universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Radiation Damage in Metalloproteins – a Scavenger Approach

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer.nat.)

Verfasserin / Verfasser:	Maria Pechlaner
Matrikel-Nummer:	0206938
Studienrichtung (lt. Studienblatt):	Molekulare Biologie
Betreuerin / Betreuer:	Kristina Djinović-Carugo

Wien, am

Radiation Damage in Metalloproteins - a Scavenger Approach

Diplomarbeit

Vorgelegt am Institut für biomolekulare Strukturchemie
an der Universität Wien 2008

Verfasst
Von

Maria Pechlaner

Betreut
Von

Kristina Djinović-Carugo
und
Sofia Macedo

Table of Contents

Table of Abbreviations....9
1 Introduction....10
1.1 What happens when ionizing radiation hits the crystal?....10
Damage spreads - a cascade of events....11
1.2 Possible reaction mechanisms....12
1.3 Consequences for the crystallographer....13
Global deterioration....13
Heating effects....14
Specific damage....14
Metal sites....15
1.4 Studying radiation damage....15
Dose....16
Dose rate....16
Energy dependence....17
Temperature....17
Crystal size....18
Scavengers....18
1.5 Living with the problem ... and even taking advantage....20
Zero-dose extrapolation....20
Multi-crystal data collection....20
Radiation-induced phasing (RIP)....21
1.6 Objectives of the project....22
2 Materials....23
2.1 The proteins....23
Human transferrin (Tf)....24
Cytochrome c from bovine heart (Cytc)....24
Cu, Zn superoxide dismutase from bovine erythrocytes (CuZnSOD)....25
Ni-containing superoxide dismutase from Streptomyces seoulensis (NiSOD)....26
Insulin from bovine pancreas....27
Azurin from Pseudomonas aeruginosa....28
Horse heart myoglobin....29
2.2 The scavengers....29
a) small organic molecules....29
b) transition metal compounds....30
c) aromatic systems....31
2.3 Crystallization utensils....32

a) Crystallization screens....32
b) Crystallization plates....32
c) Loops....32
2.4 Buffers and solutions....33
Reducing agents 33
Buffer for size exclusion chromatography....33
Electrophoresis buffers....33
2.5 Instruments....34
Rotating anode home source....34
Beamlines....35
3 Methods....37
3.1 Protein purification....37
Size exclusion chromatography....37
Concentrating proteins....37
Determination of protein concentration....37
Gel electrophoresis....38
3.2 Dynamic Light Scattering (DLS)....38
3.3 Crystallization....38
Vapor diffusion....39
Crystallization robot....40
Screening for the best conditions....40
Micro- and macroseeding....40
3.4 Soaking and freezing crystals....40
Chemically reduced crystals....41
3.5 X-ray diffraction experiments....41
Data collection....42
Dose calculation....42
3.6 Integration and Scaling....43
XDS package....43
Proteum software....44
3.7 Reindexing 44
3.8 Detwinning....44
3.9 Molecular replacement and refinement....44
3.10 Calculating maps and peak search....45
3.11 Single crystal UV-vis microspectrophotometry....45
4 Results and Discussion....48
4.1 Purification and quality check....48
NiSOD purification....48
Purity check for transferrin and cytochrome c....48

Transferrin....48
Cytochrome c....49
4.2 Dynamic Light Scattering (DLS)....49
4.3 Crystallization....50
Azurin....51
CuZnSOD....51
Insulin....52
Myoglobin....53
NiSOD....53
4.4 Scavenger solubility test....54
4.5 Soaking, freezing, data collection....56
Exclusion of NiSOD and CuZnSOD....56
Beam and crystal size....56
Oscillation range....58
Doses....58
Data collection at the home source....58
4.6 Integration and scaling – data statistics and global signs of radiation damage....59
Azurin....59
Rmrgd, Rmeas....60
I/Sigma....61
Wilson B factors....61
Completeness of data sets, cell volume....61
Mosaicity....61
Insulin....62
Myoglobin and the analysis of data collected on the home source....63
4.7 The Rd....64
4.8 Structure refinement....67
Azurin 68
Myoglobin....69
Insulin and crystal twinning....70
4.9 Analysis of local changes....73
4.9.1 General features of the Fo-Fo difference maps....73
4.9.2 Comparison and evaluation of electron difference map peaks....76
Overall peak statistics....77
4.10 Microspectrophotometry....80
4.10.1 Preparation....80
4.10.2 General....81
4.10.3 Azurin....83
A) Total peak decrease....84

a) Scaling to basal absorbance....84
b) Assuming completely oxidized samples....85
B) Rate constants....90
4.10.4 Myoglobin....92
Time course fitting....92
5 Conclusions and outlook....95
5.1 Diffraction data....95
5.2 Microspectrophotometry data....96
6 Bibliography....98
7 Appendix....104
7.1 Data collection specifics....104
7.2 Scaling statistics....105
7.2.1 Azurin....105
7.2.2 Insulin....106
7.2.3 Myoglobin....107
7.3 Rd plots....108
7.3.1 Rd regression statistics....109
Myoglobin....109
Azurin....110
Insulin....110
7.4 Microspectrophotometry - Spectra and timecourses....111
7.5 Timecourse regression statistics....118
7.5.1 Azurin....118
7.5.2 Myoglobin....119
Zusammenfassung....120
Abstract....121
Curriculum Vitae....122

Table of Abbreviations

AQ	anthraquinone
az	azurin
BESSY	Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung
BQ	benzoquinone
CCP4	Collaborative Computational Project Number 4 in Protein Crystallography
CuZnSOD	Cu, Zn superoxide dismutase
cys	cysteine
CytC	cytochrome c
DDQ	1,2-Dichloro-4,5-dicyanobenzoquinone
DLS	dynamic light scattering
DNQ	2,3-Dichloro-1,4-naphtoquinone
DPI	Dichlorophenol-Indophenol
DTNB	5,5'-Dithio-Bis (2-Nitrobenzoic Acid)
ESRF	European Synchrotron Radiation Facility
Fc	calculated structure factor
Flu	fluorescein
Fo	observed structure factor
hep	HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
ins	insulin
KF	$K_3[Fe(CN)_6]$ (potassium hexacyanidoferrate)
MA	maleic acid anhydride
MAD/SAD	multiple/single anomolous dispersion
MIR/SIR	multiple/single isomorphous replacement
MR	molecular replacement
NA	nicotinic acid
NF	$Na_2[Fe(CN)_5]*2H_2O$ (sodiumnitroprusside)
NI	nitroimidazole
NiSOD	Ni containing superoxide dismutase
NSA	sodium salicylate
PAGE	polyacrylamide gel electrophoresis
PDB	Protein Data Bank
RIP	radiation-induced phasing
SAA	sodium ascorbate
SDS	sodium dodecyl sulfate
TBQ	Tetrachloro-o-benzoquinone
Tf	transferrin
TU	thiourea

1 Introduction

X-rays are highly ionizing radiation and have a deleterious effect on all organic matter if applied in high enough dose. And yet we remain dependent on them when it comes to solving macromolecular structures at high resolution. Macromolecular X-ray crystallography is a powerful method that can show us structural details at atomic and subatomic level. As more powerful X-ray sources are being developed, the demand for more efficient means to alleviate the damage inflicted by the X-rays and thereby increase crystal lifetime in the beam becomes inevitable. The study of radiation damage has therefore been a highly active field of research since the beginnings of macromolecular X-ray crystallography.

1.1 What happens when ionizing radiation hits the crystal?

A typical X-ray photon used in protein crystallography carries an energy between 5 and 17 keV. When it hits the crystal it can interact with matter in three different ways: Rayleigh scattering, Compton scattering or photoelectric absorption[1]. The latter two, while utterly useless for a crystallographer's purposes, account for over 92% (at 1 Å) of events taking place in the beam and are responsible for radiation damage.

When a photon transfers all of its energy to an electron, which is thus expelled from its shell, we speak of the photoelectric effect. The electron carries the acquired energy as kinetic energy and will gradually lose it to whatever reaction partners it finds on its path of a few micrometers of length (for a 10 keV photoelectron[1]). It leaves behind an electron hole, which can be filled from a higher shell. The energy set free in this process can lead to the ejection of more electrons (Auger effect), resulting in a multiply charged absorbing atom and a build up of charge in the neighboring regions.

Compton (or incoherent) scattering transfers only part of a photon's energy to an electron. The photon is scattered in a different direction and with longer wavelength, adding to background scatter. The energy it deposits may be enough to lead to the ejection of the electron.

Only 8% (at 12.4 keV) of interactions between X-ray photons and electrons lead to the elastic or coherent scattering known as Rayleigh scattering that is enhanced by the regular repeats of the crystal to produce the beautiful diffraction pattern we hope to record on our detector. No energy is deposited into the sample and the photon maintains its wavelength (with a phase shift of 90°). The integrated intensity of X-ray scattering, which we wish to maximize, is dependent on the atomic number, the square of the wavelength and the

scattering angle[2].

The relative importance of the three effects – Compton and Rayleigh Scattering and the photoelectric effect – depends on absorption cross-sections. Looking at the distribution of these cross-sections for carbon (Fig. 1.1, taken from Nave[1]), we see that the commonly used wavelengths (wavelength[Å] = 12.4/energy[keV]) already exploit the region with the best cost-benefit ratio.

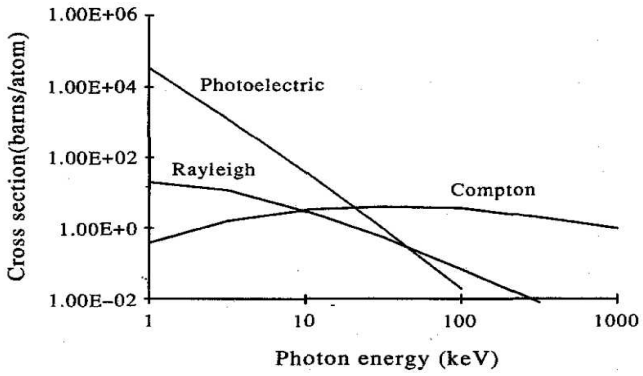


Fig. 1.1: Absorption cross section for photoelectric effect, Rayleigh and Compton scattering.

Longer wavelengths provide higher scattering intensity, but the effect is counteracted by increased absorption. Therefore only strong synchrotron beamlines can afford smaller wavelengths, because the weaker scattering can be compensated for by high flux.

Damage spreads - a cascade of events

Primary damage, i.e. ionizations and excitations inflicted by the incoming photons, is followed by a cascade of secondary and tertiary effects. Photoelectrons ejected from protein or solvent molecules (direct and indirect effects, respectively) travel through the crystal gradually losing their energy to their surroundings. Further electrons of lower energy are set free. Energy is not deposited uniformly, but in so-called spurs along the photoelectron's path. When the energy has sunk below 5 keV, spurs grow into denser "short tracks", which, at 100-500 eV, further condense to blobs with high concentrations of radical species[3],[4]. The close neighborhood of reactive species allows for reactions that might otherwise not be possible.

The main products of ionization processes are hydroxyl radicals, solvated electrons, protons and hydrogen radicals. Hydroxyl radicals are very unstable by themselves and react quickly with various organic groups, their favored reaction being the abstraction of a hydrogen atom from backbone amides[5]. Solvated electrons are highly active reductive agents and retain a considerable mobility at 100 K[6]. They have high potential to interact

with different groups of the protein and to introduce transient and persistent chemical changes. One primary photoelectron can produce up to 500 secondary lower energy electrons[7].

1.2 Possible reaction mechanisms

Studies have so far been concentrated on reactions in aqueous solution at room temperature, where indirect effects by water radicals are the main events. Those are largely inhibited at 77 K, and so direct effects on the protein and its primary hydration shell are the almost exclusive source of damage.

A neighboring $H_2O^{+\cdot}$ might, by withdrawing an electron, induce the following reaction, which is a major cause of persistent damage at 77 K[8]:

$$\begin{array}{c} \text{-CHR-C(O)NH-CHR'-} \\ \downarrow \text{ -e}^- \\ \text{-CHR-C(O)N}^{\cdot}\text{-CHR'-} + H^+ \end{array} \qquad (1)$$

At carboxylic groups the loss of an electron can lead to decarboxylation[9],[10]:

$$\begin{array}{c} \text{-CHR-CO}_2^- \\ \downarrow \text{ -e}^- \\ \text{-}^{\cdot}\text{CHR} + CO_2 \end{array} \qquad (2)$$

The main electron gain sites in a protein are carbonyls and disulfide bridges. The trapping of an electron may lead to the breakage of bonds through the following mechanisms[7],[10]:

$$\begin{array}{c} \text{-CHR-C(O)NH-CHR'-} + e^- \\ \downarrow \\ \text{-CHR-}^{\cdot}\text{C(O-)NH-CHR'-} \\ \downarrow \\ \text{-CHR-C(O)NH}_2 + {}^{\cdot}\text{CHR'-} \end{array} \qquad (3)$$

$$\begin{array}{c} \text{Cystine} + e^- \\ \downarrow \\ \text{RSSR}^{\cdot -} \\ \downarrow \\ \text{RS}^{\cdot}, \text{RSH}, \text{RS}^- \end{array} \qquad (4)$$

Chemical agents that are able to either donate electrons to an electron hole or take up an electron from an electron gain center could possibly restrict the damage and/or stop a chain of destructive events.

1.3 Consequences for the crystallographer

Global deterioration

The general effect of radiation damage is a decrease in diffraction quality up to a total loss of interpretable diffraction, thus limiting the amount of data that can be collected from a crystal. Primary damage and reactive species introduce a multitude of small changes increasing sample mosaicity. Not only protein moieties are modified by radiation damage effect. As damage accumulates, the packing of the molecules can be affected up to a total loss of crystal order. If crucial crystal contacts are concerned, crystal dissolution will happen long before damage would otherwise become noticeable.

