotation method* (see §2.1.9). X-ray exposure times are sample dependent, but an exposure of 300 seconds is normally sufficient to

obtain a high quality near atomic resolution diffraction pattern as shown in Fig. 2.9b. The X-ray detector manufacturer (Rayonix) provides a program called *marccd* (Fig. 2.11a) that not only controls the CCD165 detector but can also be configured to drive external hardware such as stepper motors and a shutter (Doyle et al. 2006). SBL1 employs the *marccd* software for data collection since it is simple to use and is capable of controlling the other devices. Configuring *marccd* to use the Huber smc9300 motor controller (Pp-electronic 2007; Pp-electronic 2011) and Thorlabs SC10 shutter controller (Thorlabs 2014) requires two components: a configuration file that defines the specific capabilities of the smc9300 and SC10 and a program script that interprets the positioning commands sent by *marccd*, translates them and then sends these commands to the appropriate controllers for execution. The data collection process is completely automatic once the dataset protocol has been defined using the *marccd* → *Acquire Dataset* software dialog, seen in Fig. 2.11b.

Since each component was purchased individually, facilitating seamless communication between the various machines was challenging. There are 3 computers used at SBL1: (i) Computer 1 is a Windows 7 PC located at the optical table and used for beam set up / alignment / optimization and sample positioning; (ii) Computer 2 is a Windows 7 PC located at the control desk and operates

the Oxford Instruments CryoJet HT liquid nitrogen cryogenic system (CryoJet), the GeniX X-ray source, the Allied Vision Technologies (AVT) X-ray reference and sample monitoring camera, 3 web cameras to monitor critical systems and a remote desktop connection to the smc9300 for monitoring motor positioning; (iii) Computer 3 is a Redhat Enterprise Linux (RHEL) PC located at the control desk that operates the detector, sets up data collection, controls the X-ray shutter and motors during data collection and performs data reduction. The smc9300 drives 3 separate stepper motors (Fig. 2.9c): a detector translation stage that positions the detector distance from 45 mm to 600 mm away from the sample, a rotation stage that rotates the phi axis of the sample and a translation stage that positions the rotation axis into the X-ray beam. All 3 computers are networked using an Ethernet switch, and the motor and shutter controllers are connected to the RHEL PC via 9-pin RS-232 serial cables. The control desk and 3 monitors are shown in Fig. 2.9d. The 2 monitors on the left are used for Computer 2 and the one on the right is for Computer 3.

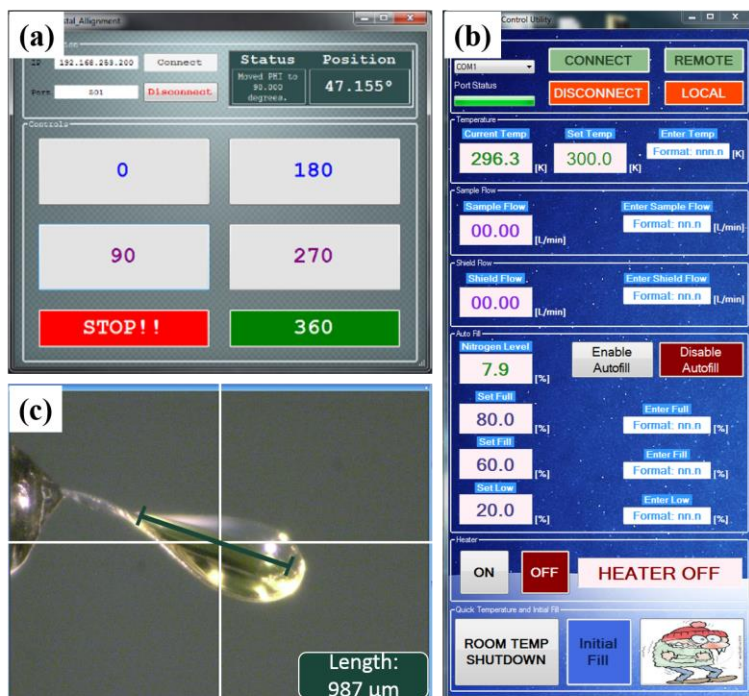


**Figure 2.9.** SBL1. (a) SBL1 setup in the lab. (b) Diffraction data from PYP obtained using SBL1. The red circle identifies 1.58 Å resolution. (c) A schematic of SBL1 is shown here; the sample is yellow and placed into the X-ray beam; the direction of movement for the translation and rotation stages are denoted by yellow arrows. (d) SBL1 control desk showing an experiment in progress.

As mentioned above, Computer 1 is located at the optical table and is used for beam set up / alignment / optimization and sample positioning. A *Visual C-Sharp (C#) Graphical User Interface* (GUI) was developed in-house to facilitate quick phi axis positioning during sample mounting and alignment into the X-ray beam using this computer, as shown in Fig. 2.10a. Once a sample has been mounted onto the goniometer the user aligns the crystal into the crosshairs of the reference camera at  $\phi=0^\circ$  by adjusting one goniometer axis. The sample is then rotated conveniently using the GUI to  $\phi=90^\circ$  and the second axis of the goniometer is adjusted to bring the crystal back into the crosshairs. At this time the sample should be aligned properly into the X-ray beam at the rotation axis, meaning that as the crystal rotates it will not leave the X-ray beam. The user double checks the alignment by rotating the sample to  $\phi=180^\circ$  and  $\phi=270^\circ$  while verifying the crystal is still positioned in the AVT camera crosshair.

Computer 2 is used primarily for monitoring, but it also controls all key components of SBL1. Two GUIs were developed for operating the CryoJet and for measuring protein crystal dimensions once mounted, and these are shown in Fig. 2.10b and Fig. 2.10c, respectively. The CryoJet did not come with any software from the manufacturer and all adjustments needed to be done locally using the front panel of the controller. This was inconvenient since experiments frequently extend into the weekend and if small adjustments are needed then someone must physically come to the lab. The CryoJet controller can interpret ASCII commands sent via a serial connection, so a GUI was created to provide near full control of the machine. The company also does not provide any means to adjust the liquid nitrogen fill parameters from a remote connection, so a solution had to be designed to deliver full remote control. Using an AC relay fitted to an Arduino UNO microcontroller board, we programmed the Arduino to turn the relay on and off when the appropriate command is received from the CryoJet GUI software. The liquid nitrogen level is





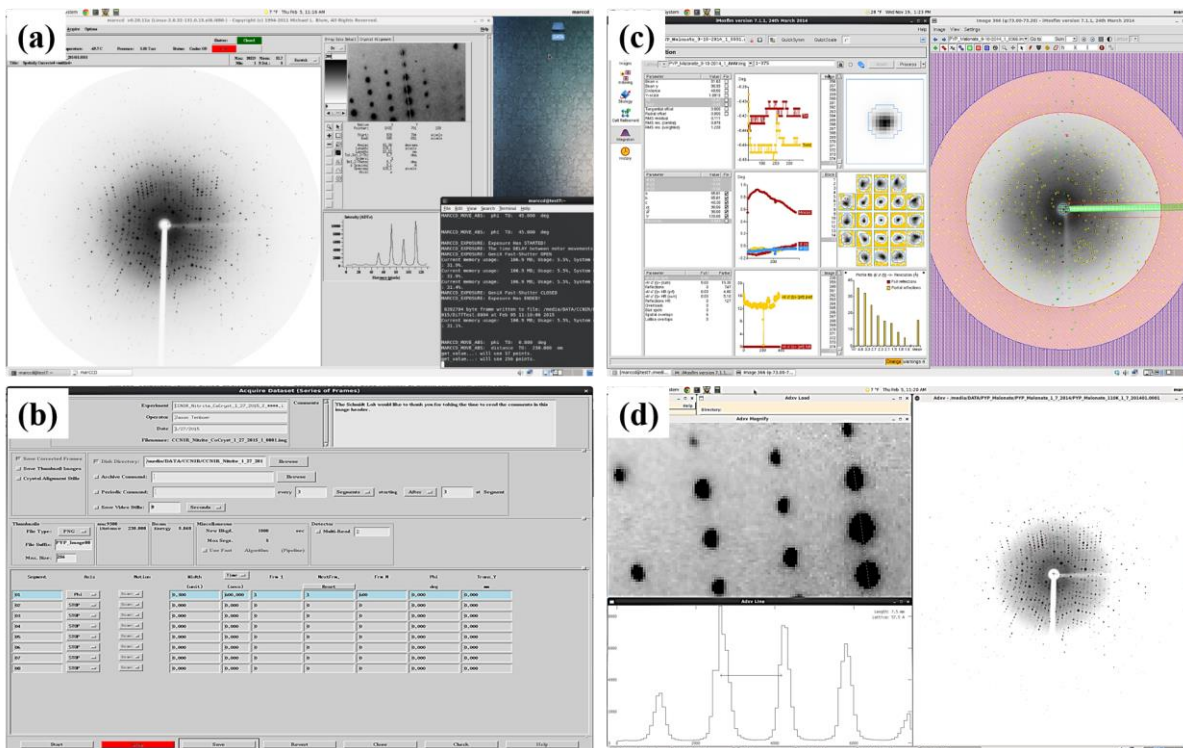
**Figure 2.10.** Graphical User Interfaces (GUIs) written in-house with Visual C#. (a) GUI to control the phi axis of the smc9300 motor controller during sample alignment. (b) GUI to set all features of the CryoJet for complete remote control. (c) A GUI to facilitate easy crystal measurements. The program is transparently overlaid on the Allied Vision Technologies (AVT) camera display. The white crosshair denotes the X-ray interaction region and a length scale is shown.

continually monitored by the software and when it falls below the fill level set by the user it activates the relay via the Arduino board which then sends 115 VAC to a solenoid valve that opens and allows liquid nitrogen to fill the instrument dewar. Once the nitrogen level climbs above the desired full level the software

turns off the relay by the same process to stop the flow of liquid nitrogen. A function was also added to the software so that if the liquid nitrogen level falls below a

user set low level (for instance, someone forgot to fill the main nitrogen storage tank) then the machine warms to room temperature, holds this temperature for 15 minutes and then shuts down the machine to protect critical components from overheating due to the lack of nitrogen. The second program provides an easy way to measure crystal dimensions after sample mounting, as seen in Fig. 2.10c.

For controlling data collection, a *tool command language / expect* (TCL/EXP) software script was created to translate commands from the *marccd* software and serve them to both the smc9300 and SC10 shutter controller. This program is installed on the RHEL PC. The TCL/EXP script



**Figure 2.11.** Data collection with *marccd* and initial data reduction using *iMosflm*. (a) Rayonix *marccd* control software. (b) *marccd* → *Acquire Dataset* dialog. (c) *iMosflm* program from the CCP4 software suite for spot finding, indexing, cell refinement and integration of reflection intensities. (d) *Adxv* program for displaying X-ray diffraction data.

receives the following commands from the *marccd* software: *marccd\_move\_abs* will move the motor(s) to the required position(s); *marccd\_set\_abs* will set the motor encoder(s) to a specific location; *marccd\_exposure* will execute a data collection procedure according to the protocol defined by the user in the *marccd* → *Acquire Dataset* software dialog; *marccd\_shutter* will open / close the X-ray shutter. The general steps for collecting a dataset with SBL1 are described here. Start the *marccd* software and set up the dataset by filling in the *Acquire* → *Dataset* dialog displayed in Fig. 2.11b, including the disk directory and root filename of the experiment, operator name, date, the width of phi rotation per exposure, time of exposure and the number of images in the dataset. Begin the data collection. Once image acquisition has started the entire process is automatic and the user can monitor progress by checking images using the *adxv* software (Fig.

2.11d). Finally, the data (Fig. 2.11c) are processed and the structure is refined as described in §2.1.10 and §2.1.11.

In Fig. 2.9a a safety guard is shown that is constructed from 3/8" thick clear polycarbonate panels mounted within grooved Aluminium channel uprights attached to the optical table. Two doors are installed for ease of access to the sample environment. These doors are interlocked with magnetic switches to the safety shutter within the GeniX X-ray source and will close the shutter when open. Although the polycarbonate panels are not capable of stopping the direct beam, they are thick enough to trap all scattered radiation inside the enclosure. This has been verified by testing with a Geiger counter all around the machine. The direct beam is completely blocked from the back by the detector.

### 2.3.2 Schmidt Beamline 2 (SBL2)

SBL2 is built to perform *time-resolved absorption spectroscopy* (TR-AS). The components of SBL2 are listed in Tab. 2.3 and shown in Fig. 2.12c. They consist primarily of a ns laser (pump), a Xenon light source (probe), a spectrometer and a *charge coupled device* (CCD) iStar detector. Although spectra can be recorded in as little 2 ns with the iStar (Andor 2008a), the pulse width of the laser is 5 ns *full-width half-max* (FWHM) so this effectively sets ~5 ns as the minimum achievable time resolution for SBL2 (Opotek 2008a; Opotek 2008b). Similarly to SBL1, SBL2 is not a turn-key solution; instead, it is a unique time-resolved near ultraviolet / visual wavelength microspectrophotometer (TR-NUV/VIS-MSP) incorporating components from various companies and assembled according to the designs drafted by Professor Schmidt and myself (Purwar et al. 2013).

The design of the TR-NUV / VIS-MSP is shown schematically in Fig. 2.12a. The MAX-302 is a 300 W Xenon light source (VIS) that generates a spectrum from 385-800 nm, as shown in Fig.

**Table 2.3. Primary components of SBL2.**

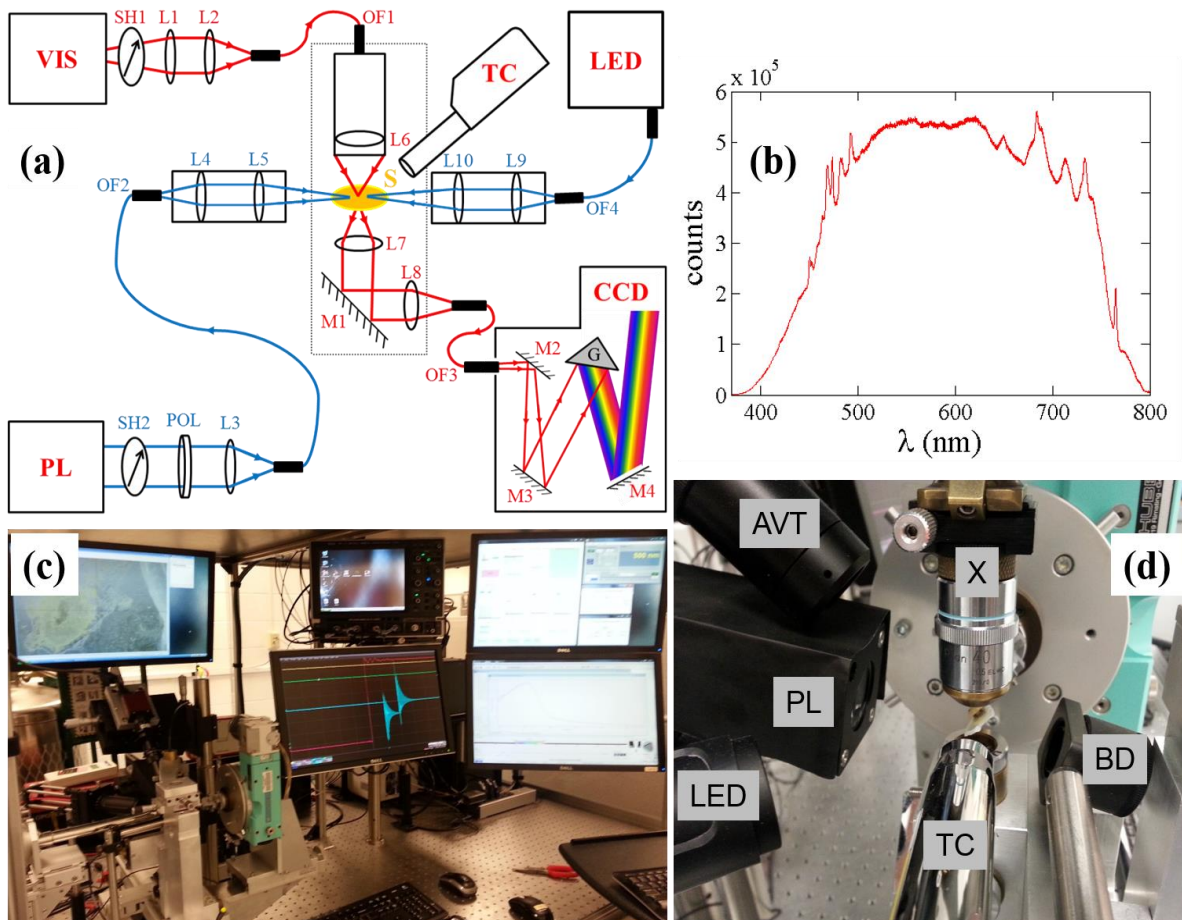
Component	Model #	Manufacturer
Xenon Light Source	MAX-302	Asahi Spectra Co., Ltd.; Tokyo, Japan
iStar ICCD Detector	DH734- 18U-A3	Andor Technology Ltd.; Belfast, UK
Shamrock Spectrometer	SR-303i-A	Andor Technology Ltd.; Belfast, UK
Microspectrophotometer	custom	Advanced Photon Source, BioCARS; Lemont, IL, USA
VIS ns Laser	Opolette 355-HE-II	Opotek, Inc.; Carlsbad, CA, USA
Shutter Controller	SC10	Thorlabs Inc.; Newton, NJ, USA
Optical Shutter	SH05	Thorlabs Inc.; Newton, NJ, USA
Camera	Stingray F- 145C	Allied Vision Technologies; Stadtroda, Germany
Optical Table	ST-series	Newport Corporation; Irvine, CA, USA
Digital Delay Generator	DG645	Stanford Research Systems; Sunnyvale, CA, USA
Calibration Light Source	CAL-2000	Ocean Optics; Dunedin, FL, USA

2.12b (Asahi Spectra 2010). The MAX-302 features a design that effectively eliminates wavelengths >750 nm thereby reducing sample damaging infrared radiation. Light intensity is controlled by an internal neutral density filter adjustable from ~18  $\mu$ W-650  $\mu$ W,

as measured at the sample. Output from the MAX-302 is coupled into a hybrid light guide with a numerical aperture of 0.57. This light guide consists of a  $\varnothing$ 5 mm bundle of  $\varnothing$ 50  $\mu$ m core fibers. The light is coupled into the fiber by  $\varnothing$ 50 mm OD focusing optics located within the instrument. The output of this fiber is controlled by a SC10 shutter controller and a SH05 optical shutter (SH1). The external focusing optics (L1 and L2) couple the light into an  $\varnothing$ 800 /  $\varnothing$ 200  $\mu$ m tapered optical fiber (OF1), which is then mounted to the microscope objective above the sample. The sample environment (Fig. 2.12d), including the two microscope objectives and the sample mount, is described elsewhere (Chen et al. 1994). In short, the microscope objective (L6) focuses the probing light to a  $\varnothing$ 220  $\mu$ m spot at the sample (S). The transmitted light is collimated and focused by two



identical  $\varnothing 10$  mm OD 4 $\times$  achromatic microscope objective lenses (L7, L8). An enhanced Aluminium coated mirror guides the collimated beam to the focusing lens (L8), which couples the light into a pickup fiber (OF3) that carries the signal to the spectrometer. The Andor Shamrock SR-303i spectrometer supports a wavelength range from 190 nm-10  $\mu$ m with a wavelength resolution of up to 0.1 nm and wavelength accuracy of  $\pm 0.2$  nm (Andor 2008b). Attached to the spectrometer is an iStar *intensified charge coupled device* (ICCD) camera with a built-in (internal) *digital delay generator* (DDG<sub>i</sub>).



**Figure 2.12.** SBL2. (a) Components of SBL2. VIS: MAX-302 light source; PL: pulsed laser; TC: temperature control; LED: high power light emitting diode; CCD: iStar camera; S: sample; OF<sub>x</sub>: optical fibers; L<sub>x</sub>: lenses; M<sub>x</sub>: mirrors; SH<sub>x</sub>: shutters; POL: polarizer; G: grating. (b) Spectrum of the Xenon light source. (c) SBL2 assembled in the lab. (d) Close-up of the sample environment (clockwise starting from lower left corner): LED optics (LED); laser optics (PL); microscope camera (AVT); Xenon focusing optics (X); laser beam dump (BD); CryoJet temperature control (TC). Fig. 2.12a from Purwar et al. 2013.

This DDG<sub>i</sub> enables control of multiple parameters: the *delay* setting is a time-delay before the intensifier is activated and a spectrum is collected; the *gate pulse width* is the time the intensifier is active, which defines the experimental time resolution; the *exposure time* is the total time the CCD is in a powered-on state. The *micro-channel plate* (MCP) gain setting controls the amplification of the signal reaching the CCD by altering the voltage across the plate. These parameters may be set either by using the *Andor Solis* proprietary software or by a script I have written in the *Andor Basic* language and executed within the *Andor Solis* program. Although the *Andor Basic* script significantly automates data collection, the downside is that while the program is running no other software control is permitted with the equipment.

This shortcoming is currently the subject of software design in our lab where we are building a complex GUI that can control all of the hardware at SBL2 without this limitation, as shown in Fig. 2.13a-2.13f. We have named this program the *GUI of Everything GUI* (GOEGUI), since this one piece of software will control all hardware and the data collection process. This solution is being designed with the Microsoft Visual Studio Professional 2010 integrated development environment using the Visual C# programming language. GOEGUI will also be capable of performing data reduction such as calculating difference spectra in addition to scaling of the spectra and performing an SVD analysis. This work is currently under development and will not be discussed here.

The laser pump pulse to start a reaction is provided by a Q-switched *neodymium-doped yttrium aluminium garnet* (Nd:YAG) pulsed laser that is tunable by an *optical parametric oscillator* (OPO) from 410 nm to 710 nm. The ns laser can be controlled either internally or externally, but for the greatest timing flexibility we choose to operate the flashlamp internally at 20 Hz while externally triggering the Q-switch via a Stanford Research Systems DG645 digital delay generator (Systems

2008). Maximum pulse intensity is achieved when the Q-switch signal is delayed by  $\sim 146 \mu\text{s}$  after the flashlamp, which provides a 5 ns FWHM pulse at 6.5 mJ (wavelength dependent) approximately 40 ns after the rising edge of the Q-switch.

Laser pulses are delivered to the sample by means of two optical assemblies. The first consists of a SC10 shutter controller and SH05 optical shutter (SH2), a calcite laser polarizer (POL), a focusing lens (L3), and an  $\varnothing 600 \mu\text{m}$  core diameter fiber (OF2). Attenuation of the laser is



**Figure 2.13.** SBL2 GUI of Everything GUI (GOEGUI). (a) The connections tab where all equipment is connected to the software. (b) Experiment setup tab which configures the iStar camera and Shamrock spectrometer and defines the experimental parameters for data collection. (c) Interface for adjusting all features of the MAX-302 Xenon light source. (d) Tab providing control of the CryoJet cryogenic system. (e) Shutter tab for controlling all parameters of the Xenon and laser shutters. (f) Tab for adjusting all features of the DG645 digital delay generator.

accomplished by the polarizer, which is attached to a Thorlabs PRM1 rotation stage enabling adjustment of the laser energy from 38  $\mu\text{J}$  up to a maximum of 2 mJ, measured at the sample with  $\lambda = 500 \text{ nm}$ . The pulses are focused to a spot size less than  $\varnothing 700 \mu\text{m}$  and are then coupled into a fiber (OF2). The fiber delivers the light to the second optical assembly, which consists of a 3-axis translation stage containing collimating and focusing optics (L4 and L5), which is used to focus and align the beam onto the sample. Typical focal spots are in the range of 1-1.5 mm.