Crystals subjected to high radiation dose can get visibly marked, showing brown spots where they were hit by the beam. On warming up they are furthermore observed to 'bubble', releasing a so far unidentified gas[11].

To assess the damage a crystal has suffered, a variety of factors have been analyzed:

- diffraction intensity (I/sigma)
- data quality as assessed by the calculation of R factors (R factors indicate the agreement of the crystallographic model with the experimental data and are of the general form $R=\frac{\sum |F_{obs}|-|F_{calc}|}{\sum F_{obs}}$ where F_{obs} and F_{calc} are observed and calculated structure factors)
- the Wilson B factor
- mosaicity
- unit cell volume: has been shown to increase with dose[12],[10],[13], but not sufficiently reproducible; it is not yet clear what causes the expansion
- resolution limits: high resolution reflections are lost

Alhough all of these factors can generally serve as indicators of radiation damage and together as a whole can give us a qualitative picture of how much a crystal has suffered, none of them can be used to evaluate radiation damage quantitatively and reliably on an absolute scale. Their behavior in different crystals can be unreliably different. Additionally they are mostly varying with the phi range of data collection due to anisotropy. To find a metric that will allow accurate and sensitive scientific evaluation of efforts to reduce radiation damage therefore remains a topic of high interest. Ideally, it should be applicable to the raw data with as little prior analysis and treatment needed as possible while retaining a maximum in sensitivity. Kay Diederichs suggested a new parameter for the

purpose that might come close to these specifications, although comparisons between different crystals are still to be taken with reservation[14].

In analogy to other R factors, Diederichs defines a new *decay R factor* that is a function of the frame difference. The exact definition is the following:

$$R_d = \frac{\sum_{hkl} \sum_{|i-j|=d} |Y_i - Y_j|}{\sum_{hkl} \sum_{|i-j|=d} (Y_i + Y_j)/2} \qquad (5)$$

where Y_i and Y_j are the observed intensities of unique reflection hkl with centroids on frames i and j. The R_d will rise as the intensity of the same reflection on different frames is changed by accumulating radiation damage. Due to the contribution of all the reflection pairs with the same frame distance a relatively smooth curve results. Only at 180° and its multiples sharp wiggles are produced, due to the inherent symmetry of the crystal diffraction pattern. The R_d can be calculated using XDSSTAT, a program developed by Kay Diederichs.

Heating effects

Crystal heating due to the deposition of energy by the X-ray beam was originally considered a possible cause for the deteriorating effects observed. However, theoretical studies have demonstrated that even at the more powerful beamlines in use today, crystal heating at cryotemperatures would not exceed 15 K[15], which is not likely to cause any significant problems. Only above 155 K[16] does increased conformational flexibility lead to a much enhanced radiation sensitivity.

Specific damage

The global loss of data quality sets a limit to the lifetime of a crystal in the X-ray beam. But even before the diffraction pattern is compromised, modifications take place at specific sites of the protein molecule. It has been shown that damage is not distributed uniformly over all the residues[17], but that some groups are more sensitive than others. More treacherous than the general degradation of the crystal, these effects may not always be recognized as what they are, and may instead be taken as genuine features of the structure.

Sulfur groups and carboxyls as well as metals are especially sensitive to radiation. Disulfide bridges break easily, equation (4) on page 12 depicts a possible reaction mechanism. Thiol and methylthiol groups of cysteines and methionines as well as carboxyl groups lose electron density (see the reaction mechanism in equation (2) on page 12).

There seems to be a difference in the sensitivity between the same groups at different sites of the protein. While one disulfide bridge might be broken early during data collection, another may hold no matter what dose applied. It is not yet clear which factors determine this sensitivity to radiation damage, surface accessibility alone can not account for it[12],[10]. Amino acids in a more stable environment like α-helices or β-sheets tend to be less damaged than those in flexible loops. A study of malate dehydrogenase from a halophilic archaebacterium (chosen for its high abundance in acidic residues)[18], proposes an inverse relation between the volume of a solvent-filled cavity and the radiation sensitivity of an acidic residue exposed to it. The same study indicates a protecting effect of stabilizing salt-bridges.

Several groups have observed a tendency for active sites to be especially radiation sensitive and prone to changes[19],[12],[20]. This might be explained by the fact that residues in the active site are often in a strained conformation. Thus the specific effects of radiation damage might on top of all be focused on the sites of highest interest of the protein.

Apart from the possible disruption of defined protein features, specific radiation damage can seriously threaten the solution of a structure by MAD (multiple wavelength anomalous dispersion), which depends on isomorphous data sets.

Metal sites

One of the first things that happen when the crystal of a metalloprotein is exposed to X-ray radiation is photo-reduction of the metal. While the oxidation state of a metal is not obvious from routine X-ray diffraction experiments, it can have considerable impact on the structure of the metal's environment[21]. Metal coordination has been studied by extended X-ray absorption fine structure (EXAFS) and X-ray absorption near-edge structure (XANES) [22],[23]. For those metals with spectroscopically distinct states, on-line microspectrophotometry has been used to follow changes in the oxidation state[24].

1.4 Studying radiation damage

Radiation damage has been an almost constant companion of the protein crystallographer, setting limits to his work that need to be overcome. For a while the problem was pushed to the background when cryocooling to liquid nitrogen temperatures became common practice about 20 years ago. At 100 K the mobility of reactive species in the crystal is reduced and crystal lifetime is extended 70-fold[11]. In 1990 Henderson predicted a lifetime of 5 years for a cryocooled crystal on a rotating anode and a lifetime of one day on the most powerful synchrotron source available at that time[25]. Today the

same dose can be reached within a single data set at a highly-brilliant third-generation undulator beamline and radiation damage has again become a limiting factor and therefore a highly active field of research.

Several factors have been explored as to their influence on the extent of radiation damage, such as dose, dose rate, photon energy, temperature, crystal size and the addition of electron scavengers.

Dose

Radiation damage is strictly dependent on radiation dose[26], which makes the determination of the total absorbed dose an indispensable prerequisite for any kind of study in the field. Regular calibration of the photon flux with the help of a photodiode to be able to relate flux in the ion chamber to the flux at the crystal position is therefore a must.

The dose is expressed in Gray, which is a measure of Joules per kg. Apart from the photon energy E, the number of photons n, duration t and area A of irradiation, it is also dependent on the mass absorption coefficient μ/ρ.

$$D = \frac{(\mu/\rho) nteE}{A} x10^{11}$$

μ ..... absorption coefficient
ρ ..... density
e ..... 1.6 x 10^{-19} J eV^{-1}

An accurate knowledge of the beam flux and size is of high importance, as are unit cell size and contents, since they determine the absorption coefficient.

According to Henderson[25], who did a theoretical calculation based on damage effects in electron microscopy, a dose of 20 MGy would reduce a crystal's diffraction intensity to half of its original value. Experiments have recently extended this limit to 40 MGy, with a proposition of 30 MGy as an allowed limit where the crystal will have lost 30% of its diffraction intensity ($I=I_0 \times \ln 2$) and structural information will be compromised[27]. Note that a high content of heavy atoms will lower the expected lifetime substantially due to large absorption cross-sections. It is therefore advisable to back-soak heavy-atom derivatives before freezing them to resolve the excess metal. In any case, the proposed dose limits can only give a theoretical maximum lifetime for the crystal, specific changes to the protein molecules are to be expected long before this dose is reached.

Dose rate

Apart from the total absorbed dose, does the rate at which the dose is delivered matter? It can be speculated that above a certain critical dose rate the steady-state free-radical concentration can reach a critical point, where recombination and diffusion are no

longer able to quench certain effects[28]. On the other hand, very high dose rates might reduce the amount of damage when it allows the collection to be completed before secondary damage becomes severe[7]. This would require collection of a full data set on the micro- to picosecond scale.

Sliz *et al*. experimentally concluded dose rate does not affect global diffraction intensity[26] at 100 K. Their results were confirmed by Leiros *et al*.[29], who administered the same overall dose at different dose rates and found no discernible differences in global damage indicators. What they did note, however, was a dose rate effect on the specific damage of sensitive side chains, where higher dose rates lead to slightly higher damage. Nonetheless, this effect was very small and they concluded that dose rate effects at the currently employed photon fluxes are insignificant compared to the effect of total absorbed dose.

At room temperature, damage continues in the protein crystal after the exposure to X-rays[13]. Very low dose rates will therefore lead to higher damage, for the simple reason that there is more time in which the radicals can diffuse through the crystal and cause damage.

Energy dependence

X-rays with different energies ranging from 6.5 to 33 keV have been investigated to see if a change of wavelength can have beneficial effects for the prevention of radiation damage[30],[31]. No correlation between photon energy and radiation damage could be observed, damage was solely dependent on the dose. Be aware, however, that the choice of wavelength has an influence on the absorbed dose, which is highly increased at the absorption edges of elements present in the crystal as described by the following relations: the absorption coefficient μ is dependent on the number of atoms per molecule (N) and the total absorption cross section δ of the atom:

$$\mu = N\,\delta$$

The absorption cross section again, can be expressed as

$$\delta = 2\,\lambda\,\Delta f''$$

where Δf" is the imaginary component of the anomolous scattering factor with a maximum at the absorption edge.

Temperature

When radiation damage became a problem also at liquid nitrogen temperatures, people started to wonder if going even lower would maybe help. Several studies with helium streams (40 K) have been done since[32],[33], but there is no definite answer yet. In some cases[34] the use of helium made all the difference, but in general the beneficial effects –

if existent at all - are considered too small and not worth the considerable additional cost and effort helium cooling implies[35].

Crystal size

Nave *et al.*[36] believe that in small crystals only a few micrometers of diameter, more radicals and free electrons produced inside the crystal will quickly find their way out and not deposit their energy inside the crystal. If reduced radiation damage can be achieved with small crystals remains yet to be shown experimentally. Special setups, possibly based on those used in electron microscopy, would be required to prevent charge build-up and electron flow from the surroundings to the crystal[36].

Scavengers

Mobile reactive species are responsible for specific secondary damage during X-ray exposure. Their mobility can only partly be restrained by lowering the temperature to 100 K, especially solvated electrons are still able to move quickly through the crystal. Secondary damage is thus not overcome and interest has risen again to find ways of mitigation. One possible way to achieve that is by the addition of small molecules with the ability to bind or deactivate free radicals, before any damage is done. Such electron or radical scavengers should be able to take on or give up one or more electrons, forming a more stable radical from the labile and reactive species produced by the radiation.

This idea is not new and has lead to success in other fields. The method is commonly employed in electron spin resonance spectroscopy, where scavengers act as so-called 'spin traps' [13]. 5,5-dimethylpyroline-N-oxide (DPMO) and 2,2,6,6-tetra-methyl-4-piperidone (TEMP) have found application there due to their ability to form stable radicals. Also in the protection of DNA in solution the approach has been successful[37].

Scavengers have been tested in macromolecular X-ray crystallography at room temperature. Zaloga and Sarma[38] found styrene to be able to extend their crystals' lifetime and improve resolution. Additionally at RT different polyethylene glycols have been reported to have a protective effect[39]. However, the beneficial effect of styrene could not be reproduced by Murray *et al.*[13] in cryocooled crystals and also no protection could be attributed to any PEG explored in an experiment using a test system[40] for the simulation of disulfide bridges.

Nonetheless not only PEGs but many other cryoprotectants like glycerol and sugars are by themselves efficient hydroxyl radical and hydrogen atom scavengers[7]. They are frequently present in the cryosolution at high concentrations and since they are necessary for the crystallization and flash-cooling experiment it is difficult to assess their influence in

radiation damage.

A selection of suggested candidate substances with scavenging properties includes the following: ethanol, thiourea[41], ascorbate, cysteine, glutathione, t-butanol, N-2-hydroxy-ethylpiperazine-NH-2-ethanesulfonic acid (HEPES), tris(hydroxy-methyl)-aminoethane (Tris), glucose and ethylene glycol[13]. Ascorbate has recently been the most extensively examined compound. It could be shown that lysozyme co-crystallized with 0.5 M ascorbate kept B-factors almost stable while a dose of 10 MGy (50% of the Henderson limit) was administered. This and a significantly lower increase of the absolute scale factor were accompanied by clear differences in electron density at the sensitive cystine bridges of the native crystal indicating a breakage of bonds that was almost not present in the ascorbate-co-crystallized one. A beneficial effect of ascorbate was also confirmed by Southworth-Davies[40], who concluded that reduced DTT, quinones, and TEMP might be useful as scavengers in X-ray crystallography. While ascorbate is known to be a hydroxyl radical scavenger and antioxidant, it is not known through which mechanism these features could be involved in the protection of disulfide bonds. This might be an indirect rather then a direct effect[13].

In order to establish a more ready-to-use and widely applicable method for employing scavengers, it would be preferable to avoid the need of co-crystallization. Promising experiments where the crystal is only briefly soaked in a solution of the scavenger prior to flash-freezing have been undertaken. Kauffmann *et al.*[42] did studies on crystals of three different proteins. They dissolved there candidate substances – 5,5'-dithio-bis-2-nitrobenzoic acid (DTNB), nicotinic acid and oxidized glutathione – directly in the well solution, soaked their crystals for less than a minute and froze them after removing excess liquid in oil for cryoprotection. To evaluate the effectiveness of a scavenger they concentrate on a comparison of R_d plots[14] (as defined on p. 14) of native and scavenger data sets, accompanied by observations of density loss at acidic and sulfur containing residues and conclude that both DTNB and nicotinic acid are able to decrease the damage suffered by their crystals[42].