Additionally, we use high-intensity (11 mW, 31 mW) fiber-coupled *light emitting diodes* (LED) operating at 490 nm and 590 nm, respectively, to switch between spectral forms of a protein, such as  *$\alpha$ -phycoerythrocyanin* ( $\alpha$ -PEC). Illumination with the LED light is used to prepare a distinct state, which is then excited by the laser pulse at a different wavelength to initiate a transition to another state. The LED light is focused by lenses L9 and L10. Focal spots of the LEDs are between 2-3 mm and they may be controlled by a 5V *Transistor-transistor Logic* (TTL) signal. Fig. 1.4b demonstrates the capabilities of SBL2 for recording high quality spectra of  $\alpha$ -PEC using this LED illumination method. PYP dark spectra (no laser excitation) is shown in Fig. 1.2b. The temperature at the sample can be adjusted from 90 K to 490 K using an Oxford Instruments CryoJet HT liquid nitrogen cryogenic stream. Software control of the CryoJet is provided by *GOEGUI*, as shown in Fig. 2.13d.

Experimental timing is perhaps the biggest challenge for SBL2. The laser, laser shutter, Xenon shutter, LEDs and iStar detector must all be synchronized and controlled in real time. Each of these devices can be synchronized by sending 5V TTL pulses on a *Bayonet Neill-Concelman* (BNC) connector. If the laser is operated completely externally, then two signals are needed to control the flashlamp and Q-switch. The laser and xenon shutter, LEDs and iStar camera each require one pulse. The DG645 is equipped with only 4 programmable pulse outputs so a creative solution is



required or another DG645. Tab. 2.4 outlines the timing signal connections for synchronizing the hardware using just one DG645.

**Table 2.4.** SBL2 timing signal connections. *EXT*=DG645 external trigger input; *FL*=laser flashlamp; *QS*=laser Q-switch; *XS*=Xenon shutter; *iS*=iStar camera external trigger input; *LED*=LED controller input.

DG645 Channel	From	To	Details
EXT	FL	EXT	Signal from laser flashlamp output to DG645 EXT
A/B	A/B	QS	Signal from DG645 channel A/B to laser Q-switch input
C/D	C/D	XS	Signal from DG645 channel C/D to Xenon shutter
E/F	E/F	iS	Signal from DG645 channel E/F to iStar camera
G/H	G/H	LED	Signal from DG645 channel G/H to LED controller

The timing scheme begins with the laser flashlamp running internally at 20Hz. The laser controller generates these signals and the pulses are fed into the external trigger input (EXT) of the DG645. The laser shutter controller (LS) is

controlled by the *Andor Basic* software script. Once the DG645 receives the pulse from the laser controller, channel A/B is delayed by ~50.146 ms and sent to the Q-switch (QS) to activate the laser. The laser pulse is generated 40 ns after the QS signal is received and the LS will be open at this time. The arrival of the laser pulse at the sample defines the experimental time  $T_0$ . The signal from channel E/F sets the data time-delay, as this is sent to the iStar camera. After the iStar receives this pulse there is an additional ~36 ns delay until the internal circuitry of the iStar is ready for data acquisition. This is called the *insertion delay*. Channel C/D activates the Xenon shutter and channel C is always set to the value of channel E – 17 ms, since 17 ms is the average time needed for the SH05 optical shutter to open. Channels C/D and E/F therefore always move together as a pair. Channel G/H activates immediately after channel F (or after a delay, if desired) and can be used either to set a protein relaxation time before the next data point or to activate an LED to prepare an initial state in a non-cyclic photo-reactive protein, such as  $\alpha$ -PEC.

The *Andor Basic* script or the *GOEGUI* may be used to define and execute all data collection parameters, and these programs will enter the appropriate settings into the DG645 for each time-delay during the experiment. The timing is monitored and fine-tuned using a LeCroy WaveSurfer (Model 40S, 400 MHz) oscilloscope (LeCroy 2009). Tab. 2.5 outlines the required pulse widths for each channel of the DG645.

The Xenon probing light is positioned approximately normal to the surface of the sample focal plane (SFP). The laser is tilted by  $\sim 30^\circ$  above the SFP, therefore, the angle between the laser and

**Table 2.5.** DG645 channel delays for SBL2. *EXT*=external trigger input; *w.r.t.*=with respect to.

DG645 Channel	Delay w.r.t.	Delay [ms]	Details
EXT	laser flashlamp	none	EXT receives a 20 Hz signal from the laser flashlamp which begins DG645 delay program
T0	EXT	none	Channel T0 is active once EXT receives a pulse from the laser
A	T0	50.146015	Laser Q-switch rising edge
B	A	0.10	Laser Q-switch falling edge
C	E	-17.0	Xenon shutter rising edge
D	C	10.0	Xenon shutter falling edge
E	A	variable	iStar rising edge which defines the time-delay
F	E	10.0	iStar falling edge
G	F	0.0010	LED/relaxation time rising edge
H	G	variable	LED/relaxation time falling edge

Xenon beams are  $\sim 60^\circ$ .

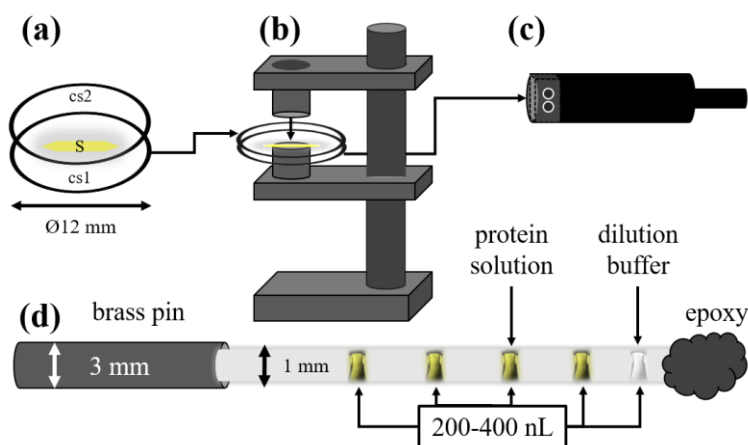
When samples in solution are studied (Fig. 2.14d) the user positions a capillary into the intersection of the SFP and the plane spanned by the Xenon and laser beams (XLP). When crushed

single crystals are used (Fig. 2.14a) the method is similar, however, once the sample has been positioned into intersection of the SFP and XLP the goniometer is then rotated by an angle of  $30^\circ$ . This rotation creates an angle of  $60^\circ$  between both the laser and the Xenon light spots with respect to the SFP.

As with SBL1, SBL2 is contained within a safety guard constructed from  $\frac{1}{4}$ " polycarbonate panels. These panels are mounted in frames with rollers so that they may slide sideways to permit

access to the instrument. There are 8 doors in total and each of them are interlocked with magnetic switches. If any of the doors are opened the laser will not function until the door is closed. These doors may be blacked out during an experiment so that no laser light penetrates to the main room. The design of the interlock enables anyone to be in the room safely when the Class IV laser is in operation.

Absorption spectroscopic methods traditionally require sizable quantities of purified protein in solution. Also, large single crystals are usually exceedingly optically dense and do not allow complete penetration of the probing light. Therefore, obtaining high quality absorption spectra is difficult with large crystals. To avoid these limitations smaller volumes of protein in solution must be acceptable and single crystals should be crushed until the path length permits full penetration of the probing light. The SBL2 instrument permits the study of both nL-sized quantities of protein



in solution and large single crystals.

SBL2 accommodates two sample mounting methods, including protein in solution and crushed single crystals as shown in Fig. 2.14a and Fig. 2.14d and described in previous work (Purwar 2013). Both of these methods mount the sample to the goniometer using the hole intended for the brass

**Figure 2.14.** SBL2 sample mounting and preparation methods. (a) A single crystal, *S*, is placed in a 1.5  $\mu\text{L}$  drop of buffer between two cover slides, *cs1* and *cs2*. (b) The crystal is crushed between the cover slides by tightening a screw on the device picture here. The edges of the cover slides are then sealed with a fast curing epoxy. (c) After the epoxy on the cover slides has cured the sample is placed into a holder and secured using two set screws. (d) A protein solution is prepared using a glass capillary mounted inside a brass pin. Protein droplets of 200-400 nL each are transferred into the capillary. A drop of dilution buffer is placed in the capillary and both ends are sealed with epoxy to prevent drying of the sample.

specimen pin. To prepare a crushed single crystal, a 1.5  $\mu\text{L}$  drop of an appropriate buffer (e.g. Na-malonate for PYP) is first placed at the center of a 12 mm diameter cover slide (Hampton Research, HR3-277). A single crystal is then positioned within this drop and a second cover slide is laid directly on top of the first. Using the device pictured in Fig. 2.14b the cover slides are pressed together by turning a screw and then held in place while they are sealed with the two-part epoxy. Once this epoxy has cured the sample is placed into the holder shown in Fig. 2.14c and secured with set screws before being mounted onto the goniometer for an experiment. The second type of sample uses 1 mm OD borate silica glass capillaries epoxied inside a brass specimen pin. The protein concentration is adjusted using the dilution buffer to  $\sim 1\text{-}12$  mg/mL, depending on the protein, so that the absorbance maximum is  $\sim 1$  *absorbance units* (a.u.). An a.u. value of 1 means that 90% of the probing light has been absorbed. The a.u. is a measure of the *optical density* of the sample and is determined by the equation (Andor 2008a)

$$\text{a. u.} = \log_{10} \frac{S_{\text{ref}}}{(S_{\text{sig}} - S_{\text{bg}})} \quad (2.38)$$

where  $S_{\text{ref}}$  is the reference spectrum of the probing light source through the dilution buffer,  $S_{\text{sig}}$  is the signal spectrum of the probing light through the sample and  $S_{\text{bg}}$  is the spectrum of the ambient room lighting with no probing light exposed to the detector. A pipette tip is connected with a short piece of rubber tubing to the brass specimen pin. Using a pipette,  $\sim 200\text{-}400$  nL drops of the diluted protein solution are transferred into the capillary until filled, as pictured in Fig. 2.14d. Lastly, a drop of the dilution buffer is also placed within the capillary before the ends are sealed with epoxy.

### 2.3.3 Schmidt Beamline 3 (SBL3) and Gas Dynamic Virtual Nozzles (GDVNs)

SBL3 is a platform to fabricate and test GDVNs. Tab. 2.6 lists the primary components of SBL3. The basic construction of a GDVN is described here. A Polymicro polymide-coated synthetic silica capillary is cut to a length of 4 m. This capillary has an OD of  $\text{Ø}365 \mu\text{m}$  and an ID selectable from  $\text{Ø}50\text{-}100 \mu\text{m}$ , depending on the sample. A conical profile is formed on one end of

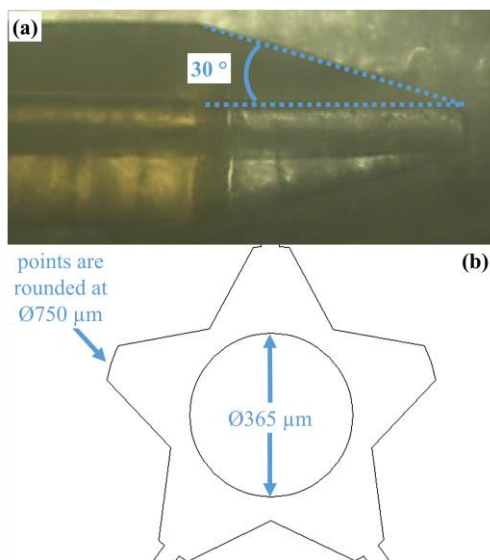
**Table 2.6.** Primary components of SBL3.

Component	Model #	Manufacturer
Optical Table	B3648F	Thorlabs Inc.; Newton, NJ, USA
Camera	Stingray F-145C	Allied Vision Technologies; Stadroda, Germany
Trinocular Microscope	52-352	Edmund Optics Inc.; Barrington, NJ, USA
Vacuum Pump	XDS-10i	Edwards; Crawley, West Sussex, UK
Fiber Shaper/Polisher	ULTRAPOL Fiber Lenser	UltraTec Manufacturing Inc.; Santa Ana, CA, USA

this capillary using an optical fiber polisher to an angle of  $30^\circ$  measured from the axis of the capillary (i.e.,  $60^\circ$  from vertical), as shown in Fig.

2.15a. A borate silica glass

capillary of  $\text{Ø}765 \mu\text{m}$  ID and  $\text{Ø}1 \text{ mm}$  OD is cut into a 10 cm length and secured into a holder mounted on a small electric motor. While spinning this motor at 30-50 rotations per minute, a butane torch is used to slowly melt one end of this capillary from an  $\text{Ø}765 \mu\text{m}$  ID to an ID of  $\text{Ø}100\text{-}175 \mu\text{m}$ . This profile can be seen in Fig. 1.1 and is used to constrict the flow of the helium gas so that it acts as a dynamic virtual nozzle to compress and stabilize the inner microcrystal stream. The borate silica glass capillary is epoxied into a stainless steel tube of dimensions  $\text{Ø}1 \text{ mm}$  ID x  $\text{Ø}1.7 \text{ mm}$  OD x 6 cm length. ~4 cm of the glass capillary is allowed to protrude out of one end of the stainless steel tube. Two Kapton spacers (MiTeGen) (see Fig. 2.15b) are placed onto the polished end of the Polymicro capillary. One spacer is positioned ~5 mm in from the beginning



**Figure 2.15.** GDVN construction. (a) A Polymicro polyimide-coated synthetic silica capillary of  $\text{\O}365 \mu\text{m}$  OD is ground into a conical profile at  $30^\circ$  by an UltraTec ULTRAPOL fiber lenser. (b) The Polymicro capillary is then mounted inside a borate silica glass capillary with an  $\text{\O}750 \mu\text{m}$  ID using star-shaped Kapton spacers of  $\text{\O}750 \mu\text{m}$  OD and an ID of  $\text{\O}365 \mu\text{m}$ . These spacers center the Polymicro capillary within the outer glass capillary. The star-shaped profile allows the helium gas flow to surround the inner microcrystal stream, acting as a gas dynamic virtual nozzle (GDVN). Two Kapton spacers are used for each nozzle construction.

of the conical profile and the second is placed a further 2 cm in from the first spacer. These spacers will center the inner capillary within the outer capillary while allowing helium gas to flow through. A 5 cm length of  $\text{\O}370 \mu\text{m}$  ID dual-bore flexible tubing is used to thread the Polymicro capillary into one of these holes. An additional Polymicro capillary of dimensions  $\text{\O}365 \mu\text{m}$  OD x  $\text{\O}100 \mu\text{m}$  ID x 4 m length is fed into the other hole. This second capillary transports the helium focusing gas and the ends are both cut square. The stainless steel tube is now placed into an Upchurch fitting and this is then threaded into a standard vacuum fitting to complete the nozzle fabrication.

This method generally describes how nozzles were constructed for both of the PYP experiments at the LCLS. A chromatography solvent delivery pump or compressed nitrogen gas supply is used to pressurize

the hydraulic sample reservoir. Trial-and-error is required to determine the optimal parameters for helium focusing gas pressure, sample flow rate / sample gas pressure and ID of the Polymicro capillary.

### 2.3.4 APS, BioCARS 14-ID-B and 14-BM-C Beamlines

X-rays are generated at the APS by first heating a cathode to  $1100^\circ\text{C}$  to generate electrons.

The electrons are then accelerated through a high voltage linear accelerator (linac) to 450 MeV at

$>0.99999c$ . Next, the electrons are fed into a booster synchrotron and further accelerated to  $\sim 7.6$  GeV. The accelerating force is provided by high electric fields in four radio frequency cavities, in synchronized coordination with bending magnets. These high energy electrons are then transferred to a 1,104 m circumference storage ring that is designed for use with insertion devices, which is the hallmark of a 3<sup>rd</sup> generation synchrotron X-ray source. Insertion devices are comprised of a linear periodic array of strong magnets arranged in an alternating fashion of north and south poles. These magnets deflect the path of the incoming electrons by the Lorentz force, and this change in direction produces the synchrotron X-ray radiation which is then fed into the experimental hutches. Further beam conditioning is accomplished by means of motorized slits and mirrors within the various beamlines, depending on the types of samples and desired experiments to be accommodated.

There are 71 different beamlines located at the APS, designed for experiments in multiple fields of research from Materials Science, Physics, Chemistry, Optics and Life Sciences. All of the experiments referenced in this work were performed at two beamlines: 14-BM-C and 14-ID-B. Both of these beamlines are operated by BioCARS, which is managed by the University of Chicago and led by Keith Moffat. 14-BM-C is a bending magnet source, producing monochromatic X-rays with a  $\Delta E/E = 3.1 \times 10^{-4}$  in the energy range from 8-14.9 keV. This instrument can produce a flux of  $6 \times 10^{11}$  photons / second at 12.668 keV into a spot size of  $130 \times 340 \mu\text{m}^2$ . Beamline 14-ID-B uses an *undulator* insertion device to generate a flux of  $5 \times 10^{10}$  photons / 100 ps pulse at 12 keV with a bandwidth of  $\Delta E/E = 5 \times 10^{-2}$  into a  $15 \times 20 \mu\text{m}^2$  spot size.

### 2.3.5 X-ray Free Electron Lasers (XFELs) and the Coherent X-ray Imaging (CXI)

#### Beamline

John Madey demonstrated in his seminal paper from 1970 (Madey 1971) that *free electron lasers* (FELs) could create finite gain spontaneous radiation from the far-infrared to visible wavelengths. Further, he established the possibility of using FELs to generate coherent X-ray radiation. Nearly 40 years later, the vision of an *X-ray free electron laser* (XFEL) in the hard X-ray regime ( $\geq 9$  keV) was finally realized when the LCLS came online at SLAC in 2009.

Production of X-rays begins with the generation of electrons in a copper photocathode by pulses from a titanium-sapphire (Ti:sapph) laser that has been frequency doubled into the ultraviolet wavelength range (Schmüser et al. 2014). These electron bunches are then accelerated by the last kilometer of the SLAC linac to a peak current / energy of 3.4 kA and 15 GeV, respectively (Emma et al. 2010). This electron beam is then inserted into a periodic arrangement of alternating magnetic dipoles, called an *undulator*. Since an optical cavity that operates in the X-ray regime has not yet been realized, amplification cannot be achieved by multiple passes through a lasing medium as with a conventional *bound-electron laser*. Instead, lasing of a FEL is accomplished through the process of *self-amplified spontaneous emission* (SASE) by a single pass through an arrangement of multiple *undulators* (Kim 1986a; Kim 1986b; Wu and Emma 2006).

The relativistic electron bunches produce spontaneously emitted radiation as they travel through these long *undulators*. The large electron current then interacts with the strong radiation field, leading to a self-organization called *micro-bunching*. In this process the electrons group together into bunches with each bunch separated by a distance  $\lambda_{\text{SASE}}$ , which is the radiation wavelength. This *micro-bunching* effect leads to brilliant, coherent X-ray radiation because this



electron bunch behaves in a manner similar to a point-particle electron, but now with a charge of  $N \times e$ , where  $N$  is the number of electrons in the bunch.

The *coherent X-ray imaging* (CXI) instrument (Boutet and Williams 2010) is located at SLAC in Hutch 5 of the Far Experiment Hall of LCLS. The LCLS experiments described in this dissertation were performed at this beamline. X-ray radiation at 4-10 keV is delivered into spot sizes selectable from 10  $\mu\text{m}$ , 1  $\mu\text{m}$  and 0.1  $\mu\text{m}$ , with  $\Delta E/E=0.1\%$  and  $\sim 10^{12}$  photons per pulse. Pulses are of  $\sim 40$  fs duration and arrive with a 120 Hz repetition rate ( $\sim 8.33$  ms between subsequent shots). Multiple sample delivery methods are available, including the GDVNs described in §1.4 and §2.3.3. Resulting 2-dimensional diffraction patterns are collected on a *Cornell-SLAC pixel array detector* (CSPAD) (Hart et al. 2012a; Hart et al. 2012b; Herrmann et al. 2013) that operates at the 120 Hz LCLS repetition rate.

## 2.4 Dynamics in Biological Macromolecules with Nanosecond Time Resolution

A TR-SFX experiment was performed at the LCLS on the model system PYP to assess the capability of the CXI instrument to collect time-resolved diffraction snapshots at near atomic spatial resolution and ns time-resolution. Prior to this experiment, studies on other photo-active proteins at the LCLS had produced low resolution, difficult to interpret time-resolved structures (Kupitz et al. 2014a). Although all intermediate structures of the PYP photocycle are known from  $\sim 100$  ps onwards by experiments conducted at synchrotrons, the *trans* to *cis* isomerization of the pCA chromophore occurs earlier and had not yet been observed by crystallographic methods.

When laser pulses are of ns duration, the time-width of these pulses is sufficient for initiating a photo-reaction in a substantial number of the PYP molecules so that subsequent X-ray pulses may measure the small structural changes. Since typically only  $\sim 20\%$  of molecules enter the photocycle, ensuring that all molecules within the X-ray interaction volume have an opportunity

to absorb a photon will contribute to signal maximization. Ideally, the laser pulses would be completely absorbed just as they reach the last molecular layer. The energy of a 450 nm wavelength photon is

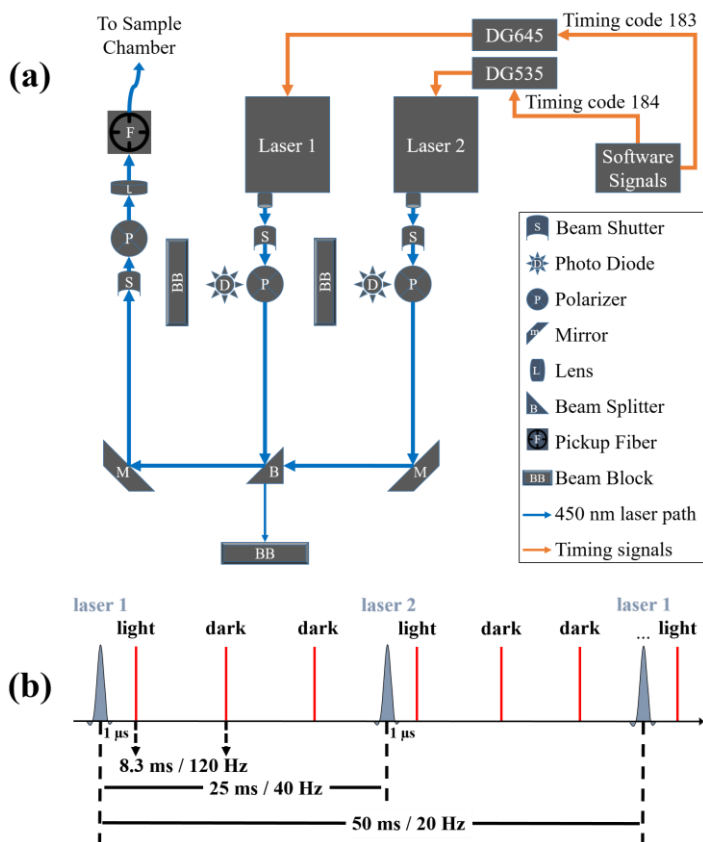
$$E = hv = \frac{hc}{\lambda} = 6.6 \times 10^{-34} \text{ Js} \times \frac{3 \times 10^8 \text{ ms}^{-1}}{450 \times 10^{-9} \text{ m}} = 4.4 \times 10^{-19} \text{ J} \quad (2.39)$$

For a typical PYP microcrystal of 2  $\mu\text{m}$  and a laser fluence of  $800 \times 10^{-6} \text{ J/mm}^2 = 8.0 \times 10^{-10} \text{ J}/\mu\text{m}^2$ , a linear cross-section of  $4 \mu\text{m}^2$  would be exposed to the laser pulse and the crystal will receive a total energy of

$$E_{\text{total}} = (4 \mu\text{m}^2) \times (8.0 \times 10^{-10} \text{ J}/\mu\text{m}^2) = 3.2 \times 10^{-9} \text{ J} \quad (2.40)$$

from  $\sim 7 \times 10^9$  photons. There are approximately  $4 \times 10^7$  molecules of PYP per  $\mu\text{m}^3$  of crystal volume (Tenboer et al. 2014)(supplementary material). Therefore, for a  $2\mu\text{m}$  crystal of  $8 \mu\text{m}^3$  total volume each PYP molecule will receive around 22 photons. When laser pulses are of ns duration each molecule has a significant opportunity to absorb at least one photon and enter the photocycle. If the time extent of the laser pulse is reduced to the fs regime, non-linear effects modify the cross-section and subsequently reduce the penetration depth and photolysis yield. Also, the ultrashort pulses do not permit multiple opportunities for photon absorption. Synchrotron experiments with ps laser pulses on PYP have demonstrated that the level of reaction initiation drops to  $<10\%$  (Schotte et al. 2012) when the laser pulses are of 300 ps duration.

Crystals were grown using the protocol described in §2.2.1. For this study, the time range from 200 ns to 100  $\mu$ s was selected for several reasons: (i) the intermediates  $pR_1$  and  $pR_2$  (Ihee et al. 2005) are considerably populated at this time (Fig. 1.2d), leading to strong DED features; (ii) the  $pR_1$  and  $pR_2$  intermediates are the only intermediates present; (iii) they are relatively long-lived;



**Figure 2.16.** TR-SFX nanosecond PYP experiment laser schematic. (a) Two optical parametric oscillator (OPO) ns lasers were controlled externally with digital delay generators (DDGs). Polarizers were utilized to ensure the laser pulse energies were nearly identical when they arrived at the beam splitter. The pulse energy was adjusted to  $\sim 800 \mu\text{J}/\text{mm}^2$  with the final polarizer before being focused onto the pickup fiber and delivered to the sample. (b) Since a ns laser operating at one-half the LCLS repetition rate of 120 Hz was not available, both ns lasers, each operating at 20 Hz, were phase-shifted so that data could be collected at 40 Hz in a light-dark-dark scheme. Therefore, twice as much dark (no laser excitation) as light (with laser excitation) data was obtained. Fig. 2.16b has been modified from a figure from Tenboer et al. 2014.

(iv) the  $pR_1 / pR_2$  population is largely temperature independent at this time. An additional time-delay of 10 ns was further selected as a challenge to determine if the mixture of the 3 intermediates  $I_{CT}$ ,  $pR_1$  and  $pR_2$  could be determined. Fig. 1.2d shows the fractional concentration of the PYP intermediates as a function of time and the dashed vertical lines denotes the two time-delays at 1  $\mu$ s and 10 ns that were chosen for this experiment.

Since no ns laser that operates at 60 Hz was available at the LCLS, three separate optical parametric oscillator (OPO) ns lasers with 5 ns FWHM pulses running at 20 Hz were planned to be installed. This would have enabled an interleaved light-dark

scheme. The Schmidt lab laser and a laser from BioCARS passed the SLAC safety inspection. The third laser from Lawrence Berkeley National Laboratory did not pass this inspection. However, by time-shifting the operation of the remaining two lasers a 40 Hz data collection strategy could be implemented, as shown in Fig. 2.16b. This light-dark-dark scheme meant that twice as much dark (no laser excitation) as light (with laser excitation) data was collected. DG645 and DG535 *digital delay generators* (DDGs) were used to drive these lasers completely externally. These DDGs received their timing triggers from the radio frequency of the LCLS, labeled as *timing code 183* and *timing code 184* in Fig. 2.16a. To ensure the pulse energies from both lasers were nearly identical, two polarizers were implemented to attenuate their energies so that they matched on average at the beam splitter. The pulses were then adjusted by a third polarizer to dial in the desired final pulse energy before being focused by a single lens onto the pickup fiber and delivered to the sample chamber. Inside the sample chamber two additional lenses, mounted on a motorized X-Y-Z stage, collimated and focused 90% of the  $\sim 15 \mu\text{J}$  laser pulses into an  $\text{\O}150 \mu\text{m}$  diameter spot size. Therefore, the laser fluence at the sample position was  $\sim 800 \mu\text{J} / \text{mm}^2$ . Laser energy was measured at the beginning and end of each shift to ensure fluence was correct and adjustments were made if necessary.

A concern was that although each image was labeled as light and dark in the data header there was no direct confirmation that a laser pulse had actually been created. Therefore, the data might potentially be mixed together. As a sanity check, 3 photo-diodes were used: one for laser 1, a second for laser 2 and a third inside the sample chamber. The diode traces were saved and checked during data reduction so that the dark and light data were properly binned. The two lasers and optics were installed at the CXI instrument by SLAC scientist Dr. Despina Milathianaki and myself.

## 2.5 Dynamics in Biological Macromolecules with Femtosecond Time Resolution

A fs time-resolved study probed the previously unknown intermediates of the *trans* to *cis* isomerization in the PYP photocycle. This experiment introduced new challenges, due primarily to the increased complexity of timing synchronization and data reduction. PYP microcrystals were grown using the protocol outlined in §2.2.1 and the injector fabrication and operation is described in §2.3.3.

The most significant modification was the fs time resolution which required laser pulses of 140 fs duration with a fluence of  $800 \mu\text{J} / \text{mm}^2$  be delivered at the full 120 Hz repetition rate of the LCLS. The pump for the TR-SFX experiment was a *titanium sapphire* (Ti:sapph) Coherent Legend CPA system (Coherent Inc.) which generated 40 fs laser pulses at an 800 nm wavelength and 120 Hz. These pulses were then split by a beam splitter so that a portion of the pulse train was sent to the *timing tool* (see §2.2.2) and the majority to a HE-TOPAS-Prime (Light Conversion Ltd.) parametric amplifier. The HE-TOPAS-Prime supplied 5  $\mu\text{J}$  pulses at an output wavelength of 1800 nm. Frequency doubling through a pair of *barium borate* (BBO) crystals delivered the desired final wavelength of 450 nm. These pulses were directed into the evacuated chamber by a series of mirrors before final focusing into an  $\text{\O}90 \mu\text{m}$  spot by a toroidal mirror on-axis with the X-rays. Chamber windows were selected so that a final chirped pulse length of 140 fs was achieved from the 40 fs initial pulse. The laser pulse characteristics described above, in addition to spectral distribution and phase, were obtained by the method of *second harmonic generation frequency-resolved optical gating* (SHG-FROG) (Kane and Trebino 1993; DeLong et al. 1997). These data were analyzed by Dr. Jasper van Thor and Dr. Chris Hutchison from Imperial College London using a homebuilt SHG-FROG instrument they installed at CXI.

Five time-delays were selected: 300 fs, 600 fs, 900 fs, 3 ps and 200 ns. The 200 ns time-delay was chosen as a control to compare with the 1  $\mu$ s data obtained previously from the ns PYP experiment. The indexed diffraction snapshots were sorted into 16 *timing bins* with ~22,000 images in each bin ranging from 100 fs to 1.2 ps. These images were then merged and the integrated intensities obtained for each reflection by *Process\_hkl*, as discussed previously.

To refine the structures, the 16 *timing bins* were combined into 8 *timing bins* with 40,000 images in each bin. Determination of interpretable DED maps at ultrafast timescales was complicated by the lack of suitable restraints for the unknown structures that accumulate in hundreds of fs. All known restraints are compiled from structures at equilibrium and were, therefore, not appropriate for this refinement. Since no restraints were available at these ultrafast timescales, excited state *quantum mechanics / molecular mechanics* (QM/MM) simulations were performed by our collaborator, Dr. Gerrit Groenhof (Groenhof et al. 2004; Groenhof 2013a; Groenhof 2013b) from the University of Jyväskylä, to guide the refinement process. The structures were refined using the calculated QM/MM restraints. Finally, extrapolated DED maps were obtained. An SVD analysis was applied to the DEDs by Prof. Schmidt to extract the time-dependent kinetic information.

## **2.6 Structure Determination of the Native Form of ccNiR**

The method of single crystal monochromatic X-ray diffraction was used to determine the native form structure of ccNiR. Data was collected at the BioCARS 14-BM-C beamline. A 150  $\mu$ m x 250  $\mu$ m crystal was harvested from the hanging drop and placed into the cryobuffer (see §2.1.8). The crystal was then transferred out of the cryobuffer using a cryoloop and magnetically mounted onto the goniometer into a 100 K nitrogen cryostream and quickly frozen. Images were collected using the *rotation method* (see §2.1.9).

## 2.7 Structure Determination of the Nitrite-bound Form of ccNiR

To determine the structure of the nitrite-bound form of ccNiR, data was collected using SBL1. The cryobuffer utilized for the native ccNiR data set did not produce desirable results when used with the nitrite-bound crystals. The crystals fractured apart immediately when placed into the cryobuffer and therefore could not be used for data collection. Attempts were made to produce a new cryobuffer: 20 mmol/L Na-nitrite was added to the existing cryobuffer; 15-30% concentrations of PEG 4k and PEG 8k with and without Na-nitrite were made; cryobuffers without TEA were also attempted. In each of these cryobuffer variations the crystals still broke apart. Since only a few crystals were remaining after these tests, a crystal was harvested from a drop and placed immediately into the 100 K nitrogen cryostream and frozen without any cryobuffer. The crystal appeared to have survived the freezing process and the mother liquor remained clear. No obvious icing had occurred. Test images revealed that the crystal still diffracted to better than 3 Å and a complete dataset was recorded. Images were collected using the *rotation method*.

Data reduction was performed as discussed previously, however when the resulting electron density map was displayed in 'Coot' it was immediately apparent that it did not fit with the 3UBR model. Therefore, *molecular replacement* was performed with 'Phaser' as described in §2.1.10. The subsequent model was then combined with the observed *structure factor amplitudes* and refined using 'Refmac', as described in §2.1.11.

## 3 Results

### 3.1 PYP Structural Dynamics with Nanosecond Time Resolution

The TR-SFX ns experiment on PYP generated large amounts of data which resulted in interpretable high quality DED maps at near atomic spatial resolution. PYP microcrystals in the 1-5  $\mu\text{m}$  range were produced by the previously described protocol and GDVNs were constructed using an  $\text{\O}75$   $\mu\text{m}$  ID Polymicro capillary to deliver the sample mixture. A Shimadzu LC-20AD solvent delivery pump supplied the pressure for the microcrystal samples. The study used over 5 g of purified protein to create enough microcrystals of size  $\sim 2 \times 2 \times 5$   $\mu\text{m}^3$ . A final microcrystal density of  $10^{10}$ - $10^{11}$  PYP crystals/mL was estimated using a Neubauer counting chamber. The crystal mixture was filtered through a 10  $\mu\text{m}$  stainless steel frit immediately before  $\sim 2$  mL was loaded into a hydraulic syringe sample reservoir and jetted through the GDVN. Jet diameters ranged from 3-5  $\mu\text{m}$  using a flow rate of 22-25  $\mu\text{L}/\text{minute}$ , meaning that almost 16 mL of crystal solution would be needed each shift if no problems occurred and the experiment ran nonstop. In reality, however, sample exchanges usually took 15-30 minutes to complete, X-rays went down due to unforeseen problems and nozzles had to be replaced due to clogging or breakage—all of these factors contributed to a decreased amount of sample needed during each shift.

The dark state model for PYP was obtained from PDB entry 2PHY. This dark state model, PYP<sub>AS</sub><sup>D</sup>, was determined from PYP crystals grown using *ammonium sulfate* (AS) as a precipitant. We used *Na-malonate* (NM) as a precipitant because AS was suspected to contribute too much background scattering at the detector and could also prove difficult to jet. Large PYP crystals were easily grown using NM. Two important concerns were raised: (i) are the crystals grown using NM as a precipitant identical to those from AS and (ii) are these crystals still photo-active. To answer



the first question, I obtained a dataset on a NM PYP crystal using SBL1 to 1.56 Å resolution. The dark state model,  $PYP_{NM}^D$ , agreed perfectly with the  $PYP_{AS}^D$  model. The *cell constants* a, b and c and  $\alpha$ ,  $\beta$  and  $\gamma$  were identical. Further, the TR-AS results I acquired with SBL2 confirmed that NM PYP crystals were indeed photo-active (results not presented). The refined  $PYP_{NM}^D$  model phases could now be used since  $PYP_{AS}^D$  was identical to  $PYP_{NM}^D$ .

Timing was straightforward, owing to the ns nature of the experiment and the long time-delays selected. No time-binning of the data was required and the value of the *ns Target Time* was the nominal time-delay. The CSPAD detector was positioned as close as possible to the sample stream (75.1 mm) so that diffraction data with the maximum resolution could be obtained. An 8.995 keV XFEL beam energy with a 1µm focal spot size was also selected. With this geometry and beam energy the resolution at the edge of the detector was 1.6 Å. To enable beam transmissions of 30% an electronic attenuation mask was applied to the detector for those pixels exposed to the intense low resolution (small angle) reflections up to 4 Å. This mask effectively increased the dynamic range of the detector by preventing pixel saturation at low resolution while maintaining single photon sensitivity at higher resolution (large angle).

In order to produce time-resolved DED maps (§2.2.4) we recorded 3,695,472 images at CXI over 5 twelve hour shifts. 2,463,620 of these images were dark data, while 1,231,852 images were collected with laser activation and were light data. *Cheetah* selected 170,911 (6.94%) of the dark frames as crystal hits, and 82,627 (6.71%) of the light frames were hits. Therefore, a hit-rate of 6.86% was achieved over both light and dark data combined. A summary of the results from data reduction is given in Tab. 3.1.

**Table 3.1.** Result summary from the TR-SFX ns PYP experiment. The columns labeled as ‘Dark’ mean that no laser excitation was used. The columns labeled as ‘Light’ mean that laser excitation was present using a wavelength of 450 nm. Parentheses: highest resolution shell (1.64 Å-1.6 Å). Angled brackets: <mean values>. Table has been modified from Tenboer et al. 2014.

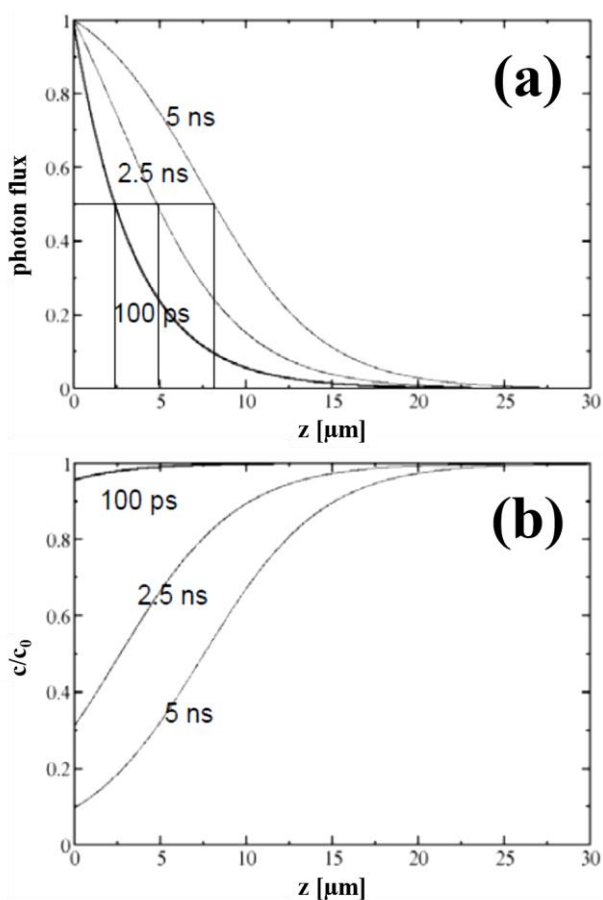
	10 ns		1 μs	
	Dark	Light	Dark	Light
<b>laser Fluence [μJ / mm<sup>2</sup>]</b>	off	800	off	800
<b># images</b>	414,326	207,168	2,049,294	1,024,684
<b># hits</b>	75,757	36,632	95,154	45,995
<b># indexed from hits</b>	46,460 {61.3%}	22,678 {61.9%}	64,496 {67.8%}	31,255 {67.9%}
<b>unit cell dimensions (a, b, c) [Å]<sup>a</sup></b>	66.96, 66.96, 40.98	66.99, 66.90, 41.00	66.97, 66.97, 40.99	66.99, 66.99, 41.00
<b>completeness [%]</b>	98.62	96.62	99.45	98.82
<b>redundancy<sup>b</sup></b>	2,112 (1,329)	1,037 (650)	2,924 (1,853)	1,420 (903)
<b>I/σ(I)</b>	6.0 (2.3)	4.3 (1.7)	12.4 (4.7)	8.7 (3.0)
<b>R<sub>split</sub> [%]<sup>c</sup></b>	14.4 (49.8)	20.2 (86.1)	6.5 (22.4)	9.2 (38.6)
<b>R<sub>cryst</sub> [%]<sup>d</sup></b>	18.8	nn	17.8	nn
<b>R<sub>free</sub> [%]<sup>d</sup></b>	21.1	nn	20.8	nn
<b>&lt; ΔF &gt;/&lt;σ<sub>AF</sub>&gt;</b>	9.1 / 9.1		6.4 / 5.1	
<b>max/min DED [σ]</b>	+6.0 / -12.4		+18.3 / -22.4	
<b>reaction initiation</b>	nd		40% total 18% pR <sub>1</sub> 22% pR <sub>2</sub>	

<sup>a</sup>unit cell angles:  $\alpha=\beta=90^\circ$ ,  $\gamma=120^\circ$ ; <sup>b</sup>average number of observations per reflection (and symmetry equivalent); <sup>c</sup>see Eqn. 2.36; <sup>d</sup>see Eqn. 2.26; **nn**: not necessary; **nd**: not done

completely through the PYP microcrystals, as shown by the simulations in Fig. 3.1 and supported by these results. The extent of reaction initiation was not determined for the 10 ns time-delay since the amount of data did not support an analysis. At this time-delay a mixture of 3 intermediates is present and the fractional concentrations are varying quickly. Therefore, many more data sets must be collected at multiple time-delays in order to unmix these structures.

The space group was determined to be P6<sub>3</sub> and diffraction data extended to a resolution of 1.6 Å. A total reaction initiation of 40% was obtained (see §2.2.7) for the 1 μs time-delay between pR<sub>1</sub> and pR<sub>2</sub> at 18% and 22%, respectively. The ns laser pulses penetrated

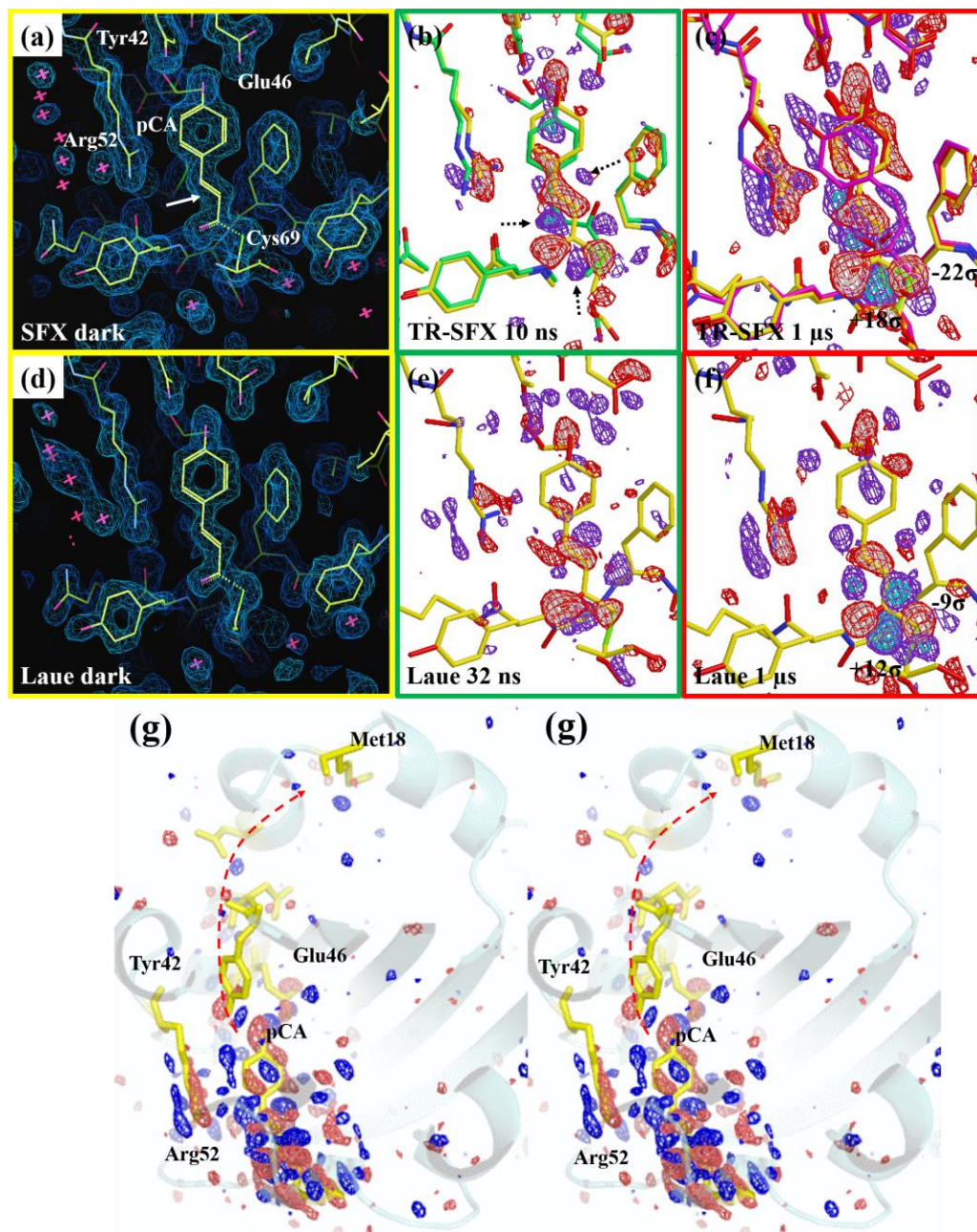
The light-minus-dark DED maps at the 1  $\mu\text{s}$  time-delay are of exquisite quality with positive



**Figure 3.1.** Simulation of ns laser penetration depth for PYP. A laser flux of  $\sim 800 \mu\text{J}/\text{mm}^2$  at 450 nm wavelength with a rectangular-shaped beam profile is assumed; extinction coefficient is  $45,500 \text{ cm}^2/\text{mmol}$  for the PYP crystal and a concentration of  $63 \text{ mmol/L}$  is used; 10% of the molecules are assumed to enter the photocycle. (a) The curves demonstrate the penetration depth of the 5 ns FWHM laser pulse at times of 100 ps, 2.5 ns and 5 ns. The vertical lines in the 3-sectioned rectangle intersect with each curve along the top horizontal line of the rectangle where half of the photons have been absorbed. (b) Percentage of bleached PYP molecules plotted vs. the penetration depth. For the 5 ns FWHM laser pulse used for the TR-SFX ns PYP experiment  $>80\%$  of the PYP molecules are bleached up to a 4  $\mu\text{m}$  penetration depth. Fig. 3.1 has been modified from a figure from Tenboer et al. 2014 (supplementary material).

and negative features extending up to  $18\sigma$  and  $-22\sigma$ , respectively, where  $\sigma$  is the root mean squared (RMS) value of the DED over the asymmetric unit. These results are shown in Fig. 3.2b-Fig. 3.2c and Fig. 3.2g. The largest DED signals are located around the significantly displaced sulfur of the chromophore at the  $1 \mu\text{s}$  time-delay. The *trans* to *cis* isomerization causes conformational changes that extend to the edge of the protein, which is observed in the DED map shown in Fig. 3.2g. A new conformation of the Glu46 residue was determined during refinement that had not been previously reported by synchrotron experiments (see Fig. 3.3). This will be discussed further below.