Some promising results have been obtained with electron scavengers, but much more data is necessary to confirm them. There is a lot of room for the discovery of new protecting agents and until their mechanism of action and thereby the specifications for an efficient electron scavenger for cryocooled macromolecular crystals are elucidated, this means a lot of – still very labor-intensive – testing.

1.5 Living with the problem … and even taking advantage

For the moment the problem of radiation damage is persisting and we have to find ways to live with it. While improving crystal size and quality – thus improving diffraction quality and being able to reduce the required dose – has been and remains the most favorable solution, it seems in many cases unreachable. To make the best of the crystals we have, we need efficient prediction of data collection strategies that will help to obtain a maximum of information with a minimum of applied dose[43]. Other theoretical efforts are focusing on scale factors that will take into account radiation damage more specifically.

Zero-dose extrapolation

Traditional scaling takes into account effects of radiation damage by correcting for the loss of diffraction intensity, assuming that to a good approximation it is decreasing monotonically. However, specific structural changes as well as the increase in cell volume and small rotations and translations of the molecules are not implied. Diederichs attempts to include such effects with a redundancy-based linear correction model[44], and later[14] also with exponential and quadratic corrections. In this study, changes due to radiation damage occur monotonically and smoothly in real space, and therefore Fourier transform and intensities will also change smoothly but not necessarily monotonically. The previously mentioned R_d was used to demonstrate the effectiveness of the corrections.

If such approaches can be extended, will it become possible to calculate snapshots of a protein at specific dose through extrapolation? Perhaps even to generate before and after X-ray burn data sets for phasing using RIP (radiation-damage induced phasing, see below). To this date only a few such snapshots have been obtained experimentally using a multi-crystal data collection procedure[45],[46].

Multi-crystal data collection

A complete data set with low radiation dose can be obtained if the protein yields enough crystals of reproducible quality. Complete data sets are collected from several crystals and the first few degrees of each are combined to one low-dose data set. The same procedure will allow us to look at the protein in several different states of low, intermediate and high dose (see Fig. 1.2 for illustration of the method). Apart from especially radiation sensitive proteins, this is especially interesting for people working with redox-sensitive proteins[45],[46]. Through the multi-crystal approach it has been possible to obtain structures of oxidized reaction intermediates that are usually already reduced by the X-rays before a complete data set can be collected. In this special case, radiation damage is no longer a burden, it becomes a very useful tool in the elucidation of reaction

mechanisms.

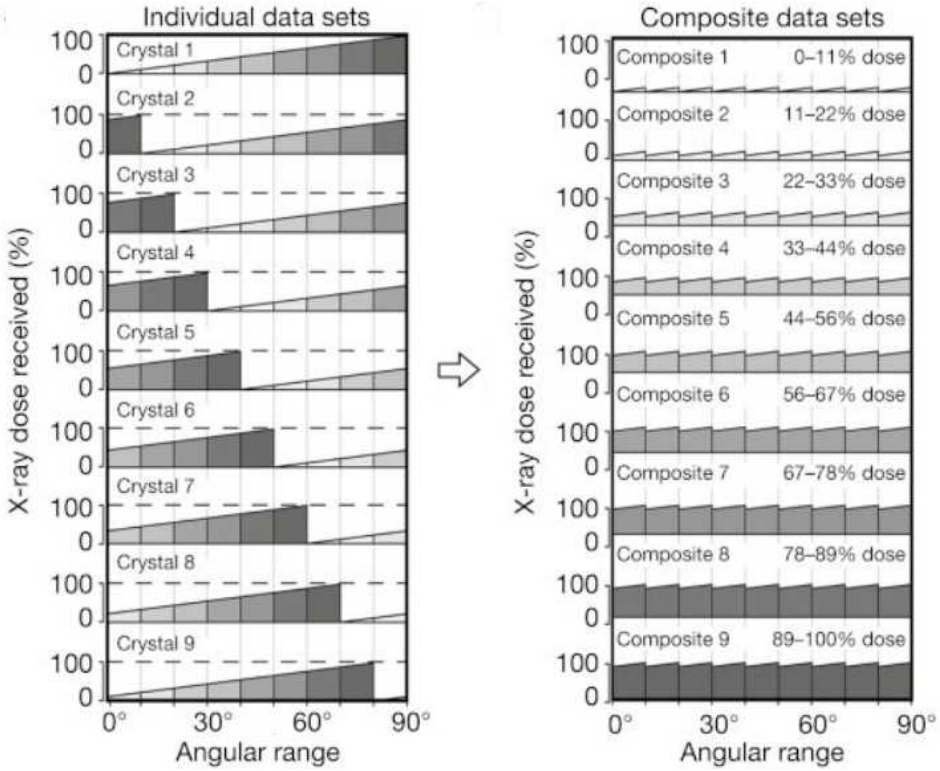


Fig. 1.2: An illustration of the multi-crystal data collection strategy[45].

Radiation-induced phasing (RIP)

Specific structural changes can be used for phasing in analogy to isomorphous replacement and anomalous dispersion using the same standard methods. Good phases can be obtained[47],[48] exploiting specific structural changes between two data sets from the same crystal that have been collected before and after an X-ray burn. Substructure determination is still more difficult than with common methods and so far tests have only been done on crystals of very good diffraction quality. Nevertheless the approach is promising. RIP signals can be large, much larger than those from sulfur SAD (single anomolous dispersion) phasing, and a more specific adaptation of the relevant software to this new method might leave much room for improvement. At the very least, RIP can yield additional information to a SAD experiment - it has been used to identify a correct MR (molecular replacement) solution[49] or a correct orientation of a ligand[50] - , but hopes are that it might on its own be a way for us to solve structures with a minimal amount of time and work invested (less than 15 min was possible in test cases!), since it requires only one native crystal.

1.6 Objectives of the project

In this work we investigated the potential of small molecule electron scavengers for the protection of metalloproteins from radiation damage effects, with a special focus on the metal's oxidation state.

This involves the following steps:

1. selection of potential scavengers and model metalloproteins
2. production of protein crystals in sufficient number and of reproducible quality
3. data collection from native and scavenger-soaked crystals
4. analysis of collected data at the level of global statistics and specific local structural changes
5. characterization of changes in the metal's oxidation state during an X-ray diffraction experiment through the use of a microspectrophotometer and evaluation of a scavenger's influence on the process

2 Materials

2.1 The proteins

Since some dropouts were to be expected, originally seven proteins were chosen for our studies. Requirements were as follows:

- the protein should be easily available, either commercially or through collaborations, but in any case
- ready to use for crystallization trials without much more time invested

furthermore:

- previously crystallized with known crystallization conditions and
- well diffracting crystals
- if possible a reasonable number of residues and monomers per unit cell which would not increase the workload unnecessarily
- a range of different metals in the active site

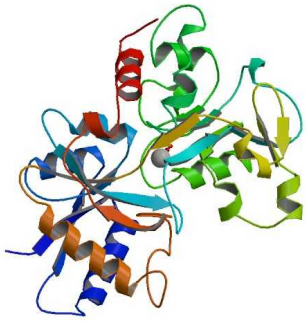
Table 2.1 gives an overview of the chosen proteins that are shortly characterized in the following.

Table 2.1: A summary of the initially chosen model proteins.

	Mw [Da]	**# amino acids**	**Swissprot accession #**	**ϵ280 [M^{-1} cm^{-1}]**	**Sigma Product #**	**PDB**
Azurin	14000	128 aa	P00282	8370		4azu
Cu,Zn superoxide dismutase	16000	152 aa	P00442	1400	S2515	1q0e
Insulin	5800	81 aa	P01317	5480	I0516	2a3g
Myoglobin	17000	153 aa	P68082	13940	M1882	1wla
Ni superoxide dismutase	13200	117 aa	P80734	16620		1q0d
Cytochrome c	12300	104 aa	P62894	10930	C2037	2b4z
Transferrin	77000	698 aa	P02787	78800	T4132	1a8e

Human transferrin (Tf)

Basically all of the iron in our blood is bound to transferrin (association constant 10^{22} M^{-1}) and in this way transported to its target location. Transferrin consists of two homologous globular domains (lobes) of 330 amino acids, each with one anion and iron-binding site in an interdomain cleft that opens and closes on iron uptake and release, respectively. The anion, usually a carbonate, is essential for metal binding, further coordination is provided by a histidine nitrogen, an aspartic acid carboxylate oxygen, and two tyrosine phenolate oxygens[51] (Fig. 2.1). The picture to the left shows a functional unit of transferrin, the N-terminal lobe. Holo-transferrin can be obtained from Sigma in the form of a lyophilized powder under the product number T4132.



Crystals have been reported to grow at 4°C in 40 mM sodium cacodylate buffer at pH 5.75 containing 20 mM sodium bicarbonate and 26% polyethylene glycol[52]. Holo-transferrin crystallizes in spacegroup $P2_12_12_1$ with one molecule in the asymmetric unit.

For crystallization trials the protein was dissolved in 20 mM TrisHCl pH 8.

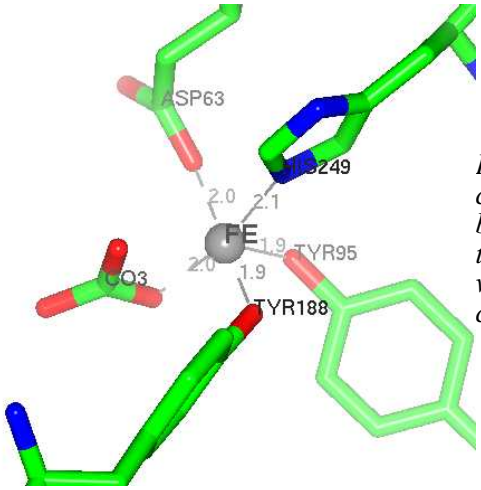
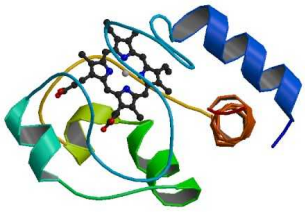


Fig. 2.1: Iron coordination in the binding pocket of transferrin (figure was created with ccp4mg[59].

Cytochrome c from bovine heart (Cytc)

The easy transition of the iron of cytochrome c between ferrous and ferric state makes it a widely used efficient catalyst and electron transporter essential in the respiration chain. The iron is coordinated axially by a histidine and a methionine residue (Fig. 2.2). The protein consists of a single chain of 104 amino acids with a covalently bound heme group



and is available from Sigma as a lyophilized powder (product number C 2037).

At 4°C and in conditions including 25-30% PEG1000 and 50 mM sodium phosphate (pH 7.0) Mirkin *et al.*[53] were able to grow crystals, optimizing them by microseeding. The structure was solved in space group $P2_12_12_1$ to 1.5 Å.

Crystallization trials were done with protein dissolved in 20 mM TrisHCl pH 7.6.

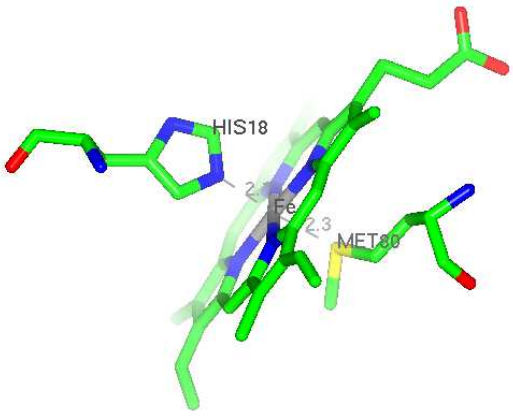
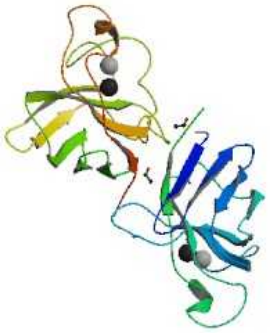


Fig. 2.2: Metal coordination in cytochrome c (figure was created with ccp4mg[59].

Cu, Zn superoxide dismutase from bovine erythrocytes (CuZnSOD)

Superoxide dismutases are important in protection from oxidative stress. They catalyze the following reaction



$$O_2^- + 2H^+ \rightarrow O_2 + H_2O_2$$

to reduce the load of highly reactive superoxide radicals (O_2^-). The family of superoxide dismutases is classified by the type of metal present in their active site. Apart from Cu and Zn containing superoxide dismutases, there are others containing manganese or iron. Recently another member with nickel in the active site has been discovered[54] (see below).

CuZnSOD is functional as a dimer, with an interesting asymmetry of the two active sites (see Fig. 2.3) that might be due to a mechanical communication between the dimers[55]. The catalytic mechanism, apart from CU(II) reduction, involves the switch from a five coordinate to a three coordinate copper[56].

CuZnSOD can be obtained from Sigma (product number S2515) as a lyophilized powder. It has been crystallized at 4°C through the sitting drop vapor diffusion method with a well solution of 15 % PEG-4000, 50 mM glycyl-glycine, 100 mM NaCl (pH 6.5). Crystals were of space group $P2_12_12_1$ and diffracted to 1.15 Å[56].

We dissolved the protein in 10 mM TrisHCl, pH 7.6 to a concentration of 10-15 mg/ml.

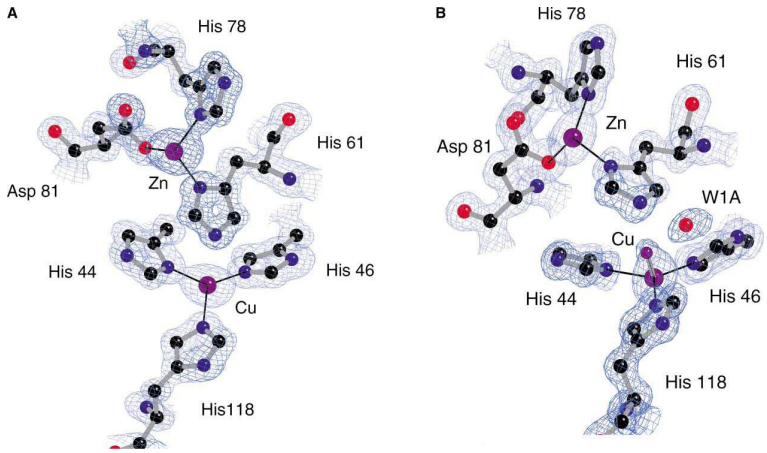
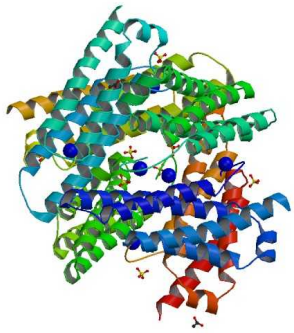


Fig. 2.3: Metal sites (reduced form) in the two monomers that make up a CuZnSOD functional unit[56].

Ni-containing superoxide dismutase from *Streptomyces seoulensis* (NiSOD)

The newest member of the family of superoxide dismutases (see also previous paragraph) was kindly provided by collaborators from South Korea in the lab of Sa-Ouk Kang, who expressed and purified NiSOD to a stage where only one final step of gel filtration was necessary to change the buffer and make it ready for crystallization[57].



NiSOD occurs as a homohexamer[54]. The Ni(III) in the active site is coordinated by the amino group of a histidine, an amide group, two cysteine thiolate groups (Fig. 2.4). An axially ligated histidine imidazolate is lost upon reduction[54].

Crystals have been reported to grow from 1.85 M ammonium sulfate, 0.1 M sodium acetate (pH 5.25), and 10% glycerol and from 2 M ammonium sulfate and 5% 2-propanol with two different unit cell parameters. Space group is $P2_12_12_1$ and diffraction was to 2.2 Å[54].

The protein we used for crystallization trials was prepared as a 10 mg/ml solution in 50 mM NaCl and 20 mM Tris pH 8.

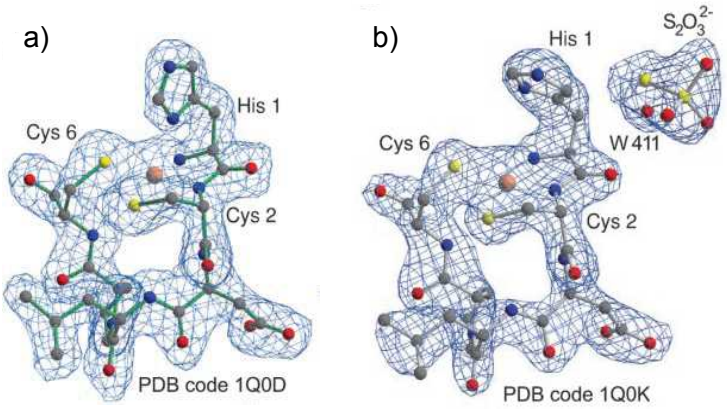
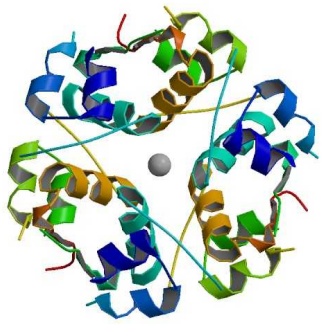


Fig. 2.4: The nickel binding site of NiSOD in oxidized (a) and reduced (b) state[54].

Insulin from bovine pancreas



The well-studied protein is an important regulator of glucose, amino acid and fatty acid uptake and storage as well as of the breakdown of glycogen, protein and fat. It comprises two short chains (21 and 30 amino acids) that are interconnected by two disulfide bridges. Hexameric insulin (T6 state as present in our crystals) is believed to be acting allosterically[58] and is equipped with two zinc ions, placed 15 Å apart along the three-fold symmetry axis. Each of the two ions is coordinated by three histidine residues (3x His10 from chain B) and three water molecules. A 10 mg/ml solution of insulin in 25 mM HEPES at pH 8.2 is available from Sigma (product number I0516).

Bovine insulin yielded rhombohedral crystals from solutions of 10 mM HCl, 7 mM zinc acetate, 50 mM sodium citrate and 17% acetone by a somewhat complicated procedure: *"The pH was raised to 8.0 to ensure complete dissolution of the bovine insulin and then reduced to 6.2. The solution was warmed to 323 K and allowed to cool to room temperature over a period of several days."* [58]. Space group was R3 (trigonal) and diffraction to 2.2 Å resolution could be recorded.

The purchased protein solution was diluted to 5 mg/ml for crystallization trials.

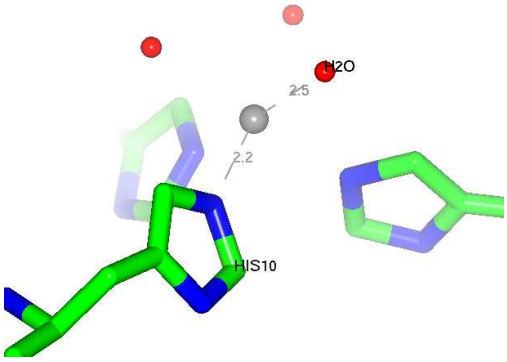
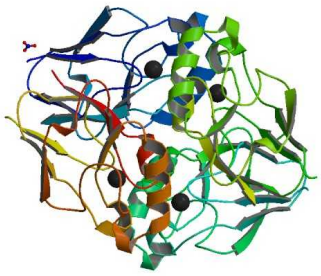


Fig. 2.5: Zinc coordination in bovine insulin (figure created with ccp4mg[59]).

Azurin from *Pseudomonas aeruginosa*

A member of the cupredoxin family, the periplasmic blue copper protein azurin is involved in bacterial electron transport. The 128 residues give a molecular weight of 14 kDa. In the active site of the monomer a copper ion is coordinated in a trigonal bipyramidal way by two histidines, a cysteine, a methionine and a backbone oxygen[60],[61] (Fig. 2.6). Reduction induces only slight changes in the geometry[62]. We received a ready-to-use purified sample of azurin (10 mg/ml) in potassium phosphate buffer pH 6.7 from Katsuko Sato and Chris Dennison (Institute for Cell and Molecular Biosciences, University of Newcastle upon Tyne).



Azurin has been crystallized from high concentrations of ammonium sulfate with added lithium nitrate in a buffer of sodium sulfate (pH 5.5-5.8)[63],[64]. The large plate-like crystals had space group $P2_12_12_1$ and diffracted to 1.85 Å resolution.

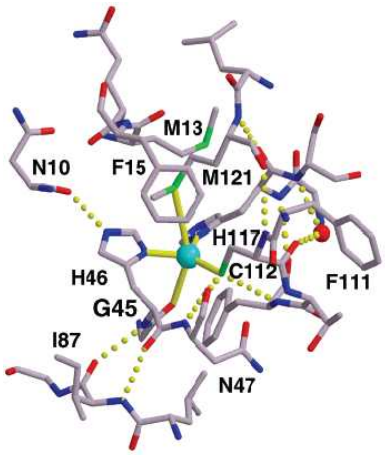
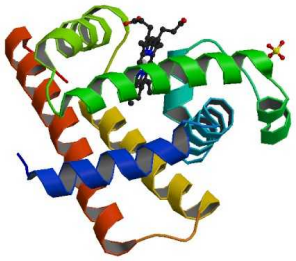


Fig. 2.6: Azurin's active site. Coordination by His117, His46, Cys112 and axially by Gly45 (carbonyl oxygen) and Met121[61].

Horse heart myoglobin

In the muscle, myoglobin has an important role in the transport of oxygen from the blood to the mitochondria. Consisting of a 153 residue polypeptide it holds a non-covalently bound heme molecule including an iron ion, which is coordinated by a histidine residue and a water (Fig. 2.7). Equine myoglobin can be bought as a dark brown lyophilized powder from Sigma under the product number M1882.



Rosetta-shaped crystals have been grown at room temperature from 2 M ammonium sulfate and 10.8% glycerol at a pH of 5.2 by the batch method[65]. They are in space group $P2_1$ and diffract to 1.35 Å resolution.

For crystallization trials the protein was dissolved in water at 10 mg/ml.

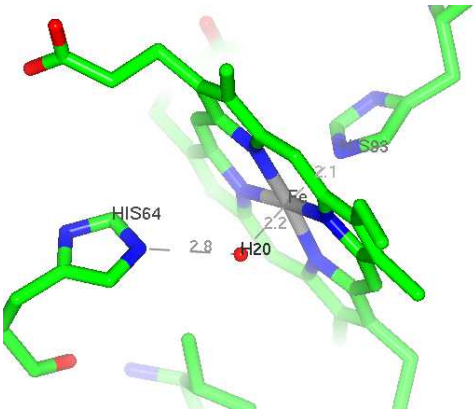


Fig. 2.7: The active site of myoglobin (figure was created with ccp4mg[59].

2.2 The scavengers

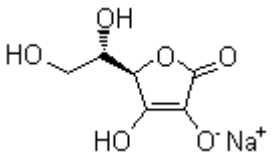
The selected scavengers fall into three groups:

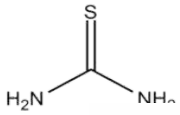
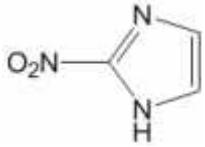
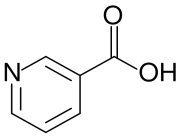
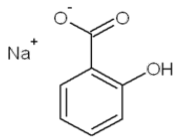
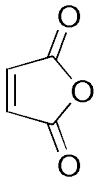
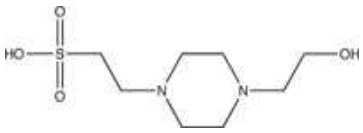
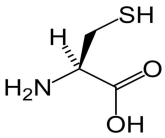
*Substances marked by * were kindly provided by Walter Schmid (Faculty of Chemistry, University Vienna).*

In parenthesis the acronyms are given that will be used to specify the scavenger throughout this text.

a) small organic molecules

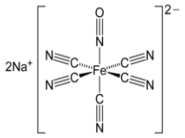
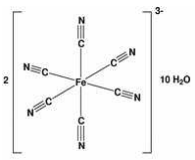
- (+)-sodium L-ascorbate (SAA)



- thiourea* (TU) 
- 2-nitroimidazole (NI) 
- nicotinic acid* (NA) 
- sodium salicylate* (NSA) 
- maleic anhydride* (MA) 
- HEPES - 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (hep) 
- L-cysteine (cys) 

b) transition metal compounds

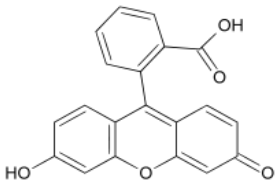
These compounds were included in our list, although the large absorption cross-section of the metals is by itself increasing X-ray absorption and might thereby counteract any beneficial effects.

- sodium nitroprusside dihydrate, $Na_2[Fe(CN)_5NO]*2H_2O$ * (NF) 
- potassium hexacyanoferrate, $K_3[Fe(CN)_6]$ * (KF) 

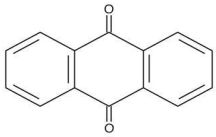
c) aromatic systems

The extended Pl-systems allow for easy uptake of electrons.

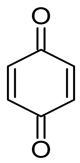
- fluorescein sodium salt (flu)



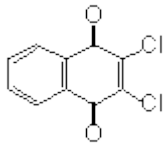
- anthraquinone* (AQ)



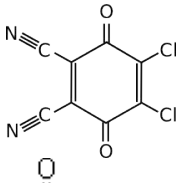
- p-benzoquinone* (BQ)



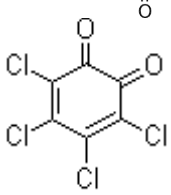
- 2,3-Dichloro-1,4-naphthoquinone* (DNQ)



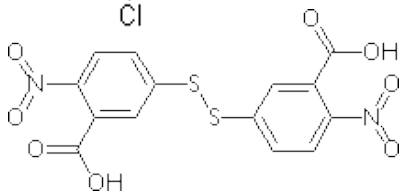
- 2,3-Dichloro-5,6-dicyano-p-benzoquinone* (DDQ)



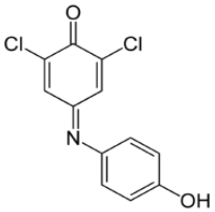
- 3,4,5,6-Tetrachloro-1,2-benzoquinone* (TBQ)



- 5,5'-Dithiobis(2-nitrobenzoic acid) (DTNB)



- 2,6-Dichlorophenolindophenol* (DPI)



2.3 Crystallization utensils

a) Crystallization screens

Commercial screens were used in early crystallization trials:

- JCSG (Qiagen)
- PACT 1 and 2 (Qiagen)
- Crystal Screen and Crystal Screen 2 (Hampton Research)
- Index Screen (Hampton Research)
- Peg/Ion Screen (Hampton Research)
- Grid Screen Ammonium Sulfate (Hampton Research): different ammonium sulfate concentrations in buffers of different pHs
- JBS6 (Jena Bioscience): ammonium sulfate screen slightly different from the Hampton one above
- Additive Screen (Hampton Research): 96 substances that are added in small amounts to previously determined crystallization conditions in an effort to optimize the crystals
- Stura Footprint Screen (Molecular dimensions): a screen designed to detect useful protein concentrations for crystallization

b) Crystallization plates

- 24 well sitting drop Linbro plates (Hampton Research and Crystalgen)
- 48 well hanging drop plates (Hampton Research and Axygen Biosciences)
- 96 well sitting drop (greiner bio-one) for use with the crystallization robot

c) Loops

- Mounted CryoLoops (Hampton Research and Molecular Dimensions)
- Magnetic CrystalCaps (Hampton Research) with bar codes to facilitate the use

Fig. 2.8: Loop, crystal caps and vials as provided by Hampton Research.

with automatic sample changers at synchrotrons (Fig. 2.8).