The TR-SFX dark structure to  $1.6 \text{ \AA}$  is of comparable quality to that obtained with TR-LX, as shown in Fig. 3.2a & 3.2d, respectively. The dark structure, PDB entry 2PHY (Borgstahl et al. 1995), may be used to interpret both maps. The  $1 \mu\text{s}$  and 10 ns

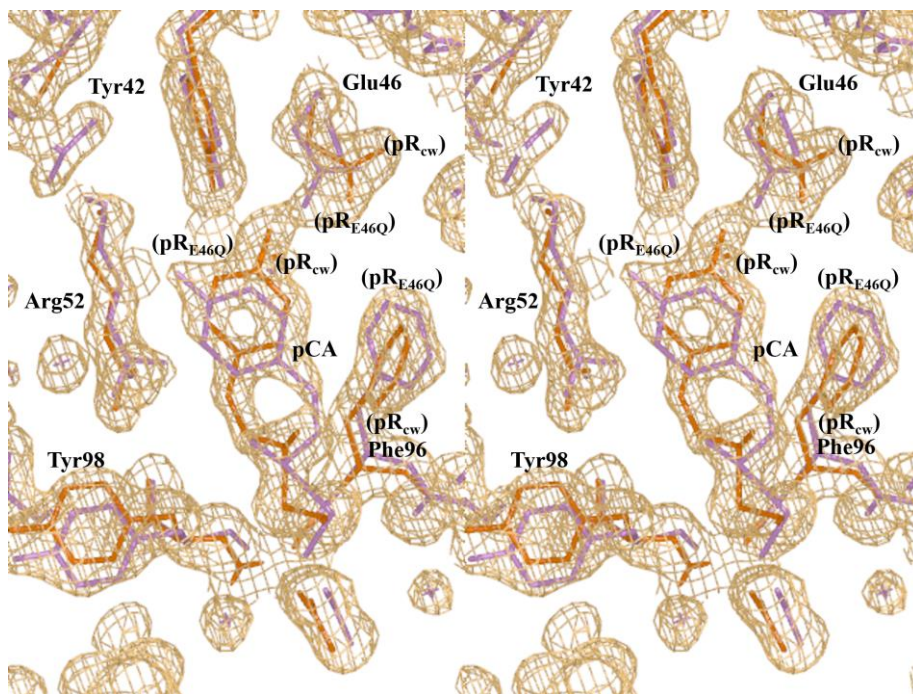


**Figure 3.2.** Yellow lines in all maps: dark state model. (a) TR-SFX electron density map of dark state model ( $1.1\sigma$  contour level) to a resolution of  $1.6 \text{ \AA}$ . The pCA chromophore and some surrounding residues are labeled. White arrow: location of the C2=C3 double bond. (b) TR-SFX DED map for 10 ns time-delay; green lines: structure of the  $I_{CT}$  intermediate; dashed arrows: features that belong to other intermediates not shown here. (c)  $1 \mu\text{s}$  time-delay, TR-SFX DED map; pink lines: pR1 intermediate; red lines: pR2 intermediate. (d) TR-LX electron density map of dark state model ( $1.1\sigma$  contour level) to a resolution of  $1.6 \text{ \AA}$ . (e) The 32 ns time-delay, TR-LX DED map and  $I_{CT}$  structure corresponds best with 10 ns TR-SFX  $I_{CT}$  intermediate. (f)  $1 \mu\text{s}$  TR-LX DED map. Note the high level of reaction initiation for TR-SFX. (g) Stereo view of  $1 \mu\text{s}$  TR-SFX map overlaid on dark state model. Dashed red arrow: DED features extend to Met18; Contour levels: (b) red/blue  $-3\sigma/+3\sigma$  (c) red/white  $-3\sigma/-4\sigma$  and blue/cyan  $+3\sigma/+5\sigma$  (g) red/blue  $-3\sigma/+3\sigma$ . Panels modified from Tenboer et al. 2014.



intermediate structures recorded by TR-SFX were compared with synchrotron results and the 10 ns TR-SFX DED map agreed best with the 32 ns TR-LX DED map. Readily apparent in Fig. 3.2b & 3.2c is the high level of reaction initiation that was achieved by TR-SFX.

These results show that a reaction initiation of 40% is possible (§2.2.7) using a ns laser with TR-SFX. A benefit of this high level of reaction initiation for photo-active proteins is illustrated in Fig. 3.2c and Fig. 3.3. Positive and negative DED features were observed at the 1  $\mu$ s time-delay near the Glu46 residue not explained by the Laue model from PDB entry 1TS7 (Ihee et al. 2005). However, a new conformation of Glu46 was determined during a restrained refinement using 'Refmac'. The extrapolated EDM for this structure is shown in Fig. 3.3. This experiment demonstrates that when comparing DED maps obtained from the same photo-reactive protein at similar spatial resolution, a clear advantage of an XFEL over a synchrotron is evident. The small



**Figure 3.3.** Stereo representation of the resulting extrapolated electron density for the 1  $\mu$ s time-delay was determined as described in §2.2.6. Contour level: 1.1 $\sigma$ . Pink lines: pR<sub>1</sub> with 22% fractional concentration; Red lines: pR<sub>2</sub> with 18% fractional concentration. Fig. 3.3 from Tenboer et al. 2014.

crystals that may be studied with XFELs enable more thorough laser excitation. The resulting larger extent of reaction initiation may enable the determination of new conformations

of nearby residues that are difficult, if not impossible with TR-LX.

### 3.2 PYP Structural Dynamics with Femtosecond Time Resolution

To reveal the fs dynamics of PYP, a total of 533,705 indexed images were recorded over the course of the beamtime, 65,289 of these images were dark data, while 468,416 snapshots were collected with laser activation. The *timing tool* information was used to sort the light data into 16 *timing bins* with ~22,000 images each from ~100 fs to ~1 ps. Another bin at 3ps contained 61,659 frames. 65,298 images were additionally obtained at a time-delay of 200 ns. A summary of results from the data reduction is given in Tab. 3.2. Sample flow rates, jet diameters, microcrystal sizes and sample densities were unchanged from the ns PYP experiment. The sample-to-detector distance was 84 mm and the XFEL beam energy was 9.5 keV. The X-rays were focused to a 1 $\mu$ m spot size and an electronic attenuation mask was applied to the CSPAD detector up to a resolution of 4.5 Å.

To determine the level of reaction initiation that could be achieved with these short laser pulses, two spectroscopic studies were performed in advance on crushed single crystals of PYP. The first experiment was conducted by a team led by Jasper van Thor using 100 fs and 300 fs laser pulses. This experiment concluded that the highest level of reaction initiation could be achieved with laser pulse lengths of ~100 fs. An additional experiment was performed by myself at the BioCARS 14-ID-B beamline using 900 fs laser pulses and the Xenon light source and spectrometer obtained from SBL2. Results from both experiments demonstrated that a 7-10% reaction initiation could be obtained from an 800  $\mu$ J / mm<sup>2</sup> laser fluence without permanent photo-bleaching of the sample using laser pulses < 1 ps in duration.

The dark reference PYP structure, PYP<sub>obs</sub><sup>D</sup>, was determined from images collected without laser excitation and the resulting EDM was identical to the dark map obtained during the ns TR-SFX

**Table 3.2.** Result summary from TR-SFX fs PYP experiment. Time-delays are the average of all calculated time-delays from the timing tool in each bin. Overall resolution is 1.6 Å. Parentheses: highest resolution shell (1.6-1.67 Å). Table modified from Pande et al. 2016.

Time-Delay [ps]	#indexed hits	R <sub>split</sub>	Completeness	Multiplicity
<b>Dark</b>	65,289	7.69 (41.08)	99.93 (100)	991.51 (22.1)
<b>0.1008</b>	19,313	12.84 (37.42)	99.84 (100)	342.76 (25.2)
<b>0.1828</b>	19,659	12.58 (38.60)	99.84 (100)	348.41 (26.3)
<b>0.2363</b>	19,647	12.93 (41.08)	100 (100)	341.26 (26.0)
<b>0.3015</b>	19,480	13.19 (36.77)	99.76 (100)	330.86 (25.3)
<b>0.3907</b>	19,090	13.70 (43.08)	99.92 (100)	312.83 (22.2)
<b>0.5210</b>	18,905	14.30 (35.20)	99.97 (100)	295.09 (30.5)
<b>0.6575</b>	21,500	11.83 (25.12)	99.84 (100)	409.97 (59.1)
<b>0.7395</b>	22,554	10.99 (21.89)	99.76 (100)	459.67 (68.2)
<b>0.7824</b>	22,816	10.95 (23.76)	99.84 (100)	468.66 (69.0)
<b>0.8149</b>	22,560	10.83 (23.70)	100 (100)	467.46 (64.6)
<b>0.8428</b>	22,507	10.97 (24.63)	99.84 (100)	463.5 (63.0)
<b>0.8701</b>	22,416	10.76 (24.18)	99.6 (100)	465.45 (60.0)
<b>0.8988</b>	22,266	10.54 (23.59)	99.84 (100)	466.38 (61.8)
<b>0.9317</b>	22,409	10.49 (21.76)	100 (100)	472.31 (64.4)
<b>0.9784</b>	22,732	10.44 (23.46)	99.84 (100)	478.35 (67.2)
<b>1.0648</b>	23,605	10.43 (20.94)	99.92 (100)	501.62 (75.9)
<b>3.0</b>	61,659	5.43	99.9	2023
<b>200,000</b>	65,298	6.99	99.9	1137

PYP experiment. To verify that a sufficient extent of reaction initiation was achieved with the 140 fs laser pulses, the 200 ns time-delay was chosen as a positive control. A DED map was calculated for the 200 ns time-delay, PYP<sub>Δ200ns</sub>, and the structures of pR<sub>1</sub> and pR<sub>2</sub> from PDB entries 4WLA and 4WL9 (Tenboer et al. 2014) were refined against the *extrapolated structure factor amplitudes* (see §2.2.6). The refined structures and resulting DED maps at the 200 ns control time-delay agree well with those obtained previously at the 1 μs time-delay during the ns PYP experiment (see Fig. 3.1c and Fig. 3.4g). To estimate the extent of reaction initiation a DED map of ‘(pR<sub>1</sub> + pR<sub>2</sub>) minus PYP<sub>obs</sub><sup>D</sup>’ was calculated and then empirically fitted to PYP<sub>Δ200ns</sub>. The

resulting extent of total reaction initiation was 12.6%, which agrees well with the 7-10% determined by the spectroscopic experiments.

Weighted DED maps were calculated for each time-delay,  $wDED_{\Delta t}$ , as were difference *structure factor amplitudes*. An SVD analysis was performed on the  $wDED_{\Delta t}$ . The first right singular vector ( $rSV_1$ ) follows the relaxation of the structure over the 16 *timing bins* and was fitted empirically by the equation

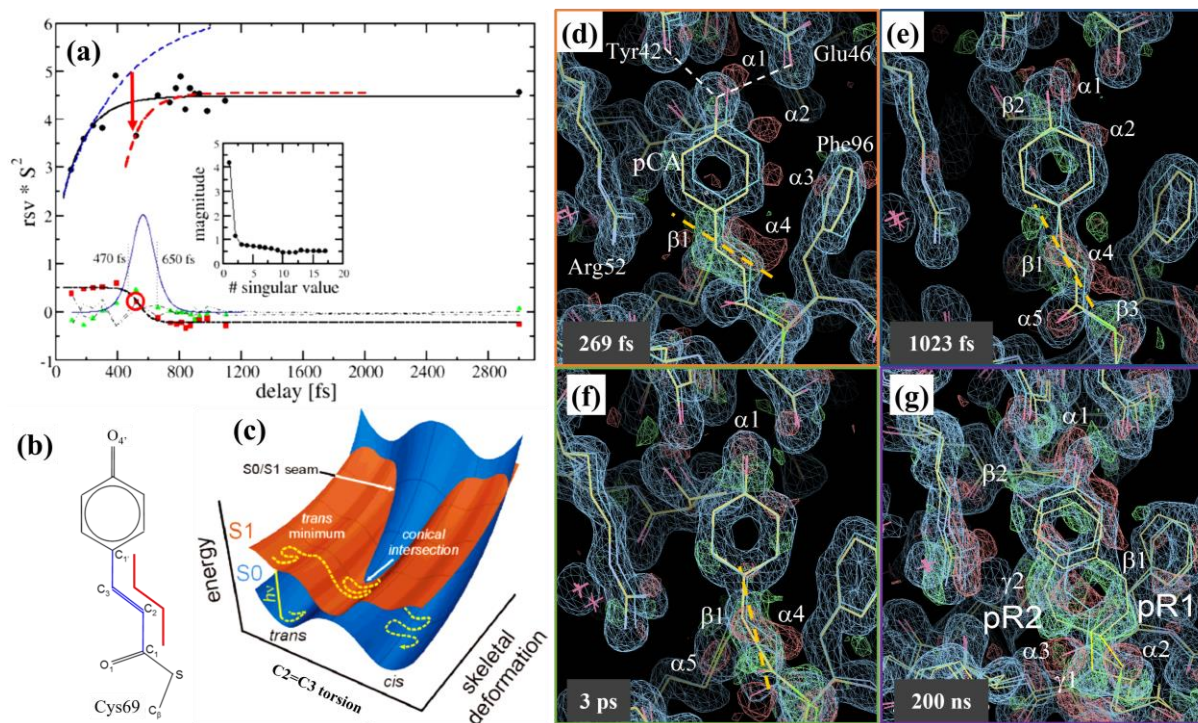
$$rSV_1(t) = A_1 + B_1 \left[ 1 - \exp\left(-\frac{t}{\tau_1}\right) \right] \quad (3.1)$$

The dashed black line in the upper plot of Fig. 3.4a is the fitting curve determined by Equation 3.1 for  $rSV_1$ . At a time  $\tau_1=155$  fs, significant structural changes lead to larger DED features that are observed to last through 3 ps. When the second right singular vector ( $rSV_2$ ) was considered an additional transition was revealed. The transition time was found by empirically fitting the  $rSV_2$  with the equation

$$rSV_2(t) = \frac{A_2}{1 + \exp(t - \tau_2/B_2)} \quad (3.2)$$

The dashed black line in the lower plot of Fig. 3.4a is the fitting curve determined by Equation 3.2 for  $rSV_2$  with a transition time of  $\tau_2=565$  fs. A difference is therefore evident in the DED maps between time-delays  $<500$  fs compared to those  $>600$  fs.  $rSV_1$  shows the average progression of the DED signal which exposes significant structural changes occurring after the  $\sim 140$  fs pump laser pulse passed through the crystal. These changes begin at  $\sim 155$  fs and are plotted by the black circles in Fig. 3.4a.  $rSV_2$  reveals a second process beginning between 500-600 fs and shown by





**Figure 3.4.** The trans-cis isomerization of the PYP chromophore. Panels (d)-(g): the extrapolated  $2F_o - F_c$  maps are displayed in blue at a  $1.1\sigma$  contour level. The weighted DED maps are shown in red/green  $-3\sigma/+3\sigma$ . Dark state model is denoted by blue lines and yellow lines define the refined intermediate structures. (a)  $Rsv_1$  is denoted by the black circles and the fitting curve from Equation 3.1 is plotted with the solid black line. The dashed blue and red lines are the fitting curves for  $rsv_1$  when the solid black curve is separated into two regions. The red arrow indicates the transition from trans to cis.  $Rsv_2$  is plotted with the red squares and the fit from Equation 3.2 is shown by the dashed black line. The red circle surrounding the  $rsv_2$  value at 521 fs designates the time-delay nearest the conical intersection. Green triangles:  $rsv_3$ .  $Rsv_4$  and  $rsv_5$  are displayed with dashed-dotted and dashed-double-dotted lines, respectively. The blue Gaussian curve indicates the transition region from 470-650 fs (FWHM  $\sim 180$  fs). (b) Structure of the pCA chromophore of PYP. The red line defines the C2=C3 torsional angle. (c) The potential energy surfaces of the excited state (S1, orange) and ground state (S0, blue) for the chromophore in the trans-cis isomerization represented as a function of the C2=C3 torsion and deformation of the bonds (labeled as skeletal deformation). The dashed yellow line represents a possible trajectory along these surfaces. The juncture of these energy surfaces forms a conical intersection where trajectories may either return to the dark state or continue towards the cis configuration. (d) DED map at a time-delay of 296 fs after laser initiation. The dashed white lines show hydrogen bonds to Glu46 and Tyr42. The dashed orange line represents the direction of elongation of the positive feature  $\beta_1$ . (e) DED map at 1023 fs. The dashed orange line again denotes the direction of elongation for positive feature  $\beta_1$ . (f) DED map at 3 ps. The dashed orange line has the same meaning. (g) DED map at 200 ns, which was used as a control to compare to DED maps from previous experiment LD62. The results from the two experiments agree well. The two intermediates, pR1 and pR2 (models not shown), are defined by the difference electron density features  $\beta_1$  and  $\beta_2$  (pR1) and  $\gamma_1$  and  $\gamma_2$  (pR2). Panels (a)-(b) & (d)-(g) from Pande et al. 2016. Panel (c) from Groenhof et al. 2003.

the red squares in Fig. 3.4a. This distinction is evident when comparing the DED maps in Fig. 3.4d for time bins <500 fs to those from Fig. 3.4e for time bins >600 fs.

Upon closer examination of this region we noted that the  $rSV_1$  values for the time-delays at 390 fs and 521 fs deviated largely from the calculated fit, as shown in Fig. 3.4a. Since all of the data was scaled together before they were separated into *timing bins* and the deviations of  $rSV_1$  was small except for two time-delays, this suggested that  $rSV_1$  should be separated into two parts. The dashed blue and red lines in Fig. 3.4a represent the separated  $rSV_1$  fitting curve. For fast time points (<390 fs) the molecules follow the dashed blue line. At ~500 fs the *trans-cis* isomerization evolves with a portion of the molecules falling back to the ground state while others relax along the dashed red line. The center of this transition lies around ~521 fs as shown by the circled red square in Fig. 3.4a. The solid blue curve denotes the period of the transition and the FWHM duration of the evolution is ~180 fs. We identify this transition with the *trans* to *cis* isomerization of the pCA chromophore.

### 3.3 Structure Determination of the Native Form of ccNiR to 1.65 Å

Data was obtained from the native form crystal of ccNiR using the following parameters: the detector was an ADSC Quantum-315; the sample-to-detector distance was 315 mm; the beam energy was 12.668 keV ( $\lambda=0.9787$  Å); a  $0.2^\circ$  rotation width was used; the exposure time was 11 seconds; a total of 600 images were collected covering a total angular rotation of  $120^\circ$ . Data reduction and structural refinement was performed as discussed previously. Phases were provided by PDB entry 3UBR (Youngblut et al. 2012). The selected space group solution was  $P2_12_12_1$  with an overall *mosaicity* of  $0.27^\circ$  and a resolution ranging from 30.77-1.65 Å. A result summary from the data reduction is given in Tab. 3.3. The refined unit cell parameters were  $a=50.18$  Å,  $b=96.07$  Å,  $c=222.60$  Å,  $\alpha=90.0^\circ$ ,  $\beta=90.0^\circ$  and  $\gamma=90.0^\circ$  from these data. The unit cell parameters from the

Laue structure PDB entry 3UBR (Youngblut et al. 2012) are  $a=51.50 \text{ \AA}$ ,  $b=95.90 \text{ \AA}$ ,  $c=223.80 \text{ \AA}$ ,

**Table 3.3.** Result summary from the structure determination of the native form of ccNiR.

	Overall	Inner Shell (low resolution)	Outer Shell (high resolution)
<b>Resolution limits</b> [ $\text{\AA}$ ]	N/A	30.77	1.65
<b>R<sub>merge</sub></b>	0.091	0.071	0.501
<b>R<sub>cryst</sub></b>	0.2234	N/A	N/A
<b>R<sub>free</sub></b>	0.2577	N/A	N/A
<b># of observed reflections</b>	382,646	16,324	7,338
<b># of unique reflections</b>	97,423	4,049	4,062
<b>unit cell dimensions</b> (a, b, c) [ $\text{\AA}$ ]	50.18, 96.07, 222.60	N/A	N/A
<b>unit cell angles</b> ( $\alpha$ , $\beta$ , $\gamma$ ) [ $^\circ$ ]	90.0, 90.0, 90.0	N/A	N/A
<b>mosaicity</b> [ $^\circ$ ]	0.27	N/A	N/A
<b>completeness</b> [%]	74.9	91.2	22.0
<b>multiplicity</b> [%]	3.9	4.0	1.8

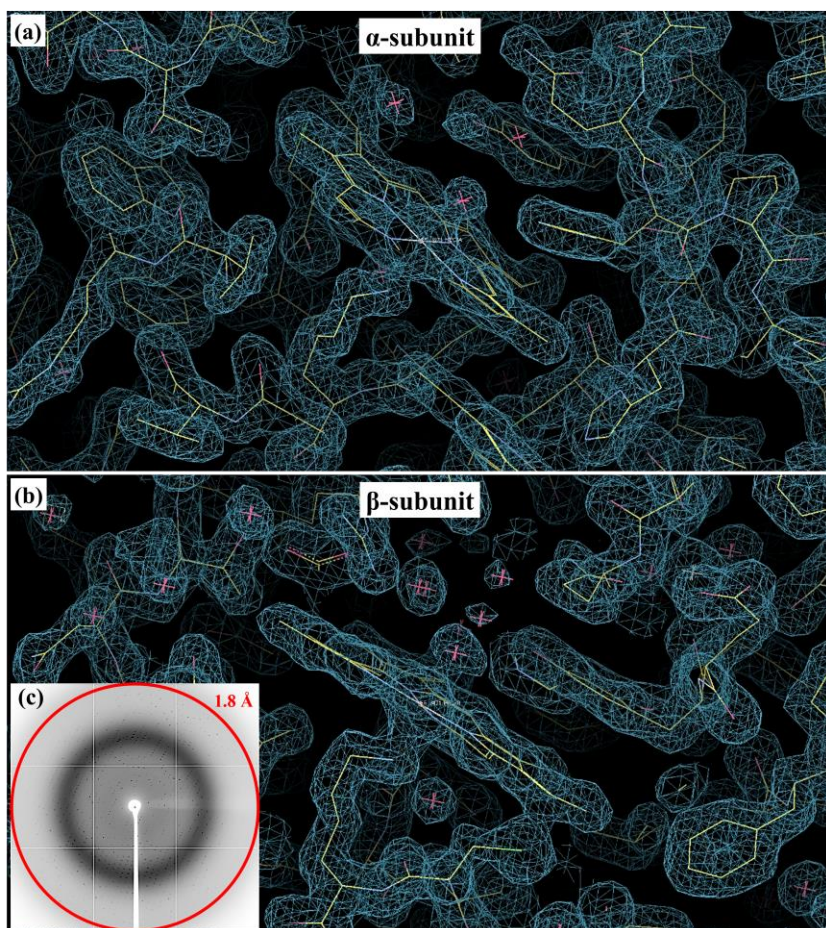
N/A: Not applicable

these data were obtained at a cryogenic temperature of 100 K. An average mosaicity of  $0.27^\circ$  is excellent and suggests the crystal was adequately protected by the cryobuffer during freezing. Previous synchrotron experiments on ccNiR at BioCARS 14-BM-C revealed difficulties with finding a suitable cryobuffer in which the crystals did not disintegrate. Those crystals that did not fracture showed an average mosaicity of  $0.9^\circ$  with a maximum resolution of  $3 \text{ \AA}$ . Substantial effort produced the cryobuffer described in §2.1.8 and the low mosaicity observed from these data suggest this cryo-protectant should be used for any future studies on the native form ccNiR at cryogenic temperatures. Additionally, the improved resolution of  $1.65 \text{ \AA}$  supports the revised crystal growth protocol developed by UWM Chemistry graduate student Dan Pauly.

$\alpha=90.0^\circ$ ,  $\beta=90.0^\circ$  and  $\gamma=90.0^\circ$ . Although the cell constants a, b and c differ slightly from the model they are within acceptable tolerances considering the 3UBR model was collected from a crystal at room temperature while



The  $R_{\text{merge}}$  value was 0.091 overall, 0.071 in the inner shell (low resolution) and 0.501 in the outer shell (high resolution). Although the overall and inner shell  $R_{\text{merge}}$  values are excellent, the outer shell is particularly high. This may be explained by considering both the completeness and



**Figure 3.5.** Structure determination of the native form of ccNiR. (a) The blue  $2F_o-F_c$  electron density map is centered on the active site heme in the  $\alpha$ -subunit. The contour level is  $1.1 \sigma$ . The red crosses surrounded by electron density features show the locations of water molecules. Yellow lines: structural model of native ccNiR. (b) The second active site heme found in the  $\beta$ -subunit is shown. Red crosses once again denote the location water molecules. (c) A typical diffraction image obtained from this experiment is shown. The red circle identifies a resolution of  $1.8 \text{ \AA}$ .

the multiplicity across all resolution shells. The overall completeness is 74.9% with a multiplicity of 3.9. The inner shell completeness is 91.2% with a multiplicity of 4.0. The outer shell completeness is 22.0% with a multiplicity of 1.8. The overall completeness is low while the inner shell is significantly better at

91.2%. However, since the outer shell completeness is only 22.0% with a multiplicity of 1.8 the

overall completeness will undoubtedly be negatively affected. The low completeness of the outer shell reflections could have been dramatically improved if more images had been collected beyond the  $120^\circ$  total angular width used for this experiment. Regardless, the  $1.65 \text{ \AA}$  resolution  $2F_o-F_c$  (2

x observed *structure factors* minus calculated *structure factors*) EDM is of reasonable quality. Waters are observed at both active site hemes as seen in Fig. 3.3, in agreement with previous results (Einsle et al. 2002). The waters were not imposed by the model since they were removed from the PDB file before refinement.

### 3.4 Structure Determination of the Nitrite-bound Form of ccNiR to 2.59 Å

Data was obtained from the nitrite-bound form crystal of ccNiR using the following parameters: the detector was a Rayonix CCD165 with a sample-to-detector distance of 90 mm; the beam energy was 8.04 keV ( $\lambda=1.5418$  Å); a  $0.3^\circ$  rotation width was used; the exposure time was 300 seconds; a total of 600 images were collected covering a total angular rotation of  $180^\circ$ .

To determine the nitrite-bound structure of ccNiR the data was reduced using standard methods. The space group solution was  $P2_12_12_1$  with an overall *mosaicity* of  $0.61^\circ$  and a resolution

**Table 3.4.** Result summary from the structure determination of the nitrite-bound form of ccNiR.

	Overall	Inner Shell (low resolution)	Outer Shell (high resolution)
<b>Resolution limits</b> [Å]	N/A	29.55	2.59
<b>R<sub>merge</sub></b>	0.196	0.178	0.449
<b>R<sub>cryst</sub></b>	0.3569	N/A	N/A
<b>R<sub>free</sub></b>	0.3770	N/A	N/A
<b># of observed reflections</b>	156,466	5,585	17,892
<b># of unique reflections</b>	27,428	1,016	3,495
<b>unit cell dimensions</b> (a, b, c) [Å]	46.18, 92.00, 210.36	N/A	N/A
<b>unit cell angles</b> ( $\alpha$ , $\beta$ , $\gamma$ ) [ $^\circ$ ]	90.0, 90.0, 90.0	N/A	N/A
<b>mosaicity</b> [ $^\circ$ ]	0.61	N/A	N/A
<b>completeness</b> [%]	95.1	97.5	84.4
<b>multiplicity</b> [%]	5.5	5.7	5.1

N/A: Not applicable

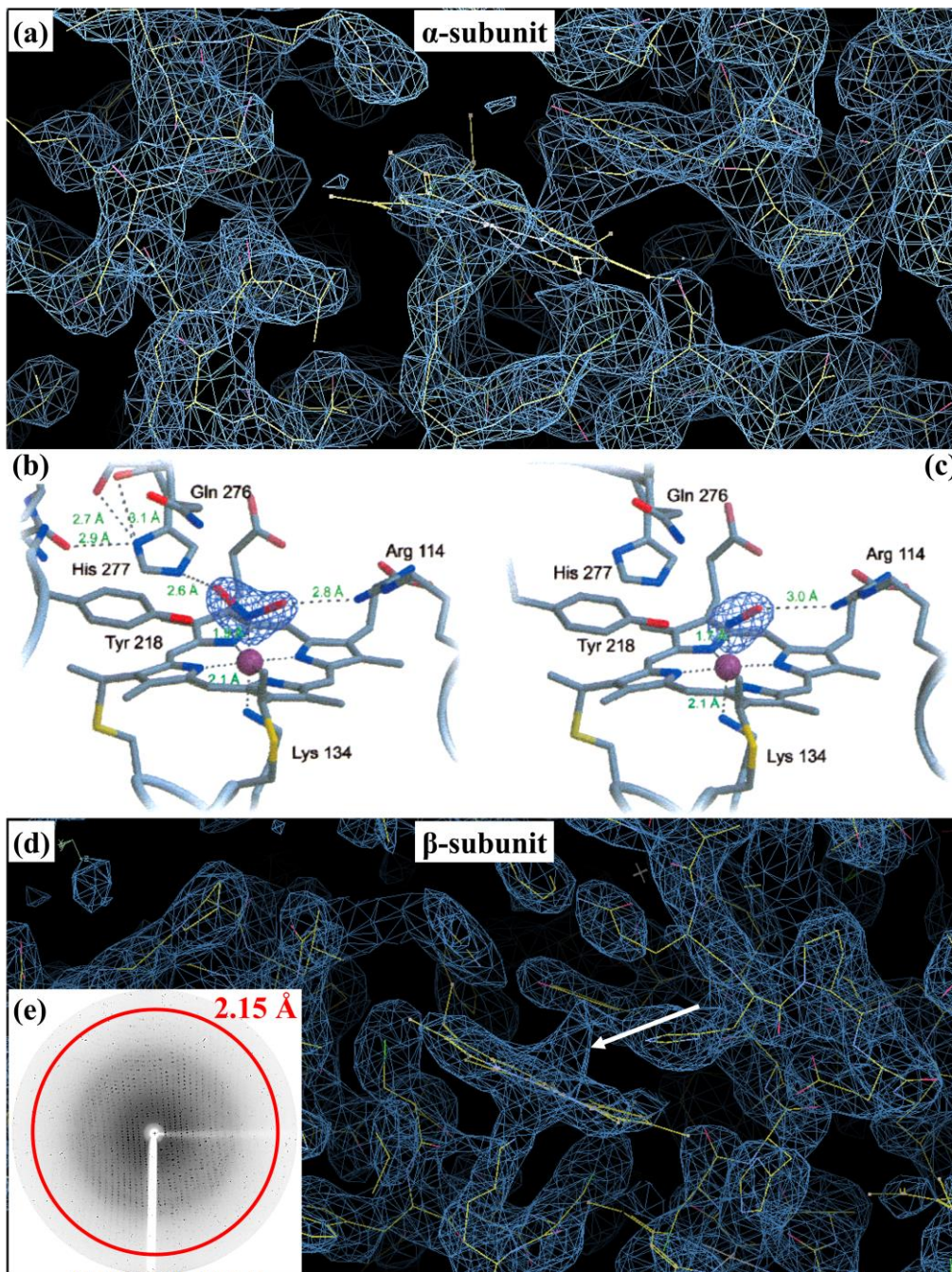
ranging from 29.55-2.59 Å. A result summary from the data reduction is given in Tab. 3.4. After refinement, an electron density feature not explained by the 3UBR model was observed at the active site heme located in the  $\beta$ -subunit, referenced by the white

arrow in Fig. 3.5. The refined *unit cell parameters* are  $a=46.18 \text{ \AA}$ ,  $b=92.00 \text{ \AA}$ ,  $c=210.36 \text{ \AA}$ ,  $\alpha=90.0^\circ$ ,  $\beta=90.0^\circ$  and  $\gamma=90.0^\circ$  from these data. The *unit cell parameters* from the Laue structure (PDB entry 3UBR (Youngblut et al. 2012)) are  $a= 51.50 \text{ \AA}$ ,  $b=95.90 \text{ \AA}$ ,  $c=223.80 \text{ \AA}$ ,  $\alpha=90.0^\circ$ ,  $\beta=90.0^\circ$  and  $\gamma=90.0^\circ$ . The cell constants  $a$ ,  $b$  and  $c$  are greatly deviated from the 3UBR model with  $\Delta a= -5.32 \text{ \AA}$ ,  $\Delta b= -3.9 \text{ \AA}$  and  $\Delta c= -13.44 \text{ \AA}$ . These large deviations demanded a *molecular replacement* strategy using 'Phaser', which determined a single solution from the observed *structure factor amplitudes* and 3UBR model. The average *mosaicity* of  $0.61^\circ$  is high but acceptable, especially considering that no cryobuffer was used to collect these data. However, any future studies with large single crystals of the nitrite-bound form of ccNiR at cryogenic temperatures will require a suitable cryobuffer to aid the freezing process.

The data quality is poor, with an  $R_{\text{merge}}$  value of 0.196 overall, 0.178 in the inner shell (low resolution) and 0.449 in the outer shell (high resolution). Although the overall and inner shell  $R_{\text{merge}}$  values are on the high range of acceptable, the outer shell is unacceptably high. Unlike the native form ccNiR data discussed in §3.3, however, the completeness and multiplicity suggests that more snapshots would not have significantly improved the values of  $R_{\text{merge}}$ . 'Scala' reported an overall completeness of 95.1% with a multiplicity of 5.5, an inner shell completeness of 97.5% with a multiplicity of 5.7 and an outer shell completeness of 84.4% with a multiplicity of 5.1. The resulting  $2.59 \text{ \AA}$  resolution  $2F_o-F_c$  electron density map shown in Fig. 3.6 is of low quality, still an electron density feature not explained by the native form model is observed at the active site heme of the  $\beta$ -subunit. Unfortunately, the resolution does not permit an interpretation of this feature with high confidence. The feature could simply be water, however the size of the electron density is rather too large. Two other possible explanations for this feature are given in Fig. 3.6b



and Fig. 3.6b which include a bound nitrite and the partially-reduced hydroxylamine form (Einsle



**Figure 3.6.** Structure determination of the nitrite-bound form of ccNiR. (a) Active site heme of the  $\alpha$ -subunit is vacant. (b) Nitrite molecule at the active site heme. The nitrogen of the nitrite molecule is bound to the Fe(III) and the two oxygen establish hydrogen bonds with Arg114 and His277. (c) Partially-reduced hydroxylamine form. The remaining oxygen maintains its hydrogen bond with Arg114. (d) An electron density feature not identified by the 3UBR model is observed at the active site heme of the  $\beta$ -subunit (white arrow). (e) Typical diffraction image. Fig. 3.6b-Fig. 3.6c from Einsle et al. 2002.

et al. 2002), respectively. If the partially-reduced hydroxylamine intermediate is considered then our data suggests the oxygen retains the hydrogen bond with the His277 residue. However, *Einsle et al.* observed that only the hydrogen bond with Arg114 remained intact with the oxygen. The hydroxylamine form appears plausible, considering the length of time needed to obtain the data from SBL1. Photo-reduction of the bound nitrite may have occurred during the nearly 3 days of data collection. Better quality data will need to be obtained before any convincing conclusion can be accepted.

An alternative to co-crystallization would be soaking the ccNiR crystals in 100 mmol/L of Na-nitrite for ~1 hour before data collection, as described elsewhere (Einsle et al. 2002). This method was attempted without success. The crystals fractured immediately upon immersion into the solution. Various Na-nitrite soak concentrations were prepared ranging from 25 mmol/L –200 mmol/L with identical results. Co-crystallization should continue to be pursued, considering these initial results, however a suitable cryobuffer will need to be developed to positively impact the data quality.

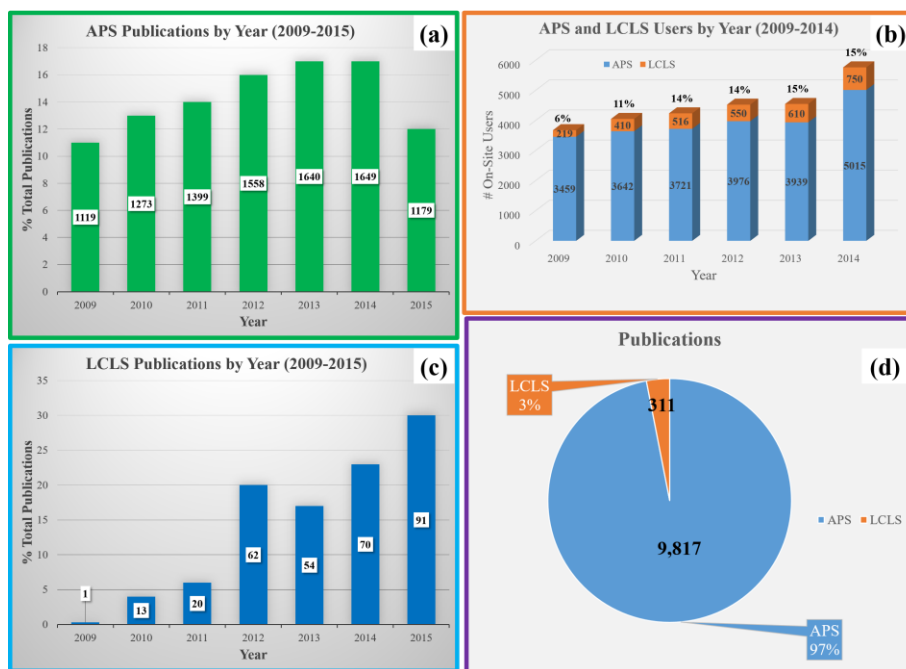


## 4 Discussion

### 4.1 LCLS and APS Scientific Output

Contemplating the potential that XFELs promise, it should be considered whether this new method for macromolecular structure determination genuinely represents a significant leap forward. After all, hard X-ray FELs (Arthur et al. 1998; Wu and Emma 2006) have only been realized for just over half a decade and techniques are presently under intense development. Comparison with the veteran synchrotrons may be tricky, considering that 3<sup>rd</sup> generation synchrotron X-ray sources now boast over 20 years of knowledge and experience.

To show the importance of these existing light sources, output in peer-reviewed journals may



**Figure 4.1.** Comparison of scientific output of the LCLS and APS. (a) APS publications by year. Total publications are shown in the white boxes. (b) Total number of on-site users for the APS (blue) and LCLS (orange). The number of on-site users are shown for each year and the ratio of LCLS Users / APS Users is given as percentages. (c) LCLS publications by year. Total publications are shown in the white boxes. (d) Breakdown of total publications between the LCLS and APS from 2009-2015.

be compared. From 2009-2015 there were 10,128 total publications combined between the APS and the LCLS (2015a; 2015b). Of these publications, 311 (~3%) were from experiments conducted at the LCLS while 9,817

(~97%) were from studies performed at the APS (see Fig. 4.1). This discrepancy appears significant, however the difference between the numbers of experimental stations at the LCLS compared to the APS is also quite large (6 versus 71, respectively). Taking this factor into account an extrapolation reveals that if the number of experimental stations between the two were similar then the annual journal publication outputs would be comparable. The overall productivity per station is therefore roughly equal. This is truly a remarkable achievement for the LCLS, however the scientific impact of X-ray free electron lasers should extend beyond these types of comparisons.

Both XFELs and synchrotrons are capable of recording high quality diffraction images from which 3-dimensional molecular structures at atomic spatial resolution may be obtained. However, an XFEL delivers superior time resolution (40 fs versus 100 ps) and their incredible brilliance and peak power offers the ability to reduce crystal sizes down to the microcrystal regime and even smaller. Consequently, the brief XFEL pulses outrun radiation damage (Neutze et al. 2000) and no freezing of the sample is necessary. Therefore, room temperature protein dynamics on the fs time scale may be probed using an XFEL.

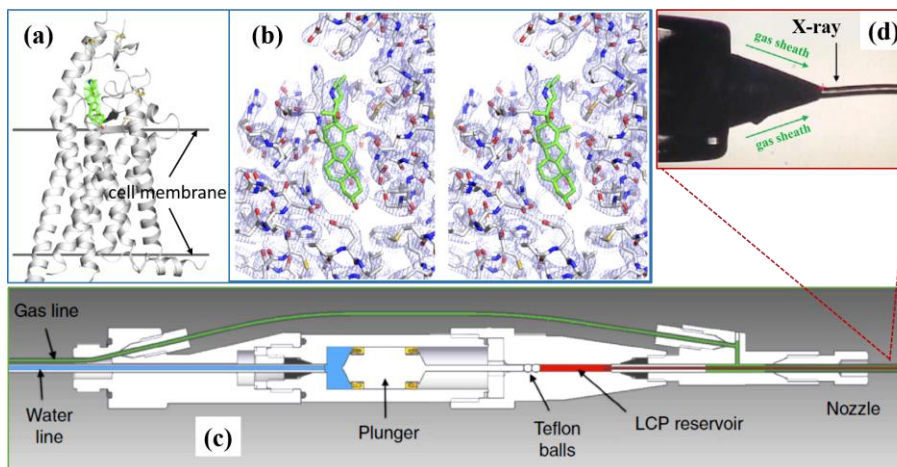
#### **4.2 Reduction of Protein Consumption and Expansion of Structural Biology with SFX**

The small crystal sizes and ultrafast time resolution that are supported by XFELs represents the hallmark departure from synchrotrons. Regrettably, the amount of purified protein required to record a sufficient number of pictures from these microcrystals to obtain structures for only a handful of time-delays is, at the moment, quite sizable. For example, the two TR-SFX studies on PYP required between 3-6 g of purified protein each (Tenboer et al. 2014, Pande et al. 2016). Therefore, the large protein volumes demanded by an XFEL experiment presents a significant challenge for research groups to produce enough sample to last through five 12 hour shifts.

Techniques for reducing protein consumption must be established to maximize the range of proteins that may be studied.

A straightforward approach to reduce protein quantities for SFX would be to dramatically improve the XFEL repetition rate so that images may be collected more quickly. The European XFEL will run at an effective X-ray repetition rate of 27,000 Hz when the facility becomes operational in 2017 (Abela et al. 2007). This faster data collection rate would lead to a >200-fold decrease in the amount of protein demanded by SFX. Another solution to this problem would be to pulse injector flows in synchronization with the arrival of XFEL pulses. This method is known as *drop-on-demand* (DOD) and uses a piezo transducer to eject small droplets of the sample into the X-rays (Weierstall 2014). Unfortunately, drop sizes are currently rather large (30  $\mu\text{m}$ ) and nozzle clogging is problematic. An alternative technique would be to slow down the sample flow rate so that the X-rays probe a greater percentage of the microcrystal mixture. In order to maintain stable streams GDVNs can only tolerate flow rates greater than 1  $\mu\text{L}/\text{min}$ , although so-called *lipidic cubic phase injectors* (LCP) (Fig. 4.2c and Fig. 4.2d) can produce sample flows in the pL/min range (Weierstall et al. 2014). In fact, LCP injectors are capable of delivering microcrystals at rates slower than the 120 Hz LCLS repetition rate. This is not an issue, however, since by increasing the driving pressure the velocity of the LCP (and microcrystals) can be tailored to the arrival time of the X-rays.

LCP injectors are ideally suited for the study of membrane proteins, which represent ~30% of the proteome and constitute ~60% of the drug targets in humans (Liu et al. 2013; Weierstall et al. 2014). One important example of a membrane protein is the human smoothened *G protein-coupled receptor* (GPCR) (Fig. 4.2a and Fig. 4.2b), which is involved in embryonic development. Although LCP injectors have been recently implemented with some success at synchrotrons (Nogly et al.



**Figure 4.2.** The Lipidic cubic phase (LCP) injector and a TR-SFX study of a GPCR. (a) The human smoothed GPCR is shown by the grey model with a green  $1\sigma$  electron density map and model of the cycloamine complex overlaid on top. (b) Stereo representation of the binding pocket with green-stick model of the cycloamine complex. (c) Schematic of the LCP injector. (d) Detail of LCP media extruding from the injector during an experiment. A helium gas sheath stabilizes the LCP flow and the approximate location of the X-rays is shown. Panels a, b and c obtained from Weierstall et al. 2014 and panel d is from Nogly et al. 2014.

2015), the small crystal sizes afforded by XFELs present a clear advantage since large crystals of membrane proteins are difficult to produce. Membrane protein crystals are grown directly in the LCP medium and the resulting

microcrystals are small but well-ordered (Liu et al. 2014). Although LCP injectors are perfectly suited for SFX studies of membrane proteins, they have also recently been proven capable of studying soluble proteins such as lysozyme and phycocyanin (Fromme et al. 2015). Therefore, if sample consumption is a barrier for the study of a particular protein with SFX, the microcrystals should be tested with the LCP medium (or perhaps with other viscous mediums) to assess whether using a microextrusion injector is possible.

Although increased rates of data collection, DOD injectors and LCP injectors lead to a decrease in the amount of sample required, improving data reduction and analysis techniques of SFX data also positively impacts protein consumption. The goal must be to lower the number of diffraction images required for structure determination. For the ns PYP experiment we recorded 253,538 indexed snapshots distributed among the dark (170,911 images) and the 1  $\mu$ s and 10 ns light

(82,627 images) data. Previous studies had suggested that anywhere from 1-30% of the total diffraction snapshots would be indexable hits. Additionally, for those experiments that had successfully determined a protein structure ~50k-100k indexed patterns were used. Therefore, we estimated that 5% of the collected images would ultimately be used for PYP structure determination and so 1 million to 2 million diffraction images were accumulated for each desired structure (e.g., dark, 10 ns and 1 $\mu$ s). Comparing these values with the 40-100 snapshots typically needed for a TR-LX study on PYP, approximately 20,000 times more XFEL than synchrotron data were obtained to produce similar resolution structures.

Most experiments to date at the LCLS have used *Crystfel* (White et al. 2012; White et al. 2013) for data reduction. This software uses the *Monte Carlo* method to integrate the Bragg spots from thousands of randomly orientated crystals. A large number of diffraction images are needed for this averaging to accurately determine the integrated reflection intensities. Recent experiments (Ginn et al. 2015) have demonstrated that alternative methods (Sauter et al. 2013; Kabsch 2014; White 2014) for data reduction can produce a high resolution (1.75 Å) structure from only 5,787 snapshots. However, it should be noted that *Crystfel* was able to determine the 1.6 Å PYP structure at the 10 ns time-delay from only 23k images. Therefore, the ns PYP study suggests that 20k-50k indexed images may be adequate for high-quality structure determination using the *Monte Carlo* method.

Regardless of the method used to integrate the reflection intensities, without quantitative measures that can provide near real-time feedback on the number of crystal hits obtained from the data, teams will certainly continue to collect more images than may be necessary. The hitfinding algorithm from *Cheetah* is accurate and quick but is not fast enough to provide simultaneous response to the 120 Hz data collection rate of the LCLS. Other software development has produced

alternative programs such as *cctbx.spotfinder* (Sauter et al. 2013) and *Hummingbird* (Single Particle Initiative Team 2015) which can also analyze snapshots for hits at fast rates. This software could help guide the decision on when enough snapshots have been collected for accurate structure determination. Near real-time feedback is critical, not only for minimizing protein consumption but also for maximizing the number of time-delays recorded in a time-series during a TR-SFX experiment.

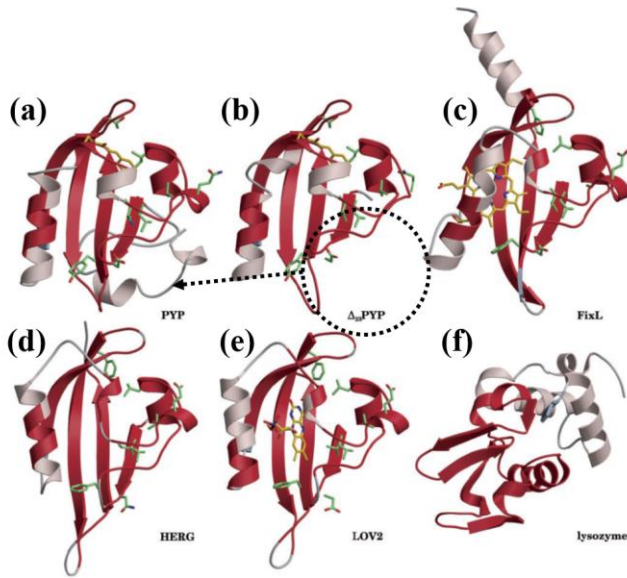
As protein volumes required for a SFX experiment continue to decrease, the number of proteins that become feasible to study using these machines will correspondingly increase. Although the ability to determine both the structure and the kinetics from the same data has been successfully demonstrated at synchrotrons (Schmidt et al. 2013), XFELs offer superior capabilities. This conclusion places X-ray free electron lasers in a unique position to remove barriers for the determination of structure and dynamics for a whole host of new proteins.