2.4 Buffers and solutions

Reducing agents

50 mM sodium dithionite (SDT)

100 mM β-mercaptoethanol (βME)

100 mM sodium thiosulfate (STS)

Buffer for size exclusion chromatography

Gel filtration buffer	
Tris-HCl pH 8	20 mM
NaCl	200 mM
β-mercaptoethanol	2 mM

Electrophoresis buffers

50 x TAE buffer	
Tris-Acetate	2 M
EDTA	50 mM
Acetic acid	1 M
pH 8.5	

SDS sample buffer (5 x)	
Tris pH 6.8	255 mM
EDTA	12.5 mM
SDS	5%
β-mercaptoethanol	700 mM
Glycerol	50%
Bromo phenol blue	0,35%

Coomassie Stain	
Methanol	50%
Acetic acid	10%
Coomassie Blue R250	0.25%

Destaining Solution	
Ethanol	30%
Acetic acid	10%
H_20 distilled	60%

2.5 Instruments

Rotating anode home source

To produce X-rays, electrons are accelerated in vacuum between a cathode and an anode. When they collide with the anode, X-rays are emitted. The spectrum of the emitted waves shows sharp peaks at wavelengths corresponding to transitions between the inner shells of the atoms of the anode material. For a copper anode the main peak is at 1.54 Å. However, most of the collision energy is converted to heat, which is why X-ray tubes with rotating anodes have been developed to allow more efficient cooling.

Fig. 2.9: An X8 Proteum installation (Bruker) as installed at the institute.

An X8 Proteum instrument produced by Bruker (Fig. 2.9), including a MICROSTAR rotation anode X-ray generator and PLATINUM 135 CCD detector was used for the data

collections carried out at the institute.

Beamlines

At the European Synchrotron Radiation Facility (ESRF) three beamlines (ID14-1 to 3) have been used. Each of them is equipped with a single phi spindle goniometer with motorised z-translation and x/y sample alignment translations, an Oxford Cryostream for 100K data collection, a high-speed CCD detector and an automatic sample changer.

A detailed characterization of the three beamlines is found in Table 2.2 and the layout of the ESRF synchrotron and a typical experimental station for macromolecular crystallography are depicted in Fig. 2.10.

Table 2.2: Details about macromolecular crystallography beamlines ID14-1 to 3 at the ESRF.

Beamline	Energy Range	Typical Beamsize [µm]	Resolution Limits	CCD detector	Maximum recommended phi speed	Typical exposure Time	Typical transmission for MAD/SAD
ID14-1 Fixed Energy	0.934Å (13.274keV)	100	0.95Å	ADSC Q210	0.3s/degree	1-5 sec	Usually 100%
ID14-2 Fixed Energy	0.933Å (13.289keV)	100	1.0Å	ADSC Q4	0.3s/degree	1-5 sec	Usually 100%
ID14-3 Fixed Energy	0.931Å (13.332keV)	100	1.0Å	ADSC Q4R	0.3s/degree	1-5 sec	Usually 100%

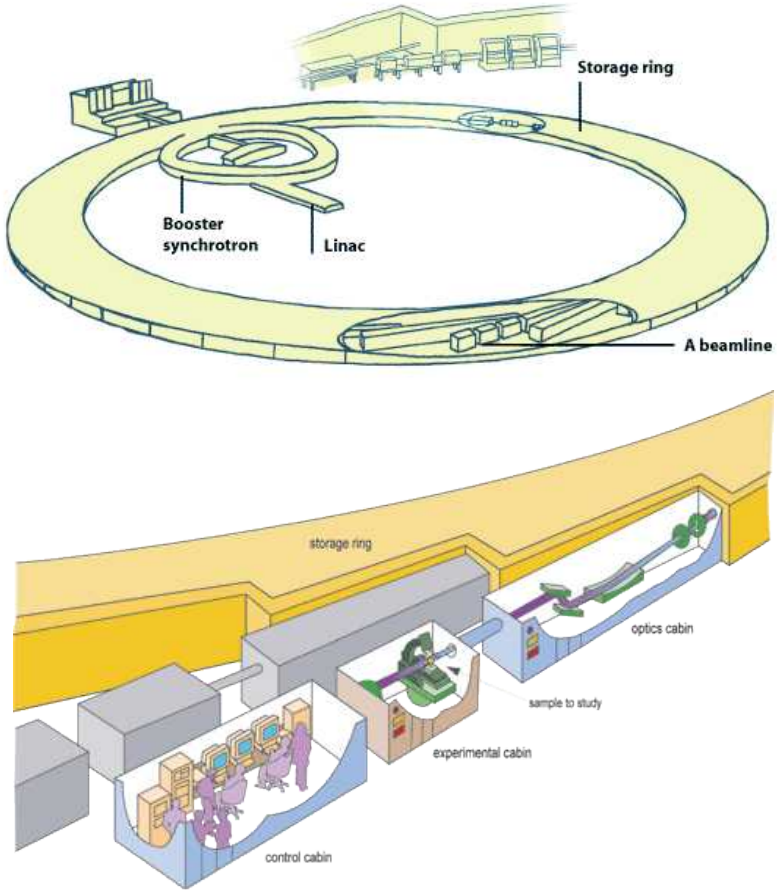


Fig. 2.10: Schematic views of the ESRf booster and storage ring (top) and a typical beamline for macromolecular X-ray crystallography (bottom).

At Bessy, we were using beamline MX 14.1 with specifications given in Table 2.3.

Table 2.3: Beamline specifications for MX-14.1 at BESSY.

Beamline	Energy Range [keV]	max. photon flux [phot/s/100mA]	Focus dim. hxv [µm]	Detector
MX-14.1	5 – 15.5	1.4×10^{11}	170x210	MAR225-MOSAIC

3 Methods

3.1 Protein purification

Size exclusion chromatography

Size exclusion chromatography, also known as gel filtration, permits the separation of particles by their apparent hydrodynamic radii. The sample is applied to a column filled with a porous material such as agarose or dextran. Smaller molecules are detained because they enter the pores while molecules that can enter the pores less deeply or not at all due to their size, pass through the column more quickly.

Gel filtration was employed as the final purification step for NiSOD (using a Superdex200 16/60 column by Amersham Bioscience), but it was also used for analytical purposes to evaluate the purity of proteins that would not crystallize (transferrin and cytochrome c, also Superdex200). Columns were run on an ÄKTA purifier.

Concentrating proteins

After gel filtration it is necessary to re-concentrate the protein solution, which we did by centrifugation. NiSOD was concentrated to 10 mg/ml in Centriprep Ultracel YM-3 centrifuge filter devices (Millipore) with a membrane holding proteins down to 3 kDa. The centrifuge was run at 3000 x g.

Azurin was also concentrated in this way when the protein buffer was changed in order to assess if it was preventing it from crystallizing. For azurin we used ultrafree MC 5,000 centrifugation filter devices (Amicon) that can be used in table centrifuges (we used 12.000 x g) and are more appropriate for the small volumes used.

Determination of protein concentration

Protein concentration was determined by absorbance measurements at 280 nm.

Gel electrophoresis

The proteins tested for purity were of small molecular weight (with the exception of transferrin) and therefore 15 or 18% SDS polyacrylamide gels were prepared according to the following recipes:

for 10 ml	15%	18%	5% (stacking)
H2O	2.3 ml	1.3 ml	1.4 ml
30% acrylamide mix	5 ml	6 ml	0.33 ml
1.5 M Tris pH 8.8 (pH 6.8 for stacking gel)	2.5 ml	2.5 ml	0.25 ml
10% SDS	0.1 ml	0.1 ml	0.02 ml
10% ammonium persulfate	0.1 ml	0.1 ml	0.02 ml
TEMED	0.004 ml	0.004 ml	0.002 ml

The gels were run at 150-200 V for 30 to 45 min and then stained with Coomassie blue.

3.2 Dynamic Light Scattering (DLS)

DLS is used for estimating the distribution of particle sizes in a sample in solution. The light that is scattered when the sample is illuminated by a monochromatic laser beam carries information about the velocities with which the particles in the sample move due to Brownian motion. An autocorrelation function is used to extract that information and then the velocity is easily translated into particle size.

The method is very useful to find out if a sample is monodisperse, i.e. present in only one oligomeric state. Monodispersity is a pre-requisite for successful crystallization.

Care has to be taken when working with particles that are known to be of elongated shape. Because Brownian motion is influenced by both shape and size, radii can not be determined correctly only by Dynamic Light Scattering.

3.3 Crystallization

That proteins – those large macromolecules of manifold shapes – can organize themselves into the regular arrangement of a crystal seems amazing. Despite the decade-

long eager research that does justice to the importance of this physical process for the determination of three dimensional structures, crystallization of proteins today is still to a large extent a matter of chance. In practice, this means extensive screening and – if one is lucky – optimizations of crystallization conditions.

The process of crystallization is influenced by protein, salt and detergent concentrations, by pH, temperature and pressure. While some proteins seem to crystallize under almost any conditions, those are the exception and mostly the success of an experiment is sensitive to even small variations in any of the involved factors.

Vapor diffusion

To convince protein molecules that are enjoying their freedom in solution to arrange themselves in a strictly determined pattern an energy barrier has to be overcome. This process becomes more favorable when a state of supersaturation can be reached in the solution. With some luck, the protein will then form well-ordered crystals instead of insoluble aggregates. While the batch method produces this state quite promptly – protein, buffer and a salt or detergent are simply mixed and everything is covered with oil – a more gradual approach to supersaturation is attempted by using dialysis against increasingly concentrated solutions. The method we used has been widely employed in the last years and is based on vapor diffusion. It involves mixing a drop of protein solution with a drop of the buffer/salt/detergent mix to be tested. A reservoir of this mix is provided and enclosed with the drop so that diffusion between the reservoir and the drop is possible. Since the mix in the drop has been diluted to some extent by the protein solution, diffusion of water from there to the reservoir will happen, gradually increasing the concentration in the drop. The experiment can be done with sitting or hanging drops, the difference and the general design of the experiment can be found in Fig. 3.1. We exclusively used the latter. Typical volumes for drops are 2 to 4 µl with a ratio of protein to reservoir of 2:1, 1:1 or 1:2. The reservoir contains 100-500 µl depending on the size of the well.

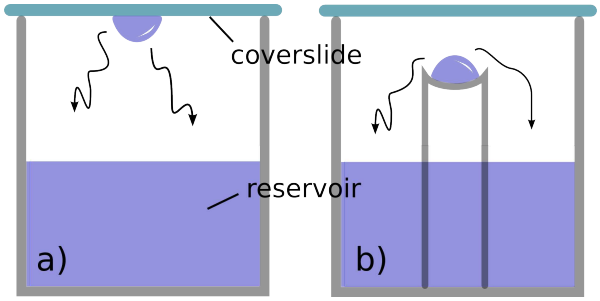


Fig. 3.1: Hanging and sitting drop vapor diffusion experimental setup.

Drop size influences the speed of equilibration. Larger drops are believed to increase the probability of nucleation. The time and effort that has to be invested in expression and purification of the protein, however, command an – as far as possible – economical usage

and setting of smaller drops.

Crystallization robot

At the start of a project crystallization robots save a lot of time and material, since they are able to set smaller drops (we commonly wanted 200 nl drops) than is manually possible, faster than it is manually possible. To identify crystallization conditions in the initial phase of our project a Cartesian nano-drop crystallization robot was used to set up screens. The robot was setting drops with ratios 1:1 and 2:1 of protein to reservoir solution.

Screening for the best conditions

If there is no previous knowledge of crystallization conditions, trials are usually started with incomplete factorial screens. Many of those are available commercially. Incomplete factorial simply means that they evenly sample the room of possible combinations of factors, because a complete coverage is not possible due to the exceeding number of factors. Complete factorials are used when it comes to optimization of promising conditions that have already yielded microcrystals or crystals in need of improvement. Then the components of the detected conditions are varied in steps of a certain size above and below their starting concentrations or starting pH.

Micro- and macroseeding

To obtain better, bigger crystals it can in some cases help to provide crystallization nuclei. This can be a smaller single crystal or a microcrystal (macroseeding) that is transferred to a new equilibrated drop. The other possibility is to touch an existing crystal/microcrystal with a hair or something similar and subsequently pass it through a freshly prepared and equilibrated drop (microseeding).

3.4 Soaking and freezing crystals

All diffraction experiments were done in a nitrogen stream at 100 K. This meant that crystals had to be fished from their mother liquor into nylon loops and quickly transferred to a prepared and previously tested cryosolution (usually the crystallization solution with an added cryoprotectant that will prevent formation of ice crystals). Potential scavengers had been dissolved into the cryosolution and should diffuse into the crystal during the short time needed to release the crystal into the cryosolution and regain it.

For experiments at home the loop was directly transferred onto the goniometer and into the nitrogen stream, for experiments at one of the synchrotrons it was plunged into liquid nitrogen to be stored at this temperature until the experiment.

Chemically reduced crystals

Chemically reduced crystals can serve as a reference when estimating structural changes in the active site during irradiation-induced reduction. For data collection crystals were soaked into cryosolution containing 50 mM sodium dithionite until they changed color (~ 1 min).

Reduced crystals were more important as a reference in the microspectrophotometry experiments. Unfortunately, sodium dithionite absorbs heavily in the interesting regions of the spectrum, which is why we had to use different reducing agents. 100 mM β-mercaptoethanol succeeded in visibly reducing azurin (color loss). Myoglobin, however, retained its red-brownish color. The crystal used in the experiment at the ESRF was soaked for 20 h in 100 mM sodium thiosulfate. The spectrum shows that reduction was not successful.