#### **4.3 Photoactive Proteins: *cis-trans* Isomerizations and E-Z Isomerizations**

Photoactive proteins represent an important group of macromolecules that play a critical role in biological success. Light-induced signaling proteins underlie a variety of reactions from vision in animals to the phototaxis response observed in other organisms, such as phototropism in plants. Many photoactive proteins are membrane proteins, such as rhodopsin which comprises ~90% of all GPCRs and is responsible for the signaling that enables vision (Palczewski et al. 2000). PYP is only one of a diverse collection of light-reactive proteins that include proteorhodopsins (Béjà et al. 2001), *blue light sensing using FAD* (BLUF) (Gomelsky and Klug 2002) and phytochromes (Wagner et al. 2005; Kennis and Groot 2007).

PYP also shares a common structural motif with the PAS family of proteins, which accepts and transduces a variety of signals. Indeed, all of the PAS domains bear a resemblance to the

conserved folds of the PYP structure (Pellequer et al. 1998; Vreede et al. 2003). PAS is an acronym derived from the names of the proteins where this repeat motif was first recognized: the *Drosophila*



**Figure 4.3.** Comparison of PAS domains. Ribbon representations are shown for (a) PYP, (b)  $\Delta_{25}$ PYP mutant, (c) FixL, (d) HERG, (e) LOV<sub>2</sub> and (f) turkey lysozyme. The residues highlighted in red represent the common motif PAS domain. Green stick models denote homologous residues at similar positions. The  $\Delta_{25}$ PYP mutant has its N-terminal helix-turn-helix cap removed, as highlighted by the dashed circle and arrow. Fig. 4.3 modified from Vreede et al. 2003.

period clock protein (**PER**), the aryl hydrocarbon receptor nuclear translocator (**ARNT**) in vertebrates and the *Drosophila* single-minded protein (**SIM**) found in insects (Taylor and Zhulin 1999). The PAS family includes proteins capable of sensing oxygen, redox potential, energy levels and light intensity (Vreede et al.

2003). The time-resolved fs experiment on PYP at the LCLS represents the culmination of over 30 years of intense study of this protein. These results may reveal insights for the role that the

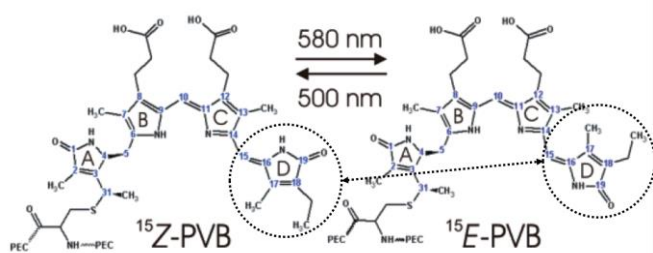
conserved folds play in the dynamics of other PAS domains.

Many proteins in the PAS family share a resemblance to PYP, such as FixL (Fig. 4.3c) (Gong et al. 1998), HERG (Fig. 4.3d) (Cabral et al. 1998), LOV<sub>2</sub> (Fig. 4.3e) (Crosson and Moffat 2001) and the turkey lysozyme (Fig. 4.3f) (PDB code 135L (Harata 1993)). However, PYP is interesting since the N-terminal helix-turn-helix feature is not present in these other proteins, as shown in Fig. 4.3. A mutant version of PYP,  $\Delta_{25}$ PYP (Fig. 4.3b), has been investigated to assess the role that this unique N-terminal cap plays on the kinetics of the PYP photocycle (Vreede et al. 2003). This study suggests that the kinetics are only moderately affected, primarily in a slower recovery of the ground



state. This conclusion is exciting because little is known about the differences in the transduction mechanism between PAS receptors like PYP and transducers such as oxygen sensors like FixL (Taylor and Zhulin 1999). Since PAS domains share a common fold there is the possibility that they also share similar dynamics, which enhances the significance of our studies (Vreede et al. 2003).

Many signaling proteins exhibit a *cis-trans* isomerization. PYP is a model system to investigate a general mechanism for how these isomerizations are initiated. Characterization of the chromophore dynamics, the excited state potential energy surfaces and the effects on the surrounding chromophore environment may reveal commonalities that other reactions share with PYP (Vengris et al. 2004). In fact, the *trans* to *cis* isomerization of the PYP chromophore is very similar to that found in retinal visual processes (Gromov et al. 2012). Our experiments on PYP may elucidate whether a *cis-trans* isomerization is a pre-requisite for the sensory transduction underlying vision in rhodopsin (Cordfunke et al. 1998). Additionally, the hydrogen bond network of the pCA chromophore with Tyr42, CYS69 and Glu46 residues (Fig. 3.2 and Fig. 3.4) is suspected to play an integral role in establishing the barriers of activation and influencing both the



**Figure 4.4.** *E-Z* isomerization of the phycoviolobilin (PVB) chromophore of  $\alpha$ -PEC. The *E*-PVB state is formed by absorption of a photon of wavelength 580 nm. The *Z*-PVB state is formed by absorption of a photon of wavelength 500 nm. The isomerization about the C15=C16 double bond is noted by the dashed black circles. Fig. 4.4 modified from Schmidt et al. 2007.

ground and excited state structures.

Understanding the relationship of the chromophore with its local protein environment may provide the ability to inhibit or promote photoisomerization in biological processes.

Other light sensing proteins, such as Phytochromes and phycobiliproteins,



have linear tetrapyrroles (also known as bile chromophores) which isomerize reversibly between so-called E-state and Z-state configurations. The protein  $\alpha$ -PEC contains a single covalently bound phycoviolobilin (PVB) chromophore that isomerizes about the C15=C16 double bond (see Fig. 4.4) (Schmidt et al. 2007). This isomerization switches reversibly between the E-state and Z-state forms by absorbing a photon of wavelength 580 nm and 500 nm, respectively.  $\alpha$ -PEC would make an excellent model for the general study of the photoreactions in phycobiliproteins and phytochromes since their chromophores are identical (Foerstendorf et al. 1997). The pigment phycoerythrin is found in red algae, including those algae responsible for toxic micro-algal blooms. Since these algae depend on this pigment for absorption of sunlight, these harmful blooms could be controlled by inhibiting their chromophore activity. The contributions from 3-dimensional time-resolved structural studies of these *cis-trans* isomerizations will positively impact our knowledge of these important biological reactions.

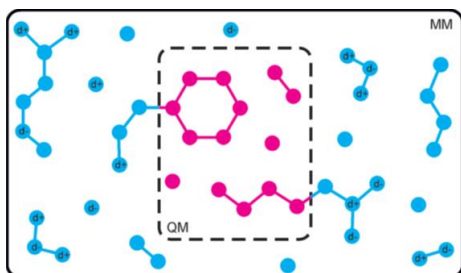
Before the development of XFELs, time-resolved structure determination of photoactive proteins was performed at synchrotrons. The fastest time scale achievable was  $\sim 100$  ps, limited by the pulse width of the X-rays. The extensive synchrotron studies on PYP captured all reaction intermediates except for those of the *trans-cis* isomerization. These experiments used large single crystals to amplify the reflection intensities so that they could be recorded. In TR-LX the crystal is typically bigger than the laser pulse and the laser does not penetrate fully into the crystal since the crystals are so large. For a ns laser at 450 nm wavelength the penetration depth for PYP is  $\sim 3$   $\mu\text{m}$  (Schmidt et al. 2013). Therefore, the high optical density of PYP along with the large crystal sizes makes uniform laser excitation difficult for TR-LX. The effect is that the X-rays probe regions of the crystal where molecules have not had the opportunity to absorb a photon and enter the photocycle. Consequently, to obtain diffraction snapshots with the maximum difference signal

3-10 pump-probe sequences must be used with a few seconds pause in between to allow the molecules to return to the dark state before the image is read from the detector (note that this means if the reaction is irreversible it may not be studied). The use of this many laser pulses comes at a cost, however, as the crystal can only tolerate a limited number before the strain becomes too great and permanent damage occurs (Schmidt et al. 2012). Additionally, the laser energy may not be increased without limit in an effort to compensate since this would cause intense stress on the crystal which can lead to streaking of the Laue diffraction spots. These limitations set a practical restriction on the extent of reaction initiation. Reaction initiations of 10%-15% (Schmidt et al. 2013) are typical for PYP experiments at synchrotrons, however, we achieved a 40% reaction initiation of the molecules with TR-SFX (Tab. 3.1). These larger levels of photo-initiation lead to stronger difference signals which may ultimately produce more accurate structure determination, as we have demonstrated by the observed new configuration of the Glu46 residue.

The serial nature of a TR-SFX experiment circumvents the limitations on reaction initiation partly because each image is acquired from a new crystal. Laser fluence may therefore be increased for serial crystallography to levels not permissible with TR-LX because the crystal only receives a single exposure (Schmidt et al. 2012). Additionally, the small sizes of the microcrystals facilitates a more uniform laser excitation with complete penetration and little local strain. The diffraction images obtained with the quasi-monochromatic XFEL for the PYP data show none of the streaking common with TR-LX. The reason for this difference may be understood when considering the *Ewald Sphere* construction described in §2.1.4 and the *mosaicity* discussed in §2.1.5. The near monochromatic nature of the X-rays produced by FELs leads to the recording of only partial reflections. Therefore, any transient strain in the crystal will lead to an increase in the *mosaicity* and subsequently the partiality of the reflection intensities observed from a reciprocal lattice point.

This increased partiality simply requires more diffraction snapshots in order to sufficiently sample reciprocal space.

Although synchrotrons are prohibited from probing fast reactions (photo-initiated or



**Figure 4.5.** The concept of the QM/MM method involves selecting the reaction region from the system to be treated quantum mechanically. The remainder of the protein is modeled by molecular-mechanics. Fig. 4.5 from Groenhof 2013.

otherwise) at timescales  $<100$  ps, other methods are available. PYP has been studied on the fs timescale by time-resolved spectroscopic techniques such as transient absorption spectroscopy (Devanathan et al. 1999), Raman spectroscopy (Creelman et al. 2014), and fluorescence spectroscopy (Nakamura et al. 2007). While these methods reveal only limited structural information (Schmidt et al. 2013) they provide insights on protein

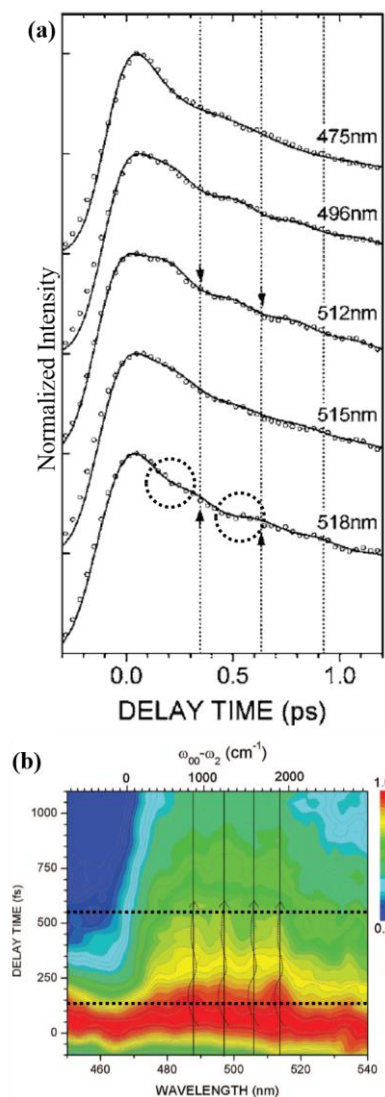
dynamics that can guide and assist fs structural studies at X-ray FELs. QM/MM simulations also offer a powerful technique for interpretation of fs structural results. In the QM/MM approach, a small region of the protein where the reaction occurs is selected for a quantum mechanical analysis while the remainder of the system is treated with molecular mechanics (Fig. 4.5) (Groenhof 2013a). The conclusions of these complex simulations will certainly be a critical component for the interpretation and analysis of fs TR-SFX data because excited state dynamics is suspected to play an integral role in driving protein configurational changes. Recent results, including the fs PYP study, have used QM/MM methods to support the structure evaluation of these ultrafast reaction intermediates (Barends et al. 2015, Pande et al. 2016).

Our fs TR-SFX study has revealed structurally uncharacterized states of the PYP photocycle and elucidates the structural progression of the early stages of the *trans-cis* isomerization of the pCA chromophore. After absorbing a photon of energy  $h\nu=4.4 \times 10^{-19}$  J = 2.76 eV, the

chromophore is elevated to an excited potential energy surface. The excited state chromophore motions for time-delays <500 fs can be explained if ~0.21 eV of the photon energy is converted into kinetic energy, which is ~10% of the total energy. The remaining 80-90% of the photon energy remains interred in the strain of the chromophore, as suspected by previous studies using stimulated Raman spectroscopy (Creelman et al. 2014). This stored energy may be used to drive the structural changes of the protein to the *cis* configuration as the molecules relax towards the ground state electronic surface (see Fig. 3.4c).

For time-delays <500 fs the chromophore remains in the *trans* configuration, although the tail of the chromophore becomes strained in a twisted *trans* configuration with the C<sub>1</sub>-C<sub>2</sub>=C<sub>3</sub>-C<sub>1</sub> torsional angle  $\varphi_{\text{tor}} = \sim 140^\circ$  (see Fig. 3.4b). QM/MM simulations have suggested a larger torsional angle closer to  $\varphi_{\text{tor}} = \sim 180^\circ$  (Groenhof et al. 2004), however this is not observed in these data. The C<sub>2</sub>=C<sub>3</sub> double bond is shifted by 1 Å behind the chromophore plane defined by the span of the Cys69 sulfur, the head atoms and the carboxyl oxygen in the tail (Fig. 3.4d). A transition with a time of ~150 fs is observed by stimulated Raman spectroscopy (Creelman et al. 2014). This characteristic transition time is also witnessed in fluorescence spectroscopy (left dashed circle in Fig. 4.6a and bottom dashed horizontal line in Fig.4.6b). Numerical calculations performed by Nakamura on time-resolved fluorescence spectra further support a transition occurring during this time. Some 400 fs later, at the conical intersection (Fig. 3.4c) of the excited state and ground state potential energy surfaces the molecules now either relax to the dark state structure or proceed to the *cis* configuration.

The *trans* to *cis* isomerization develops within 180 fs (solid blue line, Fig. 3.4a). To detect this fast transition on these slower time-scales implies that the ensemble relaxes synchronously (Pande et al. 2016). This synchronization cannot persist and the molecules eventually dephase. As the



**Figure 4.6.** Fluorescence spectroscopy of PYP. (a) Normalized fluorescence decays of PYP at various wavelengths (475 nm, 496 nm, 512 nm, 515 nm and 518 nm). The solid black lines are fitting curves. The dashed black circles indicate the location of the two characteristic times determined by our SVD analysis of the TR-SFX PYP data. (b) Contour plot of the normalized time-resolved fluorescence spectra. The dashed black horizontal lines again denote the characteristic times. Fig. 4.6a and Fig. 4.6b modified from Nakamura et al. 2007.

*trans* to *cis* isomerization progresses, a large rotation of the C2=C3 double bond occurs at  $\sim 550$  fs, in agreement with the rSV<sub>2</sub> results from the SVD analysis. Creelman additionally notes that the stimulated Raman spectra for times  $< 500$  fs were growing, while those for times  $> 500$  fs began to decay (Creelman et al. 2014). The pCA chromophore head has broken the hydrogen bond with Glu46 by pivoting so that the C2=C3 double bond aligns along the tail axis. Although we observe a broken hydrogen bond with Glu46 Creelman remarks that this bond is only weakened. However, this result is perhaps due to differences in the local chromophore environment between PYP in solution and in crystal form.

The intermediate then relaxes further to the ground state potential energy surface beginning at 1 ps as the isomerization proceeds and the torsional angle decreases to  $\varphi_{\text{tor}} = \sim 38^\circ$ . The hydrogen bond of the chromophore head is reestablished with Glu46 by 3 ps and the resulting structure shown in Fig. 3.4f is strikingly similar to the known I<sub>T</sub> (or pR<sub>0</sub>) structure determined previously by synchrotron experiments at a time-delay of 100 ps (Schotte et al. 2012; Jung et al. 2013). These results demonstrate that time-resolved near atomic resolution structure

determination is possible at ultrafast time-scales using TR-SFX. The 1.6 Å XFEL structure of PYP has been further refined to a resolution better than 1.5 Å (Schmidt et al. 2015), which to my knowledge represents the best spatial resolution to date of a biological macromolecule recorded at the LCLS. Furthermore, comparison of the XFEL structures to those obtained by synchrotron reveal that the PYP XFEL data shows practically no indication of radiation damage even though they were acquired at room temperature. The structures determined by the LCLS are comparable to high resolution (atomic) structures from the APS.

#### **4.4 Enzymes: Structure and Applications**

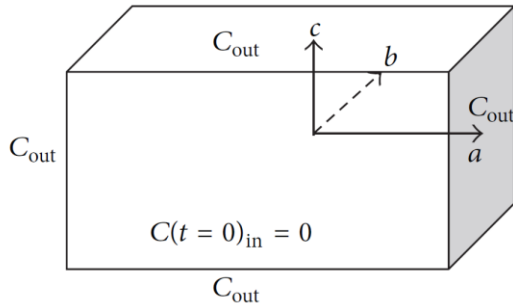
Enzymes are proteins that catalyze the reactions vital for life (Carvalho et al. 2014). Enzymes perform their function by lowering the amount of energy needed to trigger the reaction. Certainly, without the presence of these specialized proteins to catalyze reactions most would take prohibitively long to support life. Enzymes are capable of speeding up many reactions by more than a factor of a million (Berg et al. 2002). Enzyme function, and in particular their *specificity*, has long been suspected of being intimately linked with their structure (Fischer 1894). Determining the 3-dimensional structures of the reaction intermediates of these proteins in action using the *mix-and-inject* method (Schmidt 2013) would revolutionize our understanding of the mechanisms of these catalytic processes.

A *mix-and-inject* study is fundamentally a diffusion experiment. Instead of photo-activation by a laser the microcrystals are mixed with a substrate and the subsequent structural changes are

probed in a time-resolved manner. Fick's second law describes the time-dependent diffusion of a substrate in solution,

$$D\nabla^2 c = \frac{\partial c}{\partial t} \quad (4.1)$$

where  $D$  is the diffusion coefficient,  $\nabla^2$  is the Laplace operator and  $c$  is the time-dependent concentration of the diffusing substrate. While this second order partial differential equation can



**Figure 4.7.** A rectangular crystal of dimensions  $2a$ ,  $2b$  and  $2c$ . The substrate concentration outside of the crystal is  $C_{out}$  and is assumed constant. The substrate concentration inside the crystal is initially zero and changes with time,  $C_{in}(t)$ . Fig. 4.7 from Schmidt 2013.

be solved by separation of variables with appropriate boundary conditions, only the simplest cases admit an analytic solution. For other situations numerical methods must be employed.

The simplest case is now considered, with a rectangular crystal of dimensions  $2a$ ,  $2b$  and  $2c$  (see Fig. 4.7). Certain assumptions are made: (i) the substrate concentration outside of the crystal is constant; (ii) the substrate freely diffuses through the crystal; (iii) the mixing of the substrate with the crystal occurs instantaneously. The solution for Eqn. 4.1 subject to these boundary conditions is available from literature (Schmidt 2013). Importantly, all of the time dependence for the concentration inside the crystal is contained in an exponential term. We may define the characteristic time,  $\tau_D$ , as

$$\tau_D = \frac{4}{D\pi^2[(2l+1)^2/a^2 + (2m+1)^2/b^2 + (2n+1)^2/c^2]} \quad (4.2)$$

This characteristic time describes when the concentration of the substrate inside the crystal reaches a value of  $C_{out}(1 - 1/e) = 0.63C_{out}$ . The first order term ( $l, m, n = 0$ ) gives the slow limit

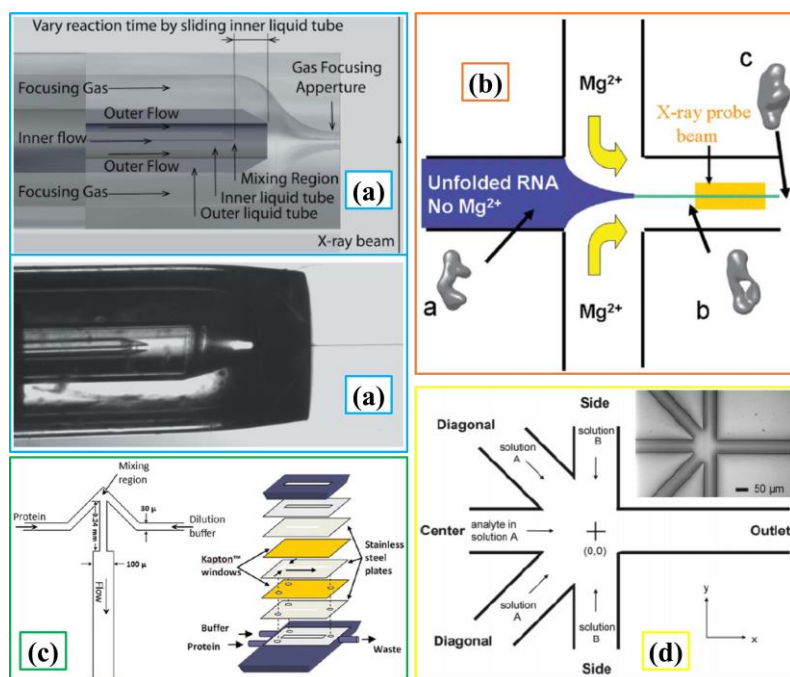


**Table 4.1.** Diffusion times,  $\tau_D$ , for various crystal sizes.

Crystal Size [ $\mu\text{m}^3$ ]	$\tau_D$
400 x 400 x 1600	16 sec < $\tau_D$ < 60 sec
300 x 400 x 500	9.5 sec
10 x 20 x 30	15 ms
3 x 4 x 5	1 ms
<sup>a</sup> 1 x 2 x 3	150 $\mu\text{s}$
0.5 x 0.5 x 0.5	17 $\mu\text{s}$
<sup>b</sup> 0.1 x 0.2 x 0.3	1.5 $\mu\text{s}$

<sup>a</sup>For smaller crystal sizes the mixing time could be slower than the diffusion time.

<sup>b</sup>*Shewanella oneidensis* ccNiR crystals would have ~4000 unit cells.



**Figure 4.8.** Mixing injectors. (a) A double-focusing mixing jet modeled after GDVNs. The inner capillary delivers the microcrystal mixture surrounded by an outer substrate-carrying capillary surrounded by a third capillary delivering the focusing gas. Time resolution is accomplished by varying the position of the inner capillary. (Wang et al. 2014) (b) A cross-shaped mixer in which the substrate serves a double purpose to activate and focus the sample stream (Pollack 2011). (c) A y-shaped mixer similar to (b) (Graceffa et al. 2013). (d) A 5-channel mixer that provides the ability to deliver multiple substrates (Park et al. 2006).

diffusion time,  $\tau_D$ , and all other higher order terms are faster. Diffusion times for various crystal sizes are shown in Tab. 4.1.

A *mix-and-inject* experiment is perfectly suited for an XFEL because of the serial nature of a TR-SFX study, the fast substrate diffusion times stemming from the small sizes of the microcrystals

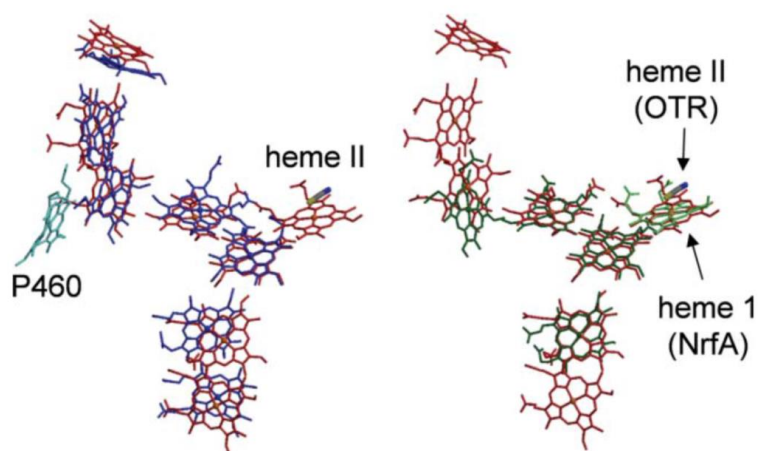
and the capability to study proteins at ambient temperatures. Additionally, this method is easily adaptable to any pH level and substrate concentration. Nearly all mixers designed for enzymatic studies share common components (Fig.4.8): samples and substrate are initially stored in separate reservoirs; they are brought together in a mixing chamber designed to quickly and thoroughly combine sample with substrate; time-resolution is



accomplished by allowing the sample to travel a distance  $\Delta L$  before the reaction is probed. The possibility to determine the 3-dimensional structures of enzymes in action is exciting to consider.

The benefits of *mix-and-inject* are numerous when compared to established methods for determining chemical kinetics, such as *stopped-flow UV/VIS spectroscopy* (SFS-UV/VIS) (Gibson 1954), *stopped-flow NMR spectroscopy* (SFS-NMR) (Grimaldi et al. 1972) and *rapid-freeze-quench EPR* (RFQ-EPR) (Ballou and Palmer 1974). As mentioned previously, spectroscopic methods provide only limited structural information.

Similar to the PAS domains discussed for PYP, ccNiR shares a conserved structural motif of heme arrangements with other enzymes, such as hydroxylamine oxidoreductase (HAO) and



**Figure 4.9.** A comparison of the heme arrangement of HAO (left side, dark blue), OTR (both sides, red) and ccNiR (right side green), also known as NrfA by some researchers. Fig. 4.9 from Mowat and Chapman 2005.

tetrathionate reductase (Fig. 4.9) (Bewley et al. 2013; Stein et al. 2015). The heme structure agreement for ccNiR and OTR even extends to the active site (right side, Fig. 4.9), although these proteins share no obvious global structural resemblances (Mowat and

Chapman 2005). HAO catalyzes the oxidation of  $\text{NH}_2\text{OH}$  to  $\text{NO}_2^-$ , as shown in lower right quadrant of the nitrogen cycle in Fig. 1a (Kostera et al. 2010). This process is essentially the reverse direction of the reaction that ccNiR catalyzes.

Since enzymes are highly selective and sensitive to their respective reactants, these proteins could be used in amperometric electrodes for the detection or measurement of unwanted chemical

species (Strehlitz et al. 1996). For example, the ability to accurately quantify nitrite concentrations has a direct application for the monitoring of nitrite in drinking water and waste water treatment facilities. Such devices would require an electron donor to function. Therefore, understanding the electron pathways of these reduction processes could also catalyze the design of novel artificial electron donors.

CcNiR is particularly attractive for these types of applications because it uses 6 electrons in its reduction of nitrite to ammonia. The more electrons that are involved in any catalytic reaction the larger the reduction current. Although this reduction current could be used for sensors, it may also be exploited as an energy source. These biological power plants would clearly have a wide variety applications. Since all that is needed is an enzyme and a suitable electron donor the possibilities are exciting to consider: bio-batteries; waste treatment facilities doubling as electric plants; reduction of nitrate and nitrite in food crops while simultaneously producing electricity. To use enzymes for these purposes will require a detailed understanding of these reduction processes. CcNiR is an excellent model for study, considering its similarities with other heme proteins. *Mix-and-inject* experiments on enzymes at FELs is the right tool for determining the 3-dimensional structure and dynamics of these reactions.

#### **4.5 Conclusion and Outlook**

The experiments performed at the LCLS over the last 5 years demonstrate the unique capabilities an XFEL provides to propel time-resolved crystallography (and other fields) into uncharted regions. By facilitating the study of both photo-active and (eventually) enzymatic

reactions of macromolecules, XFELs are poised to deliver amazing discoveries in structural biology.

However, synchrotrons are invaluable tools that should be capitalized on into the foreseeable future. Although techniques such as sample injection and microcrystal creation were born out of necessity at an XFEL, they provide corridors for research and development at synchrotrons. While not currently able to probe the small microcrystals used for a TR-SFX experiment, synchrotrons have shown they are qualified to determine 3-dimensional structures of proteins whose crystals are in the  $10 \times 10 \times 40 \mu\text{m}^3$  range (Stellato et al. 2014; Gati et al. 2014; Botha et al. 2015). This new potential may expand the range of proteins that can be studied. As synchrotron X-ray focal sizes are reduced and photon fluxes improved, the crystal sizes acceptable for study will correspondingly decrease.

An argument may be made that as TR-SFX methods develop and more XFELs come online the usefulness of synchrotrons will diminish. Although this is possibly inevitable, the timeframe will certainly be on the order of years and will most likely extend over decades. In the meantime, the most effective course should be to cultivate TR-SX techniques in tandem at both synchrotrons and XFELs. As XFEL technology improves the possibility of retrofitting synchrotrons with this technology should become more feasible. Time-resolved crystallography can only benefit from this combination.

## 5 References

- Abela R, Aghababayan A, Altarelli M (2007) The European X-ray Free Electron Laser Technical Design Report. Hamburg
- Alberts B, Johnson A, Lewis J (2002) Molecular Biology of the Cell, 4th edn. New York: Garland Science
- Alexander M (1977) Introduction to soil microbiology / Martin Alexander., 2d ed.. New York : Wiley, New York
- Amorós JL, Buerger MJ, de Amorós MC (1975) The Laue Method. Academic Press, New York
- Andor (2008a) Users Guide Andor Technology iStar. Belfast, UK
- Andor (2008b) Users Guide Andor Technology Shamrock SR-303i. Belfast, UK
- Aotore E, Kraabel B, Rats D (2007) GeniX Installation and Operation Manual. Sassenage
- Arnlund D, Johansson LC, Wickstrand C, et al (2014) Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser. *Nat Meth* 11:923–926.
- Arthur J, Bane K, Bharadwaj V, et al (1998) Linac Coherent Light Source (LCLS) Design Study Report.
- Asahi Spectra (2010) Instruction Manual Xenon Light Source 300W MAX-302. Torrance, CA
- Bailey S (1994) THE CCP4 SUITE - PROGRAMS FOR PROTEIN CRYSTALLOGRAPHY. *Acta Crystallogr Sect D-Biological Crystallogr* 50:760–763.
- Ballou DP, Palmer GA (1974) Practical rapid quenching instrument for the study of reaction mechanisms by electron paramagnetic resonance spectroscopy. *Anal Chem* 46:1248–1253. doi: 10.1021/ac60345a034
- Ban N, Nissen P, Hansen J, et al (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* (80- ) 289:905–920.
- Barends TRM, Foucar L, Ardevol A, et al (2015) Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* (80- ). doi: 10.1126/science.aac5492
- Barty A, Kirian R, Maia F, et al (2014) Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J Appl Crystallogr* 47:1118–1131. doi: 10.1107/S1600576714007626
- Béjà O, Spudich EN, Spudich JL, et al (2001) Proteorhodopsin phototrophy in the ocean. *Nature*

411:786. doi: 10.1038/35081051

- Berg JM, Tymoczko JL, Stryer L (2002) *Biochemistry*. 5th edition. New York: W.H. Freeman
- Bergfors TM (ed) (1999) *Protein crystallization : techniques, strategies, and tips : a laboratory manual / edited by Terese M. Bergfors*. La Jolla, Calif. : International University Line, c1999., La Jolla, Calif.
- Berman HM, Battistuz T, Bhat TN, et al (2002) The Protein Data Bank. *Acta Crystallogr Sect D* 58:899–907. doi: 10.1107/S0907444902003451
- Bewley KD, Ellis KE, Firer-Sherwood MA, Elliott SJ (2013) Multi-heme proteins: Nature's electronic multi-purpose tool. *BBA - Bioenerg* 1827:938–948. doi: 10.1016/j.bbabi.2013.03.010
- Bionta M, Hartmann N, Weaver M, et al (2014) Spectral encoding method for measuring the relative arrival time between x-ray/optical pulses. *Rev Sci Instrum*. doi: 10.1063/1.4893657
- Blake CCF, KOENIG DF, MAIR GA, et al (1965) Structure of Hen Egg-White Lysozyme: A Three-dimensional Fourier Synthesis at 2 [angst] Resolution. *Nature* 206:757–761.
- Blundell TL (1996) Structure-based drug design. *Nature* 384:23.
- Borgstahl GE, Williams DR, Getzoff ED (1995) 1.4 Å structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site, and chromophore. *Biochemistry* 34:6278.
- Botha S, Nass K, Barends TRM, et al (2015) Room-temperature serial crystallography at synchrotron X-ray sources using slowly flowing free-standing high-viscosity microstreams. *Acta Crystallogr Sect D Biol Crystallogr* 71:387–397. doi: 10.1107/S1399004714026327
- Boutet S, Lomb L, Williams GJ, et al (2012) High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* (80- ) 337:362–364. doi: 10.1126/science.1217737
- Boutet S, Williams GJ (2010) The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). *New J Phys* 12:35024.
- Bragg WH, Bragg WL (1949) *The crystalline state. Vol. 1 : A general survey*. Bell : London
- Bragg WL (1913) The Structure of Some Crystals as Indicated by Their Diffraction of X-rays. *Proc R Soc London A Math Phys Eng Sci* 89:248–277.
- Bragg WL (1912a) The diffraction of short electromagnetic waves by a crystal. *Nature*. doi: 10.1038/090400a0
- Bragg WL (1912b) The specular reflection of X-rays. *Nature*. doi: 10.1038/090410b0
- Bragg WL (1914) The X-ray Spectrometer. *Nature* 94:199–200. doi: 10.1038/094199a0

- Bragulla HH, Homberger DG (2009) Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. *J Anat* 214:516–559.
- Brehm W, Diederichs K (2014) Breaking the indexing ambiguity in serial crystallography. *Acta Crystallogr Sect D-Biological Crystallogr* 70:101–109. doi: 10.1107/S1399004713025431
- Brunger A (1992) FREE R-VALUE - A NOVEL STATISTICAL QUANTITY FOR ASSESSING THE ACCURACY OF CRYSTAL-STRUCTURES. *Nature* 355:472–475.
- Brunger A (1993) ASSESSMENT OF PHASE ACCURACY BY CROSS VALIDATION - THE FREE R-VALUE - METHODS AND APPLICATIONS. *Acta Crystallogr Sect D-Biological Crystallogr* 49:24–36.
- Bykov D, Neese F (2011) Substrate binding and activation in the active site of cytochrome c nitrite reductase: a density functional study. *JBIC J Biol Inorg Chem* 16:417–430. doi: 10.1007/s00775-010-0739-6
- Bykov D, Neese F (2012) Reductive activation of the heme iron–nitrosyl intermediate in the reaction mechanism of cytochrome c nitrite reductase: a theoretical study. *JBIC J Biol Inorg Chem* 17:741–760. doi: 10.1007/s00775-012-0893-0
- Cabral JHM, Lee A, Cohen SL, et al (1998) Crystal Structure and Functional Analysis of the HERG Potassium Channel N Terminus: A Eukaryotic PAS Domain. *Cell* 95:649–655. doi: 10.1016/S0092-8674(00)81635-9
- Carvalho ATP, Barrozo A, Doron D, et al (2014) Challenges in computational studies of enzyme structure, function and dynamics. *J Mol Graph Model* 54:62–79. doi: 10.1016/j.jmglm.2014.09.003
- Chapman HNH, Fromme P, Barty A, et al (2011) Femtosecond X-ray protein nanocrystallography. *Nature* 470:73–U81. doi: 10.1038/nature09750
- Chayen NE, Saridakis E (2008) Protein crystallization: from purified protein to diffraction-quality crystal. *Nat Methods* 5:147–153.
- Chen Y, Šrajer V, Ng K, et al (1994) Optical monitoring of protein crystals in time-resolved x-ray experiments: Microspectrophotometer design and performance.
- Chrispeels MJ, Crawford NM, Schroeder JI (1999) Proteins for Transport of Water and Mineral Nutrients across the Membranes of Plant Cells. *Plant Cell Online* 11:661–675. doi: 10.1105/tpc.11.4.661
- Cole JA, Brown CM (1980) Nitrite reduction to ammonia by fermentative bacteria: A short circuit in the biological nitrogen cycle. *FEMS Microbiol Lett* 7:65–72. doi: [http://dx.doi.org/10.1016/S0378-1097\(80\)80001-2](http://dx.doi.org/10.1016/S0378-1097(80)80001-2)

- Cordfunke R, Kort R, Pierik A, et al (1998) Trans/cis (Z/E) photoisomerization of the chromophore of photoactive yellow protein is not a prerequisite for the initiation of the photocycle of this photoreceptor protein. *Proc Natl Acad Sci U S A* 95:7396.
- Creelman M, Kumauchi M, Hoff WD, Mathies R a. (2014) Chromophore dynamics in the PYP photocycle from femtosecond stimulated raman spectroscopy. *J Phys Chem B* 118:659–667. doi: 10.1021/jp408584v
- Crosson S, Moffat K (2001) Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction. *Proc Natl Acad Sci U S A* 98:2995–3000.
- Cruise AJ (1957) Anisotropic Structures in Human Elastin. *Nature* 179:674–675.
- DeLano WL, Schrödinger L The PyMOL Molecular Graphics System.
- DeLong KW, Fittinghoff DN, Sweetser JN, et al (1997) Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating. *Rev Sci Instrum* 68:3277. doi: 10.1063/1.1148286
- Devanathan S, Pacheco A, Ujj L, et al (1999) Femtosecond Spectroscopic Observations of Initial Intermediates in the Photocycle of the Photoactive Yellow Protein from *Ectothiorhodospira halophila*. *Biophys J* 77:1017–1023. doi: 10.1016/S0006-3495(99)76952-3
- Doak DPD and UW and KS and JW and DS and JCHS and RB (2008) Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J Phys D Appl Phys* 41:195505.
- Doak RB, DePonte DP, Nelson G, et al (2012) Microscopic linear liquid streams in vacuum: Injection of solvated biological samples into X-ray free electron lasers.
- Doyle RA, Muchmore CRA, Blum M (2006) *marccd Software Manual*. Evanston, IL
- Duerring M, Huber R, Bode W, et al (1990) REFINED 3-DIMENSIONAL STRUCTURE OF PHYCOERYTHROCYANIN FROM THE CYANOBACTERIUM MASTIGOCLADUS-LAMINOSUS AT 2.7-Å. *J Mol Biol* 211:633–644.
- Einsle O, Messerschmidt A, Huber R, et al (2002) Mechanism of the Six-Electron Reduction of Nitrite to Ammonia by Cytochrome c Nitrite Reductase. *J Am Chem Soc* 124:11737–11745. doi: 10.1021/ja0206487
- Einsle O, Messerschmidt A, Stach P, et al (1999) Structure of cytochrome c nitrite reductase. *Nature* 400:476–480. doi: <http://dx.doi.org/10.1038/22802>
- Einsle, O., Stach, P., Messerschmidt, A., Simon, J., Kroger, A., Huber, R., Kroneck PM, Stach P, Messerschmidt A, et al (2000) Cytochrome c nitrite reductase from *Wolinella succinogenes* - Structure at 1.6 angstrom resolution, inhibitor binding, and heme-packing motifs. *J Biol Chem* 275:39608–39616. doi: 10984487



- Emma P The Linac Coherent Light Source at SLAC. Menlo Park, CA
- Emma P, Akre R, Arthur J, et al (2010) First lasing and operation of an angstrom-wavelength free-electron laser. *Nat Photonics* 4:641–647. doi: 10.1038/NPHOTON.2010.176
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and Development of Coot. *Acta Crystallogr Sect D - Biol Crystallogr* 66:486.
- Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr Sect D* 62:72–82. doi: 10.1107/S0907444905036693
- Feld GK, Heymann M, Benner WH, et al (2015) Low- Z polymer sample supports for fixed-target serial femtosecond X-ray crystallography. *J Appl Crystallogr* 48:1072–1079. doi: 10.1107/S1600576715010493
- Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der Dtsch Chem Gesellschaft* 27:2985–2993. doi: 10.1002/cber.18940270364
- Foerstendorf H, Parbel A, Scheer H, Siebert F (1997) Z,E isomerization of the  $\alpha$ -84 phycoviolobin chromophore of phycoerythrocyanin from *Mastigocladus laminosus* investigated by Fourier-transform infrared difference spectroscopy. *FEBS Lett* 402:173–176. doi: 10.1016/S0014-5793(96)01521-9
- Fromme R, Ishchenko A, Metz M, Chowdhury R (2015) Serial femtosecond crystallography of soluble proteins in lipidic cubic phase. 545–551. doi: 10.1107/S2052252515013160
- Galloway JN, Townsend AR, Erisman JW, et al (2008) Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions. *Science* (80- ) 320:889–892. doi: 10.1126/science.1136674
- Gañán-Calvo AM (1998) Generation of Steady Liquid Microthreads and Micron-Sized Monodisperse Sprays in Gas Streams. *Phys Rev Lett* 80:285–288.
- Gati C, Bourenkov G, Klinge M, et al (2014) Serial crystallography on *in vivo* grown microcrystals using synchrotron radiation. *IUCrJ* 1:87–94. doi: 10.1107/S2052252513033939
- Gibson QH (1954) Stopped-flow apparatus for the study of rapid reactions. *Discuss Faraday Soc* 17:137. doi: 10.1039/df9541700137
- Ginn HM, Messerschmidt M, Ji X, et al (2015) Structure of CPV17 polyhedrin determined by the improved analysis of serial femtosecond crystallographic data. *Nat Commun* 6:6435. doi: 10.1038/ncomms7435
- Glowacki JM, Cryan J, Andreasson J, et al (2010) Time-resolved pump-probe experiments at the LCLS. *Opt Express* 18:17620. doi: 10.1364/OE.18.017620



- Gomelsky M, Klug G (2002) BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem Sci* 27:497–500. doi: 10.1016/S0968-0004(02)02181-3
- Gong W, Hao B, Mansy SS, et al (1998) Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proc Natl Acad Sci U S A* 95:15177.
- Graber T, Anderson S, Brewer H, et al (2011) BioCARS: a synchrotron resource for time-resolved X-ray science. *J Synchrotron Radiat* 18:658–670. doi: 10.1107/S0909049511009423
- Graceffa R, Nobrega RP, Barrea RA, et al (2013) Sub-millisecond time-resolved SAXS using a continuous-flow mixer and X-ray microbeam. *J Synchrotron Radiat* 20:820–825. doi: 10.1107/S0909049513021833
- Grimaldi J, Baldo J, McMurray C, Sykes BD (1972) Stopped-flow nuclear magnetic resonance spectroscopy. *J Am Chem Soc* 94:7641–7645. doi: 10.1021/ja00777a006
- Groenhof G (2013a) Solving Chemical Problems with a Mixture of Quantum-Mechanical and Molecular Mechanics Calculations: Nobel Prize in Chemistry 2013. *Angew Chemie-International Ed* 52:12489–12491. doi: 10.1002/anie.201309174
- Groenhof G (2013b) Introduction to QM/MM Simulations. In: Monticelli L, Salonen E (eds) *Biomolecular Simulations SE - 3*. Humana Press, pp 43–66
- Groenhof G, Bouxin-Cademartory M, Hess B, et al (2004) Photoactivation of the Photoactive Yellow Protein: Why Photon Absorption Triggers a Trans-to-Cis Isomerization of the Chromophore in the Protein. *J Am Chem Soc* 126:4228–4233. doi: 10.1021/ja039557f
- Gromov E V, Burghardt I, Köppel H, Cederbaum LS (2012) Native hydrogen bonding network of the photoactive yellow protein (PYP) chromophore: Impact on the electronic structure and photoinduced isomerization. *J Photochem Photobiol A Chem* 234:123–134. doi: 10.1016/j.jphotochem.2012.01.007
- Haddock BA, Jones CW (1977) Bacterial respiration. *Bacteriol Rev* 41:47–99.
- Han K editor, Han K (2014) *Protein Conformational Dynamics* edited by Ke-li Han, Xin Zhang, Ming-jun Yang. Cham : Springer International Publishing : Imprint: Springer, 2014., Dordrecht
- Harata K (1993) X-ray structure of monoclinic turkey egg lysozyme at 1.3 Å resolution. *Acta Crystallogr Sect D* 49:497–504. doi: 10.1107/S0907444993005542
- Hart P, Boutet S, Carini G, et al (2012a) The CSPAD megapixel x-ray camera at LCLS. *Proc. SPIE* 8504:85040C–85040C–11.