3.5 X-ray diffraction experiments

X-ray photons can be scattered elastically on electrons when they strike an atom. The repeating regular arrangement of molecules in a crystal is necessary to enhance this signal so that we are able to record it on a detector.

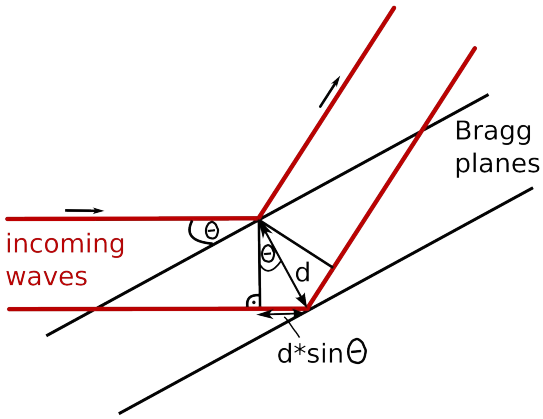


Fig. 3.2: Bragg's law. Only if $2d\sin\Theta$ is a multiple of λ will the waves scatter in phase.

The phenomenon of diffraction by a crystal can be described well as the reflection on a set of planes (Fig. 3.2). Only planes obeying the following rules can contribute to observable diffraction:

- the planes cut the axes of the unit cell into equal pieces without remainder, and as a consequence an atom that comes to lie on a plane in one unit cell will do so in every unit cell of the crystal
- the spacing between the planes obeys Bragg's law: $d = \frac{n\lambda}{2\sin(\Theta)}$ (see also Fig. 3.2), where d is the plane spacing, λ the wavelength, Θ the scattering direction and n an integer.

Waves that are scattered under different conditions will be canceled out. Spots that appear on the detector, on the other hand, carry information about the composition of the crystal. One set of Bragg planes produces one reflection on the screen, the intensity of the spot being proportional to the scattering factors of the atoms situated on or close to this set of planes. Recording the diffraction pattern of a crystal in different orientations can therefore reveal detailed structural information about the molecules in the crystal.

The closest spacing of Bragg planes that can still give measurable diffraction indicates the maximal resolution that can be obtained from a crystal.

Data collection

The same or similar procedure was employed at all the beamlines and instruments used in the course of the project. Rotation was about only one axis (phi) and the crystal was turned until a certain dose was reached. Redundant data sets were be collected to be afterwards split into five complete subsets of increasing dose.

The different fluxes and detectors at the various sites demanded different exposure times and numbers of images. Oscillation ranges were one degree for all examples except the myoglobin crystals collected at home.

Dose calculation

The ESRF provides a photodiode for their users and the local contact will help with the calibration of the beam. On some of the beamlines (ID14-1, ID14-2), the flux is written to a log file, providing the most accurate measure because it is recorded throughout the experiment. On the beamlines where this is not possible (ID14-3), the flux was determined before and after the experiment to calculate an average for the whole data collection.

At the home source, the number of photons can be counted by a special procedure aiming an attenuated beam directly onto the detector without a beamstop.

At BESSY no photodiode for calibration of the beam was available. The flux at BESSY was therefore approximated from the photon count in an upstream ion chamber.

The dose a crystal has taken can be calculated from knowledge about the beam (flux, dimensions, wavelength, shape) and the crystal sample (size, contents, density,

cryosolution contents) using *RADDOSE*, a program developed by Murray *et al*.[66]. An accurate estimate of the dose is difficult to obtain if the crystal is bigger than the beam in a direction perpendicular to the rotation axis. In this case parts of the crystal will enter and leave the beam during rotation, decreasing the total dose obtained and leading to a non-uniform distribution of dose. While we could think of ways to account for this, calculating how much of the crystal is in the beam at which times, this is of little use for our purposes. It reduces the amount of deposited dose per time significantly in an experiment in which beamtime is a limiting factor for us anyways.

The program is fed with details about the protein, the crystal, the cryosolution and the diffraction experiment as mentioned at the beginning of the previous paragraph. On the basis of that it calculates the dose that will be absorbed by the crystal and the time that will be needed to reach the Henderson limit under those conditions, as well as the percentage of the crystal that will be in the beam and the percentage of the beam that will be seen by the crystal on average.

3.6 Integration and Scaling

XDS package

The next step after collecting images from the crystal is to find the reflection spots, index them, which means assigning them to the respective Bragg plane they stem from, and evaluate their intensities. Several programs have been developed for that purpose, the one we used (with the exception of the data collected at home, see below) was *XDS,* developed by Wolfgang Kabsch[67]. The program does not employ a user interface and is controlled by text input files. It was run with default settings except for the merging of Friedel's pairs, which we turned off because we wanted these differences to remain for the subsequent calculation of R_d values through XDSSTAT[14] (for further information see also the introduction, chapter 1.3).

We ran XDS on each of the five subsets and on the complete data set of each crystal. Integration and correction was done in space group P1, then the correct space group and reindexing card as obtained from IDXREF.LP were filled into the input file and the CORRECT step alone repeated.

After integration and assignment of the correct space group and unit cell a scaling step refines the result and merges symmetry related reflections and Friedel pairs to obtain a set of unique indexed reflections. A resolution limit was set in the scaling step according to the CORRECT output. We did scaling mainly with Kabsch's XSCALE[67] and used the

output statistics for the analyses and comparison of data sets. Also part of the XDS package, XDSCONV provides a pathway to the CCP4 package[a,68], allowing us to convert XDS *.hkl or *.ahkl format to *.mtz.

Proteum software

Data collected at the home source was integrated and scaled with the software provided with the diffractometer. Data was cut to the appropriate resolution limit and converted to scalepack format using XPREP.

3.7 Reindexing

For myoglobin and insulin it was important to ensure the same indexing scheme for all subsets of one crystal, which was done with the combined help of POINTLESS[69] and REINDEX (both CCP4 suite[68]).

3.8 Detwinning

Insulin made the use of TRUNCATE[70], SFCHECK[71] and DETWIN[72] necessary, when we noticed that some of the crystals were merohedrally twinned. Merohedral twinning is a special form of twinning where different crystal lattices overlap three-dimensionally[72]. The molecules of the two lattices are not related by a symmetry operation belonging to the relevant space group. The risk of having merohedrally twinned crystal is higher, when two or more crystal axes are of similar or equal length, which is for example a space group requirement for H3. Merohedral twinning does not manifest itself by clearly overlapping or elongated spots, while at the same time making the solution of the structure difficult to impossible, depending on the twinning fraction. If the two lattices are of approximately equal intensity, detwinning is not possible.

TRUNCATE and SFCHECK yield statistics that help to detect or confirm twinning (cumulative intensity distributions and twinning fractions obtained from observed and calculated reflections), DETWIN attempts to do what its name proclaims.

3.9 Molecular replacement and refinement

To solve a structure, information about intensity and phases of the diffracted X-rays is necessary. While the former is easily read from the detector, the latter is more difficult to obtain. Several methods exist: multiple isomorphous replacement (MIR) that depends on

a CCP4 stands for Collaborative Computational Project Number 4, and is a comprehensive collection of software used in the field of macromolecular crystallography. (http://www.ccp4.ac.uk)

heavy atom derivatives of the protein, single or multiple anomalous dispersion (SAD/MAD) that exploit the anomalous signal of heavy atoms (sulfurs or metals) and molecular replacement (MR). All of them can be used to obtain preliminary phases that have to be close enough to the real ones to serve as the starting point for iterative cycles of model building and refinement.

Molecular replacement uses phases of homologous proteins as a starting point. It is a widely used method because it can succeed even if sequence identity is low. Structures for all of our investigated proteins are available in the PDB (Protein Data Bank, www.rcsb.org) and so structure solution was possible by molecular replacement using MOLREP[73], a program included in the CCP4 package[68]. For azurin this proved to be a dispensable step, our data being already so close to the published structure that a few cycles of refinement by REFMAC5[74] were enough to get started. Refinement of 3D models was then done in alternate turns of restrained refinement using REFMAC5 and manual rebuilding in COOT[75].

3.10 Calculating maps and peak search

After refinement of the first (low dose) data set, CAD[68] was employed to create a single mtz-file containing observed structure factors (Fobs) and corresponding standard variations (sigFobs) for all the subsets and the calculated phases of this refined first model. FFT[76] was used to calculate Fo_x-Fo_1 ($x \in \{2,3,4,5\}$) difference maps between the first and all the other subsets, while at the same time outputting all peaks above 3 sigma found in these maps. Afterwards with the help of CONTACT[68] those peaks could be filtered in a way that only those closer than 1 Å to a protein atom would be selected.

3.11 Single crystal UV-vis microspectrophotometry

UV-visible microspectrophotometry allows the recording of spectroscopic features directly from a crystal. It is especially suited as a complement to crystallographic data where spectroscopically discernible reaction intermediates are involved. The oxidation state of a metal can not be deduced from its appearance in an electron density map derived from X-ray diffraction data, which makes microspectrophotometry a valuable complementary experiment for structural studies on metalloproteins.

An online microspectrophotometer can be mounted at beamline ID14-2 at the ESRF (Fig. 3.3). It makes the simultaneous irradiation of the crystal and measurement of spectra possible. In this way a time course of the whole spectrum can be inspected. The focus of the photometer is far smaller than that of the X-ray beam, assuring that only

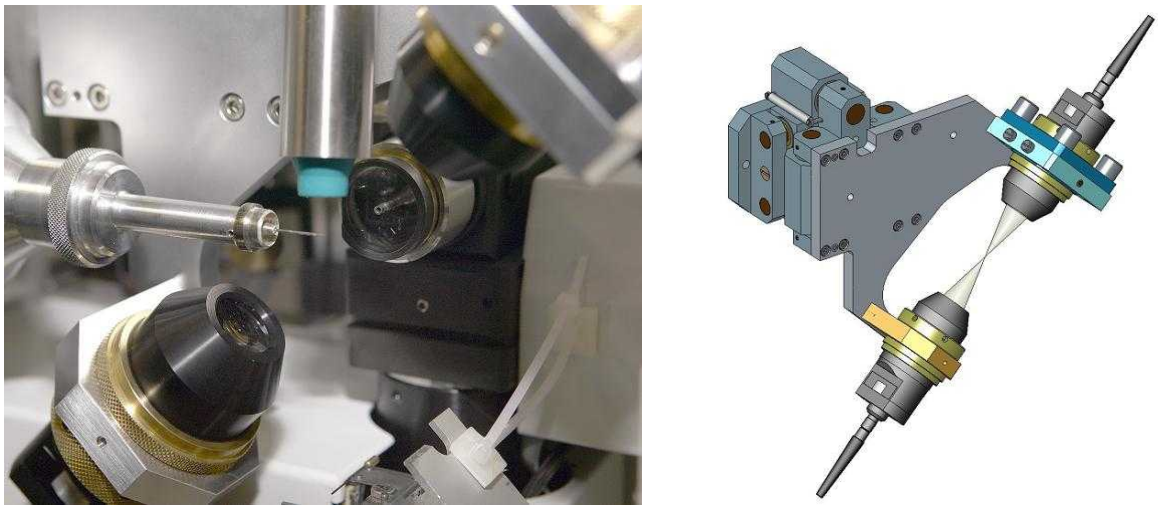


Fig. 3.3: The microspectrophotometer setup at ESRF beamline ID14-2[77].

irradiated area contributes to the recorded spectrum.

There are some caveats to the method though, that mainly stem from the inherent nature of a crystal. While in solution all protein orientations are present in equally large populations, in the crystal only a few orientations determined by symmetry are possible. As a consequence, crystals are anisotropic and their spectral features vary with the direction of the incident light. Furthermore, crystal prism effects can distort the spectrum. Another complication is the uncertainty of crystal dimensions, which can not be measured accurately. Unlike spectroscopic measurements of solutions in a cuvette, the path length the light travels is therefore not accurately known, not in the crystal itself and even less in the surrounding cryosolution.

It is therefore clear that absolute peak heights can not be compared between different crystals under these circumstances, and only relative changes in the spectrum are exploited for analysis.

Simultaneous data collection and recording of spectra is not advised, since this requires the rotation of the crystal, which involves constant change of the path length. Experimenters have to choose: either to do the two measurements alternately, e.g. go back to the original orientation and take a spectrum after every 10 images, or to do them in separate experiments. We went for the latter option for time considerations, since beamtime with the microspectrophotometer was limited and this approach allowed to test more samples.

In our experiments we used two lamps (tungsten halogen and deuterium) to cover wavelength ranges between 200 and 1000 nm.

Our procedure: in preparation for the experiment at the synchrotron, solution spectra of the proteins in oxidized and reduced state were measured to determine the characteristic spectrum and which changes to monitor.

For an optimal result, at the beamline we always carefully adjusted the orientation of each crystal to get a maximally pronounced peak with a minimal basal absorption. Then the spectrophotometer was turned on and left for about 30 sec to make sure that the measured absorption was stable before opening the shutter for the X-ray beam. One spectrum per second was recorded (10 images/sec were taken and averaged to smoothen the spectrum). At an average flux of 10^{11} photons per second the target dose was reached after about 20 min. By then also the changes in the spectra had become marginal.

4 Results and Discussion

4.1 Purification and quality check

NiSOD purification

NiSOD was eluted in a nice symmetric peak from the gel filtration column and fractions from the peak corresponding to the hexameric form were loaded onto an SDS PAGE gel to confirm that it was pure and of the expected size of one monomer. Only very minor peaks are present at a distance to the main peak.

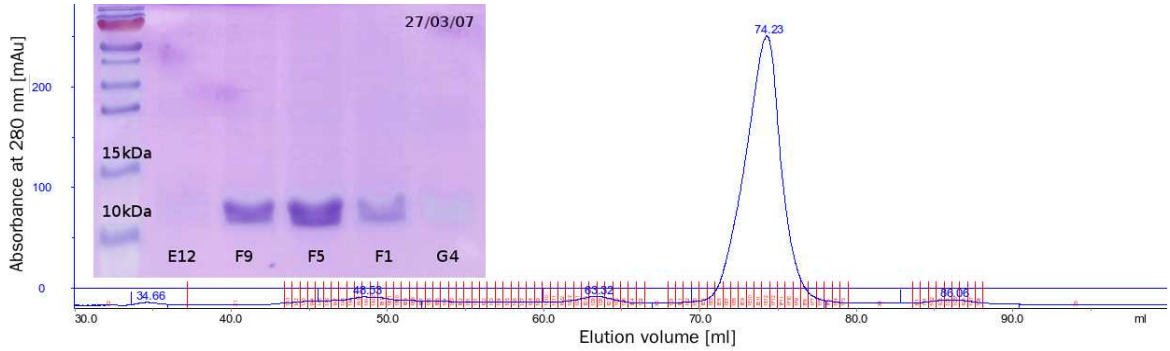


Fig. 4.1: NiSOD purification: sample fractions (10 µl) of the small peak to the left (E12), the main peak (F9, F5, F1) and the small peak to the right (G4) were loaded onto the gel.

Purity check for transferrin and cytochrome c

Gel filtration chromatography together with SDS PAGE gels of peak fractions was used to check the purity of transferrin and cytochrome c.

Transferrin

A smaller peak is present with bigger particle size then the main peak. The fractions taken from this first peak are of low concentration, but we see weak bands on the gel that indicate it contains particles of the same molecular size as the main peak, probably assemblies of more than one monomer.

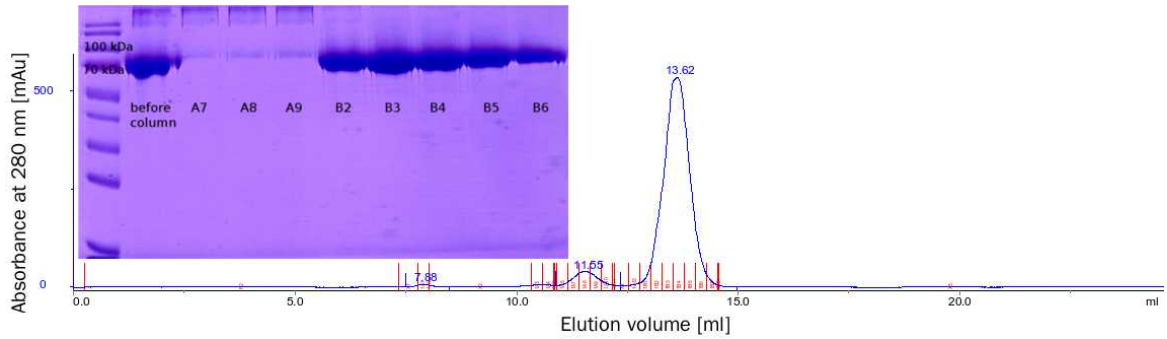


Fig. 4.2: Transferrin analytical gel filtration: lanes A7-A9 on the gel show fractions from the first smaller peak eluted after 11 ml, the other lanes are from the main peak.

Cytochrome c

10 µl samples of the main peak were loaded onto the gel. There is no contamination.

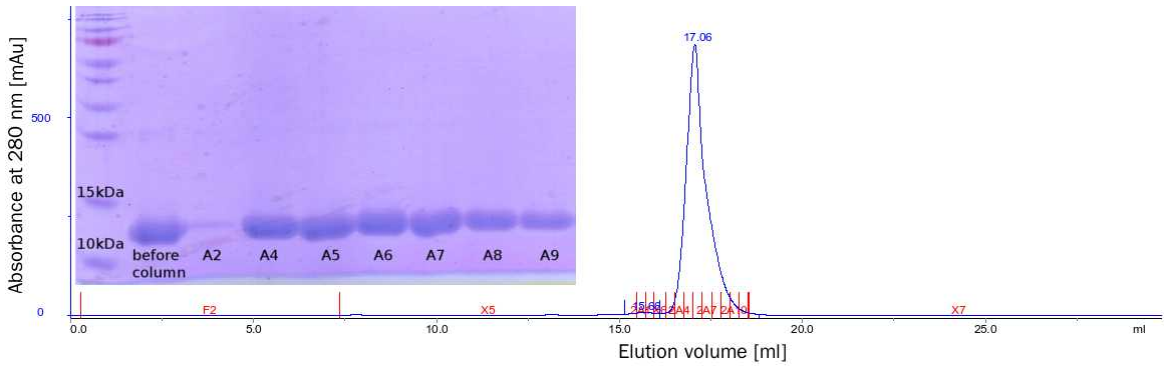
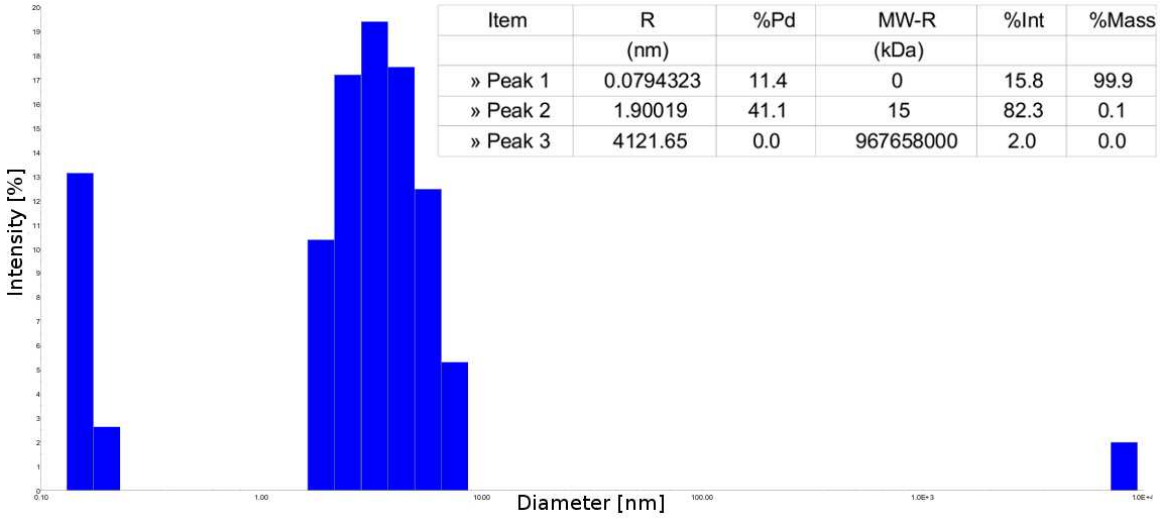


Fig. 4.3: Cytochrome c analytical gelfiltration.

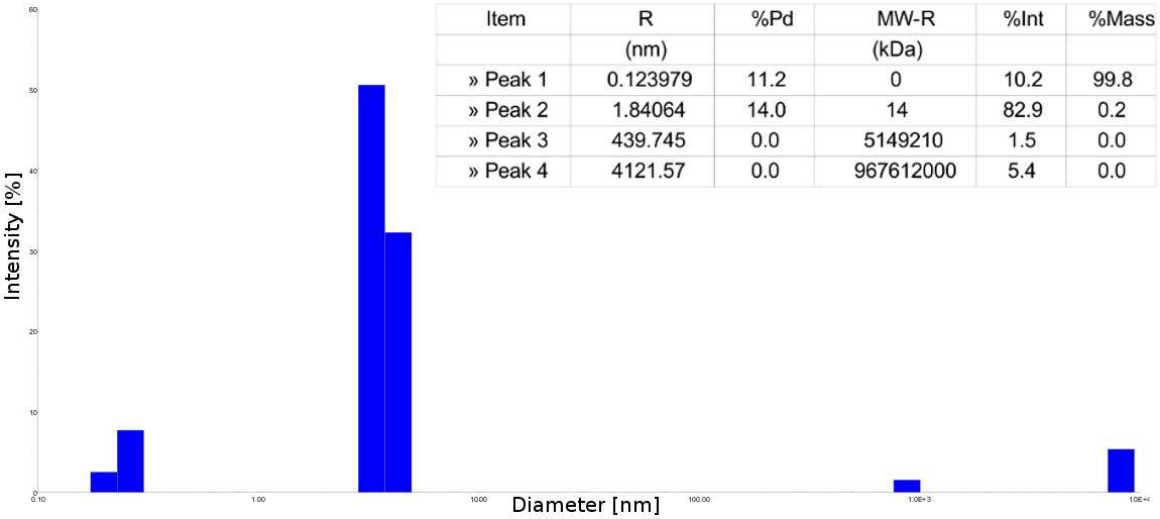
4.2 Dynamic Light Scattering (DLS)

We performed DLS measurements of cytochrome c samples before and after the analytical gel filtration to check whether poor monodispersity was the source of crystallization problems. While a clearly monodisperse peak is present for the unfiltered sample, the sample after gel filtration shows a much broadened peak and more imprecise estimation of particle size. There is thus no need to gel-filtrate. In our case the protein might even have suffered in the column.



Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
» Peak 1	0.0794323	11.4	0	15.8	99.9
» Peak 2	1.90019	41.1	15	82.3	0.1
» Peak 3	4121.65	0.0	967658000	2.0	0.0

Fig. 4.4: Particle size distribution after size exclusion column in a 1 mg/ml sample of cytochrome c in 20 mM TrisHCl pH 8. Acquisition time 10 sec.



Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
» Peak 1	0.123979	11.2	0	10.2	99.8
» Peak 2	1.84064	14.0	14	82.9	0.2
» Peak 3	439.745	0.0	5149210	1.5	0.0
» Peak 4	4121.57	0.0	967612000	5.4	0.0

Fig. 4.5: Particle size distribution in a 1 mg/ml sample of cytochrome c in 20 mM TrisHCl pH 8. Acquisition time 10 sec.

4.3 Crystallization

Despite our efforts we were not able to crystallize cytochrome c. The protein seems to be very stable since it did not precipitate even from solutions of 100 mg/ml. Transferrin was also dropped, even though some crystal clusters were obtained.

Azurin

At first azurin resisted crystallization in screens designed on the basis of previously described conditions[78], producing only bubble-like blue shapes (see Fig. 4.6). Only a considerable increase of ammonium sulfate and $LiNO_3$ concentrations lead to success. Crystals grow within one day in the following conditions:

3.7-3.9 M ammonium sulfate

1 M $LiNO_3$

100 mM sodium acetate, pH 5.8

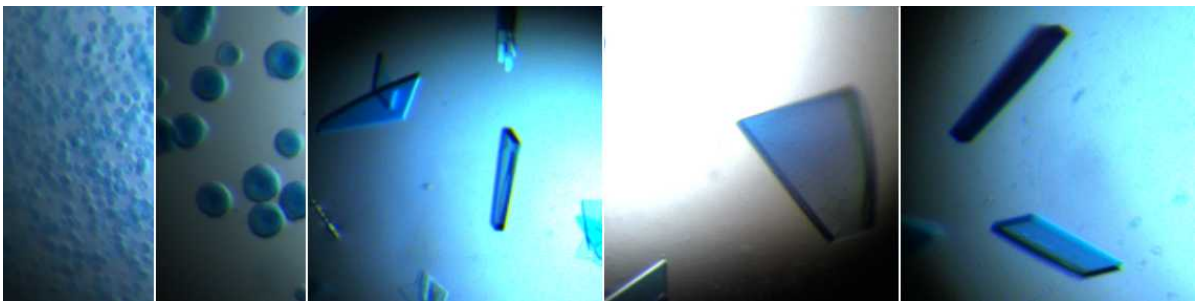


Fig. 4.6: Azurin crystal pre-stages and final crystals.

The protein sample used for the hanging drop experiments contained 10 mg/ml of azurin in potassium phosphate buffer at pH 6.7.

Due to the high salt content the solution was already cryoprotected. The crystals grow as large blue plates of usually about 200 x 150 µm and 10 to 40 µm thickness. Often they were single, but rosetta-like clusters also occurred.

The ammonium sulfate concentrations are close to the solubility limit and salt crystals tend to grow nearly as soon as the well is opened, which makes fishing the crystals for data collection difficult.

CuZnSOD

Cu,Zn superoxide dismutase was crystallized from a solution of 10 or 15 mg/ml of protein under the following conditions (found and optimized by summer student Leo Geist):

17% PEG 4000

100 mM MES, pH 5.5

3% isopropanol

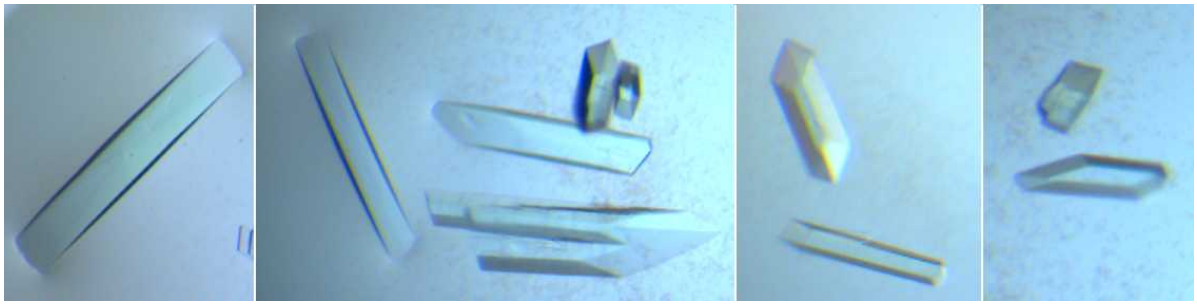


Fig. 4.7: Cu,Zn superoxide dismutase crystals.