- Hart P, Boutet S, Carmi G, et al (2012b) The Cornell-SLAC pixel array detector at LCLS. In: Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE. pp 538–541
- Hartmann N, Helml W, Galler A, et al (2014) Sub-femtosecond precision measurement of relative X-ray arrival time for free-electron lasers. *Nat Photonics* 8:706–709. doi: 10.1038/nphoton.2014.164
- Hellingwerf KJ, Hendriks J, Gensch T (2002) On the Configurational and Conformational Changes in Photoactive Yellow Protein that Leads to Signal Generation in *Ectothiorhodospira halophila*. *J Biol Phys* 28:395–412. doi: 10.1023/A:1020360505111
- Hendrickson WA (1985) Analysis of Protein Structure from Diffraction Measurement at Multiple Wavelengths. *Am Cryst Assoc* 21:11–21.
- Hendrickson WA (1997) Phase Determination from Multiwavelength Anomalous Diffraction Measurements. *Methods Enzymol* 276:494–523.
- Henry ER, Hofrichter J (1992) Singular value decomposition: Application to analysis of experimental data. *Methods Enzymol* 210:129–192. doi: 10.1016/0076-6879(92)10010-B
- Herrmann S, Boutet S, Duda B, et al (2013) CSPAD-140k: A versatile detector for LCLS experiments. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip* 718:550–553. doi: <http://dx.doi.org/10.1016/j.nima.2013.01.057>
- Hong Q, Zhao K, Scheer H (1993) TWO DIFFERENT TYPES OF PHOTOCHEMISTRY IN PHYCOERYTHROCYANIN  $\alpha$ -SUBUNIT\*. *Photochem Photobiol* 58:745–747. doi: 10.1111/j.1751-1097.1993.tb04962.x
- Housley RM, Hess F (1966) Analysis of Debye-Waller-Factor and Mössbauer-Thermal-Shift Measurements. I. General Theory. *Phys Rev* 146:517–526. doi: 10.1103/PhysRev.146.517
- Hunter MS, Segelke B, Messerschmidt M, et al (2014) Fixed-target protein serial microcrystallography with an x-ray free electron laser.
- Ihee H, Rajagopal S, Šrajer V, et al (2005) Visualizing reaction pathways in photoactive yellow protein from nanoseconds to seconds. *Proc Natl Acad Sci U S A* 102:7145–7150. doi: 10.1073/pnas.0409035102
- Judd ET, Youngblut M, Pacheco AA, Elliott SJ (2012) Direct electrochemistry of *Shewanella oneidensis* cytochrome c nitrite reductase: evidence of interactions across the dimeric interface. *Biochemistry* 51:10175. doi: 10.1021/bi3011708
- Jung YO, Lee JH, Kim J, et al (2013) Volume-conserving trans-cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nat Chem* 5:212–220. doi: 10.1038/nchem.1565

- Kabsch W (2014) Processing of X-ray snapshots from crystals in random orientations. *Acta Crystallogr Sect D* 70:2204–2216. doi: 10.1107/S1399004714013534
- Kane D, Trebino R (1993) CHARACTERIZATION OF ARBITRARY FEMTOSECOND PULSES USING FREQUENCY-RESOLVED OPTICAL GATING. *IEEE J Quantum Electron* 29:571–579.
- Karplus M, Weaver DL (1976) Protein-folding dynamics. *Nature* 260:404–406.
- Kendrew JC (1958) Architecture of a Protein Molecule. *Nature* 182:764.
- Kendrew JC, BODO G, DINTZIS HM, et al (1958) A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* 181:662–666.
- Kennis JTM, Groot M-L (2007) Ultrafast spectroscopy of biological photoreceptors. *Curr Opin Struct Biol* 17:623–630. doi: 10.1016/j.sbi.2007.09.006
- Kim K (1986a) Three-dimensional analysis of coherent amplification and self-amplified spontaneous emission in free-electron lasers. *Phys Rev Lett*. doi: 10.1103/PhysRevLett.57.1871
- Kim K-J (1986b) An analysis of self-amplified spontaneous emission. *Nucl Inst Methods Phys Res A* 250:396–403. doi: 10.1016/0168-9002(86)90916-2
- Kim PS, Baldwin RL (1982) Specific Intermediates in the Folding Reactions of Small Proteins and the Mechanism of Protein Folding. *Annu Rev Biochem* 51:459–489. doi: 10.1146/annurev.bi.51.070182.002331
- Kim TW, Lee JH, Choi J, et al (2012) Protein Structural Dynamics of Photoactive Yellow Protein in Solution Revealed by Pump–Probe X-ray Solution Scattering. *J Am Chem Soc* 134:3145–3153. doi: 10.1021/ja210435n
- Kirian RA, White TA, Holton JM, et al (2011) Structure-factor analysis of femtosecond microdiffraction patterns from protein nanocrystals. *Acta Crystallogr A* 67:131. doi: 10.1107/S0108767310050981
- Kort R, Hoff WD, Van West M, et al (1996) The xanthopsins: a new family of eubacterial blue-light photoreceptors. *EMBO J* 15:3209–3218.
- Kostera J, McGarry J, Pacheco AA (2010) Enzymatic interconversion of ammonia and nitrite: the right tool for the job. *Biochemistry* 49:8546. doi: 10.1021/bi1006783
- Krissinel E (2015) CCP4 SOFTWARE SUITE: HISTORY, EVOLUTION, CONTENT, CHALLENGES AND FUTURE DEVELOPMENTS. *Arbor-Ciencia Pensam Y Cult*. doi: 10.3989/arbor.2015.772n2006
- Kupitz C, Basu S, Grotjohann I, et al (2014a) Serial time-resolved crystallography of

- photosystem II using a femtosecond X-ray laser. *Nature* 513:261–265.
- Kupitz C, Grotjohann I, Conrad CE, et al (2014b) Microcrystallization techniques for serial femtosecond crystallography using photosystem II from *Thermosynechococcus elongatus* as a model system.
- Laue M, Friedrich W, Knipping P (1913) Interferenzerscheinungen bei Röntgenstrahlen. *Ann Phys* 41, 971 41:971–988.
- LeCroy (2009) LeCroy WaveSurfer Oscilloscope Getting Started Guide. Chestnut Ridge, NY
- Liu W, Wacker D, Gati C, et al (2013) Serial femtosecond crystallography of G protein--coupled receptors.(REPORTS)(Report)(Author abstract). *Science* (80- ) 342:1521.
- Liu W, Wacker D, Wang C, et al (2014) Femtosecond crystallography of membrane proteins in the lipidic cubic phase. *Philos. Trans. R. Soc. B-Biological Sci.* 369:
- Lodish H, Berk A, Zipursky L, et al (2000) *Molecular Cell Biology*, 4th edn. New York: W.H. Freeman
- Madey JMJ (1971) Stimulated Emission of Bremsstrahlung in a Periodic Magnetic Field. *J Appl Phys* 42:1906. doi: 10.1063/1.1660466
- Mccooy A (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr Sect D-Biological Crystallogr* 63:32–41. doi: 10.1107/S0907444906045975
- Mccooy A, Grosse-Kunstleve RW, Adams PD, et al (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674. doi: 10.1107/S0021889807021206
- McPherson 1944- A (1999) *Crystallization of biological macromolecules* / A. McPherson. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, c1999., Cold Spring Harbor, N.Y.
- Meyer TE (1985) Isolation and characterization of soluble cytochromes, ferredoxins and other chromophoric proteins from the halophilic phototrophic bacterium *Ectothiorhodospira halophila*. *Biochim Biophys Acta - Bioenerg* 806:175–183. doi: [http://dx.doi.org/10.1016/0005-2728\(85\)90094-5](http://dx.doi.org/10.1016/0005-2728(85)90094-5)
- Meyer TE, Yakali E, Cusanovich MA, Tollin G (1987) Properties of a water-soluble, yellow protein isolated from a halophilic phototrophic bacterium that has photochemical activity analogous to sensory rhodopsin. *Biochemistry* 26:418–423. doi: 10.1021/bi00376a012
- Moffat K (1989) Time-Resolved Macromolecular Crystallography. *Annu Rev Biophys Biophys Chem* 18:309–332. doi: 10.1146/annurev.bb.18.060189.001521
- Mowat CG, Chapman SK (2005) Multi-heme cytochromes--new structures, new chemistry.

Dalton Trans 3381.

- Murshudov G, Vagin AA, Dodson E (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr Sect D-Biological Crystallogr* 53:240–255.
- Nakamura R, Hamada N, Ichida H, et al (2007) Coherent oscillations in ultrafast fluorescence of photoactive yellow protein. *J Chem Phys* 127:215102.
- Neutze R, Wouts R, van der Spoel D, et al (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* 406:752–757. doi: 10.1038/35021099
- Nogly P, James D, Wang D, et al (2015) Lipidic cubic phase serial millisecond crystallography using synchrotron radiation. *IUCrJ* 2:168–76. doi: 10.1107/S2052252514026487
- Ōnuki H, Elleaume P (eds) (2003) *Undulators, wigglers and their applications / edited by Hideo Ōnuki and Pascal Elleaume*. London : Taylor & Francis, 2003., London
- Opotek (2008a) ICE 450 Power Supply User's Manual. Carlsbad, CA
- Opotek (2008b) Opolette 355 HE II Tunable Laser System User's Manual. Carlsbad, CA
- Palczewski K, Kumasaka T, Hori T, et al (2000) Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor.(Illustration). *Science* (80- ) 289:739.
- Park HY, Qiu X, Rhoades E, et al (2006) Achieving uniform mixing in a microfluidic device: hydrodynamic focusing prior to mixing. *Anal Chem* 78:4465.
- Pellequer J-L, Wager-Smith KA, Kay SA, Getzoff ED (1998) Photoactive yellow protein: A structural prototype for the three-dimensional fold of the PAS domain superfamily. *Proc Natl Acad Sci U S A* 95:5884–5890.
- Perutz MF, Rossmann MG, Cullis AF, et al (1960) Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature* 185:416. doi: 10.1038/185416a0
- Pettersen EF, Goddard TD, Huang CC, et al (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605.
- Phillips DC (1967) THE HEN EGG-WHITE LYSOZYME MOLECULE. *Proc Natl Acad Sci U S A* 57:483–495.
- Pollack L (2011) Time resolved SAXS and RNA folding. *Biopolymers* 95:543–549. doi: 10.1002/bip.21604
- Pp-electronic (2007) SMC9300 motor controller command reference. Rimsting, Germany
- Pp-electronic (2011) SMC9300 motor controller hardware reference. Rimsting, Germany
- Prince E (ed) (2006) *International Tables for Crystallography Volume C: Mathematical, physical*

and chemical tables.

- Purwar N (2013) Structure and Function of Proteins Investigated By Crystallographic and Spectroscopic Time-Resolved Methods. University of Wisconsin Milwaukee
- Purwar N, Tenboer J, Tripathi S, Schmidt M (2013) Spectroscopic Studies of Model Photo-Receptors: Validation of a Nanosecond Time-Resolved Micro-Spectrophotometer Design Using Photoactive Yellow Protein and  $\alpha$ -Phycoerythrocyanin. *Int J Mol Sci* 14:18881–18898. doi: 10.3390/ijms140918881
- Raicu V, Popescu A (2008) Integrated Molecular and Cellular Biophysics, 1st edn. Springer Netherlands
- Ren Z, Bourgeois D, Helliwell JR, et al (1999) Laue crystallography: coming of age. *J Synchrotron Radiat* 6:891–917. doi: 10.1107/S0909049599006366
- Ren Z, Moffat K (1994) Laue Crystallography for Studying Rapid Reactions. *J Synchrotron Radiat* 1:78–82. doi: 10.1107/S0909049594006692
- Ren Z, Perman B, Srajer V, et al (2001) A molecular movie at 1.8 Å resolution displays the photocycle of photoactive yellow protein, a eubacterial blue-light receptor, from nanoseconds to seconds. *Biochemistry* 40:13788.
- Richardson DJ (2000) Bacterial respiration: a flexible process for a changing environment. *Microbiology* 146:551–571.
- Röntgen WC (1898) Ueber eine neue Art von Strahlen. *Ann Phys* 300:1–11. doi: 10.1002/andp.18983000102
- Rupp B (2009) Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology, 1st edn. Garland Science, New York
- Rupp B, Wang J (2004) Predictive models for protein crystallization. 34:390–407.
- Sauter NK, Hattne J, Grosse-kunstleve RW, Echols N (2013) New Python-based methods for data processing. *Acta Crystallogr Sect D* 69:1274–1282. doi: 10.1107/S0907444913000863
- Schluenzen F, Tocilj A, Zarivach R, et al (2000) Structure of Functionally Activated Small Ribosomal Subunit at 3.3 Å Resolution. *Cell* 102:615–623. doi: 10.1016/S0092-8674(00)00084-2
- Schmidt I, Steenbakkens PJM, op den Camp HJM, et al (2004a) Physiologic and Proteomic Evidence for a Role of Nitric Oxide in Biofilm Formation by *Nitrosomonas europaea* and Other Ammonia Oxidizers. *J Bacteriol* 186:2781.
- Schmidt M (2008) Structure Based Kinetics by Time-Resolved X-ray Crystallography. In: Braun M, Gilch P, Zinth W (eds) Ultrashort Laser Pulses in Biology and Medicine SE - 9.



Springer Berlin Heidelberg, pp 201–241

- Schmidt M (2013) Mix and Inject: Reaction Initiation by Diffusion for Time-Resolved Macromolecular Crystallography. *Adv Condens Matter Phys* 2013:1–10. doi: 10.1155/2013/167276
- Schmidt M, Graber T, Henning R, Srajer V (2010) Five-dimensional crystallography. *Acta Crystallogr Sect A Found Crystallogr* 66:198–206. doi: 10.1107/S0108767309054166
- Schmidt M, Ihee H, Pahl R, Šrajer V (2005) Protein-Ligand Interaction Probed by Time-Resolved Crystallography. In: Ulrich Nienhaus G (ed) *Protein-Ligand Interactions SE - 7*. Humana Press, pp 115–154
- Schmidt M, Krasselt A, Reuter W (2006) Local protein flexibility as a prerequisite for reversible chromophore isomerization in alpha-phycoerythrocyanin. *Biochim Biophys Acta-Proteins Proteomics* 1764:55–62. doi: 10.1016/j.bbapap.2005.10.022
- Schmidt M, Pahl R, Srajer V, et al (2004b) Protein kinetics: Structures of intermediates and reaction mechanism from time-resolved x-ray data. *Proc Natl Acad Sci U S A* 101:4799–4804. doi: 10.1073/pnas.0305983101
- Schmidt M, Pande K, Basu S, Tenboer J (2015) Room temperature structures beyond 1.5 Å by serial femtosecond crystallography. *Struct Dyn* 2:–. doi: <http://dx.doi.org/10.1063/1.4919903>
- Schmidt M, Patel A, Zhao Y, Reuter W (2007) Structural basis for the photochemistry of alpha-phycoerythrocyanin. *Biochemistry* 46:416.
- Schmidt M, Rajagopal S, Ren Z, Moffat K (2003) Application of Singular Value Decomposition to the Analysis of Time-Resolved Macromolecular X-Ray Data. *Biophys J* 84:2112–2129. doi: 10.1016/S0006-3495(03)75018-8
- Schmidt M, Srajer V, Henning R, et al (2013) Protein energy landscapes determined by five-dimensional crystallography. *Acta Crystallogr Sect D Biol Crystallogr* 69:2534–2542.
- Schmidt M, Šrajer V, Purwar N, Tripathi S (2012) The kinetic dose limit in room-temperature time-resolved macromolecular crystallography. *J Synchrotron Radiat* 19:264–273. doi: 10.1107/S090904951105549X
- Schmüser P, Dohlus M, Rossbach J, Behrens C (2014) *Free-Electron Lasers in the Ultraviolet and X-Ray Regime*, 2nd edn. Springer International Publishing
- Schotte F, Cho HS, Kaila VRI, et al (2012) Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proc Natl Acad Sci U S A* 109:19256–19261. doi: 10.1073/pnas.1210938109



- Sidler WA (1994) *The molecular biology of cyanobacteria* / edited by Donald A. Bryant. Dordrecht, Dordrecht
- Sierra MJB and DS and CYH and RG (2010) Single-particle coherent diffractive imaging with a soft x-ray free electron laser: towards soot aerosol morphology. *J Phys B At Mol Opt Phys* 43:194013.
- Sierra RG, Laksmono H, Kern J, et al (2012) Nanoflow electrospinning serial femtosecond crystallography. *Acta Crystallogr Sect D* 68:1584–1587. doi: 10.1107/S0907444912038152
- Single Particle Initiative Team (2015) Hitfinding with Hummingbird. <http://spidocs.readthedocs.org/en/latest/examples.html#hitfinding>.
- Sprenger WW, Hoff WD, Armitage JP, Hellingwerf KJ (1993) The eubacterium *Ectothiorhodospira halophila* is negatively phototactic, with a wavelength dependence that fits the absorption spectrum of the photoactive yellow protein. *J Bacteriol* 175:3096–3104.
- Stein N, Love D, Judd ET, et al (2015) Correlations between the Electronic Properties of *Shewanella oneidensis* Cytochrome c Nitrite Reductase (ccNiR) and Its Structure: Effects of Heme Oxidation State and Active Site Ligation. *Biochemistry* 54:3749. doi: 10.1021/acs.biochem.5b00330
- Stellato F, Oberthür D, Liang M, et al (2014) Room-temperature macromolecular serial crystallography using synchrotron radiation. *IUCrJ* 1:204. doi: 10.1107/S2052252514010070
- Strehlitz B, Gründig B, Schumacher W, et al (1996) A nitrite sensor based on a highly sensitive nitrite reductase mediator-coupled amperometric detection. *Anal Chem* 68:807. doi: 10.1021/ac950692n
- Systems SR (2008) *Stanford Research Systems DG645 Digital Delay Generator User Manual*. Sunnyvale, CA
- Taylor BL, Zhulin IB (1999) PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light. *Microbiol Mol Biol Rev* 63:479.
- Tenboer J, Basu S, Zatsepin N, et al (2014) Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* (80- ) 346:1242–1246. doi: 10.1126/science.1259357
- Thorlabs (2014) *Thorlabs SC10 Shutter Controller User Guide*. Newton, NJ
- Tripathi S, Šrajter V, Purwar N, et al (2012) PH dependence of the photoactive yellow protein photocycle investigated by time-resolved crystallography. *Biophys J* 102:325–332. doi: 10.1016/j.bpj.2011.11.4021

- Vagin AA, Steiner R, Lebedev AA, et al (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr Sect D-Biological Crystallogr* 60:2184–2195. doi: 10.1107/S0907444904023510
- Van Beeumen JJ, Devreese B V, Van Bun SM, et al (1993) Primary structure of a photoactive yellow protein from the phototrophic bacterium *Ectothiorhodospira halophila*, with evidence for the mass and the binding site of the chromophore. *Protein Sci* 2:1114–1125. doi: 10.1002/pro.5560020706
- van der Rest M, Garrone R (1991) Collagen family of proteins. *FASEB J* 5:2814–2823.
- Vengris M, van der Horst MA, Zgrablić G, et al (2004) Contrasting the Excited-State Dynamics of the Photoactive Yellow Protein Chromophore: Protein versus Solvent Environments. *Biophys J* 87:1848–1857. doi: 10.1529/biophysj.104.043224
- Venkateswaran K, Moser DP, Dollhopf ME, et al (1999) Polyphasic taxonomy of the genus *Shewanella* and description of *Shewanella oneidensis* sp. nov. *Int J Syst Bacteriol* 49:705–724. doi: 10.1099/00207713-49-2-705
- Vreede J, van Der Horst MA, Hellingwerf KJ, et al (2003) PAS domains. Common structure and common flexibility. *J Biol Chem* 278:18434.
- Wagner JR, Brunzelle JS, Forest KT, Vierstra RD (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 438:325. doi: 10.1038/nature04118
- Walters EC (ed) (2011) Protein folding Eric C. Walters, editor. New York. : Nova Science, ©2011
- Wang D, Weierstall U, Pollack L, Spence J (2014) Double-focusing mixing jet for XFEL study of chemical kinetics. *J Synchrotron Radiat* 21:1364–1366. doi: 10.1107/S160057751401858X
- Weierstall U (2014) Liquid sample delivery techniques for serial femtosecond crystallography. *Philos Trans R Soc L B Biol Sci* 369:20130337. doi: 10.1098/rstb.2013.0337
- Weierstall U, James D, Wang C, et al (2014) Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography.
- Weierstall U, Spence JCH, Doak RB (2012) Injector for scattering measurements on fully solvated biospecies. *Rev Sci Instrum* 83:-. doi: <http://dx.doi.org/10.1063/1.3693040>
- White TA (2014) Post-refinement method for snapshot serial crystallography. *Philos Trans R Soc Lond B Biol Sci* 369:20130330. doi: 10.1098/rstb.2013.0330
- White TA, Barty A, Stellato F, et al (2013) Crystallographic data processing for free-electron

- laser sources. *Acta Crystallogr Sect D* 69:1231–1240. doi: 10.1107/S0907444913013620
- White TA, Kirian R, Martin A, et al (2012) CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* 45:335–341.
- Wimberly BT, Brodersen DE, Clemons WMJ, et al (2000) Structure of the 30S ribosomal subunit. *Nature* 407:327.
- Winick H, Brown G, Halbach K, Harris J (1981) Wiggler and undulator magnets. *Phys Today* 34:50. doi: 10.1063/1.2914568
- Winn MD, Ballard CC, Cowtan KD, et al (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67:235–242.
- Wu J, Emma P (2006) *The Linac Coherent Light Source (LCLS) Accelerator*. Menlo Park, CA
- Youngblut M (2013) Towards the use of time-resolved X-ray crystallography in mechanistic studies of cytochrome c nitrite reductase from *Shewanella oneidensis*.
- Youngblut M, Judd E, Srajer V, et al (2012) Laue crystal structure of *Shewanella oneidensis* cytochrome c nitrite reductase from a high-yield expression system. *J Biol Inorg Chem* 17:647–662. doi: 10.1007/s00775-012-0885-0
- Youngblut M, Pauly DJ, Stein N, et al (2014) *Shewanella oneidensis* Cytochrome c Nitrite Reductase (ccNiR) Does Not Disproportionate Hydroxylamine to Ammonia and Nitrite, Despite a Strongly Favorable Driving Force. *Biochemistry* 53:2136–2144. doi: 10.1021/bi401705d
- Zhao Y, Schmidt M (2009) New software for the singular value decomposition of time-resolved crystallographic data. *J Appl Crystallogr* 42:734–740. doi: 10.1107/S0021889809019050
- (2015a) LCLS Publications: Journal Papers Statistics.  
[https://portal.slac.stanford.edu/sites/lcls\\_public/Pages/pub\\_stats.aspx](https://portal.slac.stanford.edu/sites/lcls_public/Pages/pub_stats.aspx).
- (2015b) APS Science 2014. Argonne, IL

# Curriculum Vitae

## Jason Tenboer

University of Wisconsin-Milwaukee, Department of Physics, Kenwood Interdisciplinary Research Complex (KIRC), 3135 North Maryland Ave., Milwaukee, WI, 53211 |262-302-0507| [jtenboer@uwm.edu](mailto:jtenboer@uwm.edu)

## EDUCATION

University of Wisconsin-Milwaukee 2011-2015

### PhD in Physics

Dissertation: Time-resolved, near atomic resolution structural studies at the free electron laser

University of Wisconsin-Milwaukee 2005-2011

### BSc In Physics

Received with Honors

## HONORS AND AWARDS

**The David Lichtman Scholarship** (for outstanding performance in Experimental Physics), May 2014

**Distinguished Dissertator Fellowship**, September 2015

## PUBLICATIONS

### *REFEREED PAPERS*

2013

1. Purwar, N., **Tenboer, J.**, Tripathi, S. & Schmidt, M. "Spectroscopic studies of model photo-receptors: validation of a nanosecond time-resolved micro-spectrophotometer design using photoactive yellow protein and alpha-phycoerythrocyanin". *Int J Mol Sci* **14**, 18881-98 (2013).
2. Schmidt, M., Srajer, V., Henning, R., Ihee, H., Purwar, N., **Tenboer, J.**, Tripathi, S. "Protein energy landscapes determined by five-dimensional crystallography". *Acta Crystallogr D Biol Crystallogr* **69**, 2534-42 (2013).

2014

3. van Thor, J.J., Warren, M.M., Lincoln, C.N., Chollet, M., Lemke, H.T., Fritz, D.M., Schmidt, M., **Tenboer, J.**, Ren, Z., Srajer, V., Moffat, K., Graber, T. "Signal to noise considerations for

single crystal femtosecond time resolved crystallography of the Photoactive Yellow Protein”. *Faraday Discuss* **171**, 439-55 (2014).

4. **Tenboer, J.** et al. “Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein”. *Science* **346**, 1242-6 (2014).

2015

5. Schmidt, M., Pande, K., Basu, S., **Tenboer, J.** “Room temperature structures beyond 1.5 Å by serial femtosecond crystallography”. *Structural Dynamics* **2**, 041708 (2015)

6. Pande, K., Hutchison, C.D.M., Groenhof, G., Aquila, A., Robinson, J.F., **Tenboer, J.**, et al. “Femtosecond Structural Dynamics Drives the Trans/Cis Isomerization in Photoactive Yellow Protein”, *Science* (submitted, under review) (2016)

#### SPECIAL PRESENTATIONS AND INVITED TALKS

1. **Jason Tenboer**, *Design and Preliminary Results of an In-House Micro-Focus X-ray Source*, UWM Biophysics Seminar, March 2013

2. **Jason Tenboer**, *SBL1 and SBL2: In-House Tools for XFEL and Synchrotron Experiment Preparation*, UWM Biophysics Seminar, January 2013

3. **Jason Tenboer**, *Time-Resolved Serial Femtosecond Crystallography Reveals High Resolution Intermediates of PYP from High Quality Difference Electron Density Maps*, UWM Biophysics Seminar, October 2014

4. **Jason Tenboer**, *Making Movies of Biological Macromolecules using Time-Resolved Serial Femtosecond Crystallography*, FEIS-2 Conference, Michigan State University, May 2015

5. **Jason Tenboer**, *Protein Dynamics*, 3rd Ringberg Workshop on Structural Biology with FELs – from first results to future perspectives, February 2016

#### POSTERS

1. **Jason Tenboer**, *Time-Resolved Serial Femtosecond Crystallography Reveals Intermediates of Photoactive Yellow Protein*, 2<sup>nd</sup> International BioXFEL Conference, Ponce, Puerto Rico, January 2014

2. **Jason Tenboer**, *Time-Resolved Microspectrophotometer and Micro-Focus X-ray Source*, University of Wisconsin-Milwaukee Undergraduate Research Symposium, April 2011