18% glycerol were added for cryoprotection. Crystals grow within one week up to 1 mm in the longest direction, sometimes thinner and plate-like as on the image on the left, sometimes more robust like the two images on the right of Fig. 4.7.

Insulin

Crystallization conditions containing zinc could be established relatively fast on the basis of conditions found in the literature[79]. The protein was used at 5 mg/ml in 25 mM HEPES pH 8.2 and crystallized in hanging drops mixed 1:1 with the following:

5 mM Zn acetate

8-10 mM HCl

50 mM sodium citrate

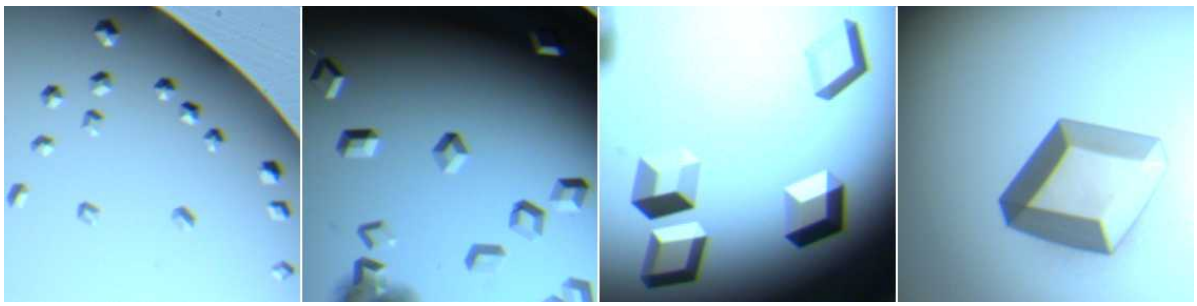


Fig. 4.8: Insulin crystals of different sizes.

30% glycerol were added for cryoprotection. Colorless rhombohedral crystals occur after two days and continue to grow for two weeks to final dimensions with edge lengths between 60 and 200 µm.

Myoglobin

Crystal growth is relatively unreliable for myoglobin. Through Hampton's additive screen it was found that small amounts of PEG 3550 could turn the previously exclusively observed bushy arrangements of needles into rosettes of thin plates. The polyethylene glycol was found not to mix well with the rest of the solution and so all the drops were set by adding 0.4 µl of PEG directly to a 2:2 drop rather than mixing it with the reservoir solution that consisted of:

3-3.2 M ammonium sulfate

100 mM TrisHCl, pH 7.2-7.4

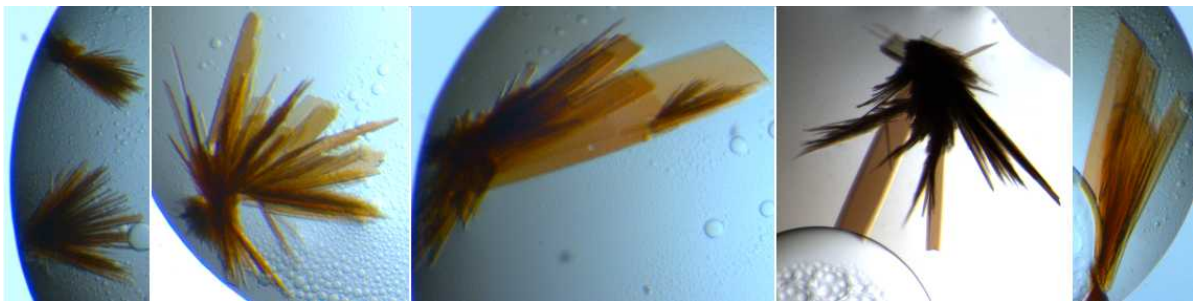


Fig. 4.9: Typical myoglobin crystal clusters.

Crystals grow at 22°C within a few days. For the cryosolution 5% glycerol was added to the reservoir solution. In the drop PEG formed a separate phase, which often seemed to become a point of nucleation for crystals. Additionally it might be responsible for the skin that covered the drops after a few weeks. The plates were very thin (mostly < 10 µm) and often difficult to separate. Only occasionally crystals were thick enough to give reasonable diffraction.

NiSOD

NiSOD crystallized in two different conditions at 22°C:

a)

1.8-2 M ammonium sulfate

3-5 % isopropanol

b)

1.7 M ammonium sulfate

100 mM sodium acetate

100 mM of a buffer system described by Janet Newman[80], comprised of succinic acid, NaH_2PO_4 and glycine at pH 5-5.5

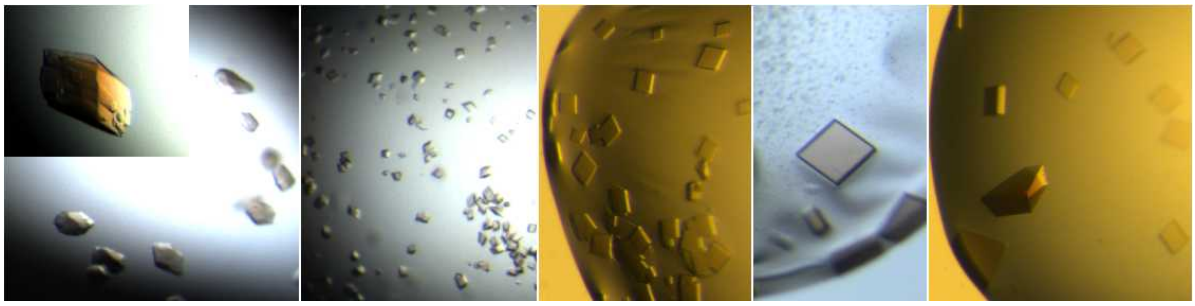


Fig. 4.10: NiSOD crystals grown in the presence of sodium acetate (left image plus small one) and isopropanol (four images to the right), respectively.

20% glycerol were added to prepare the cryosolution. Crystals from the two conditions have different forms. While the crystals grown with isopropanol had sharp edges and even faces, those grown in the presence of sodium acetate were sometimes larger. In both cases, however, crystal growth was very unreliable. Macro- and microseeding were attempted, but did not improve the outcome.

Test sets have been collected for NiSOD (isopropanol type) diffracting to 2.3 Å, but the crystals proved very fragile and were therefore abandoned to concentrate on more promising crystals of other proteins.

4.4 Scavenger solubility test

In the long run, we would like to provide the crystallographic community with a kind of recipe, which scavenger might be useful and how it can be applied. One obstacle that has to be coped with are high concentrations of PEGs, glycerol or salt that decrease the solubility of a scavenger. There is still too little data in the literature to predict the minimum useful concentration for a scavenger and there might not be a general answer to this question. Therefore we aimed at scavenger concentrations of 200 mM, as this had been previously shown to work for ascorbate[40]. Such concentrations could not be reached for the quinones in any of the tested cryosolutions and also some of the other substances had lower solubility limits in some or all of the solutions. When 200 mM could not be

reached, further dilution was tried and if the substance could not be dissolved at 50 mM experiments were done with a saturated solution.

Solubility was tested in the four cryosolutions we used in our experiments. Additional tests with a subfraction of the better-soluble scavengers in selected solutions of common cryo-protectants showed it is mostly higher salt concentrations that diminish solubility, PEGs are less perturbing. Not surprisingly, solubility was best in the insulin cryosolution, which contains only minor salt concentrations plus 25% of PEG 400. Details of the solubility tests are found in Table 4.1.

Table 4.1: Scavenger solubility in the cryosolutions used for azurin, CuZnSOD, insulin and myoglobin.
Concentrations are in M. The color bar at the right side should give an impression of the color of the solution, which is always of different intensities in the different buffers indicating the different solubility. Abbreviations used for the scavengers are defined in chapter 2.2.

	Azurin	CuZnSOD	Insulin	Myoglobin
KF	0.1	0.3	0.2	0.15
NF	0.2	0.3	0.2	0.2
Hep	0.2	0.3	0.2	0.2
TU	0.2	0.3	0.2	0.2
NSA	0.2	0.3	0.2	0.2
Flu	slurry	0.1	0.2	slurry
NA	0.04	<0.05	0.2	<0.05
NI	<0.05	<0.05	0.05	<0.05, color
C	0.2	0.3	0.2	0.2
SAA	0.2	0.3	0.2	0.2
MA	0.1	0.3	0.2	0.2
DT	<0.05	color	0.2	<0.05
AQ	<0.05	<0.05	<0.05	<0.05
TBQ	<0.05	<0.05, color, not clear	<0.05	<0.05
DNQ	<0.05	<0.05	<0.05	<0.05
DPI	<0.05, v. slight color	strong color, xtal not visible	0.01	<0.05, slight color
BQ	<0.05, color	<0.05, color	0.05	<0.05, floating stuff
DDQ	<0.05, color	strong color	0.05 (strong color)	<0.05, color

Fluorescein is very insoluble in higher ammonium sulfate concentrations. It turns the solution slurry and sticky and can therefore not be used with myoglobin or azurin. Also BQ is not optimal, although clearly soluble to some extent as evidenced by the color change of the solution, it is impossible to get rid of small floating particles even by very thorough down-spinning.

4.5 Soaking, freezing, data collection

In general the crystals did not show any signs of suffering when they were soaked into the scavenger-cryosolution. In the few cases where crystals showed cracking signs, this was not reproduced by the next crystal of the same type soaked into the same scavenger. The crystals we tested in the X-ray beam did not show any sign of serious degradation directly due to the scavenger. Quality was comparable between native and scavenger-soaked crystals of azurin, myoglobin, insulin or CuZnSOD. NiSOD and cytochrome c, however, were more problematic.

Exclusion of NiSOD and CuZnSOD

NiSOD was very fragile in itself and often started to crack already in the cryosolution without any scavenger. Diffraction was also not especially good, which in addition to the difficult reproducibility made us concentrate more on the other proteins. The situation was even worse for cytochrome c as our collaborators from the US reported to us. All of the scavengers seemed to destroy the crystals' diffraction capability while the native crystals diffracted as usual.

This left us with four proteins for our further studies. The first data sets we collected were of CuZnSOD and insulin because those were the first crystals available.

CuZnSOD crystals were found belonging to three different space groups, one of them – possessing one rather long cell axis – prevailed in crystal trays that were prepared later. These crystals proved to be difficult and often impossible to index and so CuZnSOD had to be set aside as well.

Data sets that do meet all the requirements – comparable quality (within crystals of one protein), assignable dose and presence of the metal – have been collected from azurin, insulin and myoglobin. All the crystals and specifics about data collection are listed in Table 4.2. Further information like oscillation range, exposure time and crystal size is listed in detail in Appendix 7.1.

Beam and crystal size

Beam size chosen was always the biggest possible for the respective beam line, since it determines the size limit for the crystal. Azurin crystals were mostly growing as large plates and the almost immediate growth of ammonium sulfate crystals made any attempts to cut them into smaller pieces useless. Plates therefore had to be screened for the smallest crystals, which were mostly also the thinner ones and very difficult to fish from

their mother liquor. It was therefore often not possible to avoid small ammonium sulfate contamination in the loop and salt spots in the diffraction pattern. This was a problem especially in the four crystals collected at the second date (ID14-1), but as the salt spots are very prominent the software is able to identify and exclude them without difficulty. Statistics are comparable to the other crystals.

Table 4.2: Data collection specifications. The specified resolution is the resolution limit at the edge of the detector.

		Facility	Beamline	Beamsize [mm²]	Flux [ph/sec]	Wave-length [Å]	Energy [keV]	Reso-lution [Å]	End Dose [MGy]
Azurin	HEP	ESRF	ID14-3	0.12x0.12	7.7E+10	0.93	13.32	1.9	10
	Native	ESRF	ID14-3	0.12x0.12	6.5E+10	0.93	13.32	1.9	8.4
	NI	ESRF	ID14-3	0.12x0.12	7.3E+10	0.93	13.32	1.9	9.5
	SAA	ESRF	ID14-3	0.12x0.12	5.6E+10	0.93	13.32	1.9	7.3
	TU	ESRF	ID14-3	0.12x0.12	7.2E+10	0.93	13.32	1.9	9.6
	Native	ESRF	ID14-1	0.1x0.1	3.2E+10	0.93	13.28	1.8	6.5
	Cys	ESRF	ID14-1	0.12x0.12	7.1E+10	0.93	13.28	1.8	11
	DNQ	ESRF	ID14-1	0.12x0.12	5.9E+10	0.93	13.28	1.8	9
	NA	ESRF	ID14-1	0.12x0.12	6.0E+10	0.93	13.28	1.8	9
Myoglobin	Native	ESRF	ID14-3	0.1x0.1	2.6E+10	0.93	13.32	1.8	5.7
	NSA	ESRF	ID14-3	0.12x0.12	3.9E+10	0.93	13.32	1.8	6.1
	Native	Home	Proteum X8	0.22x0.22	6.5E+8	1.54	8.05	2.2	7
	Native	Home	Proteum X8	0.22x0.22	6.5E+8	1.54	8.05	2.2	7
	NSA	Home	Proteum X8	0.22x0.22	6.5E+8	1.54	8.05	2.2	7
CuZnSOD	Native	ESRF	ID14-3	0.12x0.12	6.3E+10	0.93	13.32	2.1	8.2
Insulin	Native	ESRF	ID14-2	0.1x0.1	2.4E+10	0.93	13.29	1.8	7.6
	NSA	ESRF	ID14-2	0.1x0.1	2.9E+10	0.93	13.29	1.6	5.9
	Cys	BESSY	BL14.1	0.17x0.23	6.4E+10	0.92	13.51	1.4	4.5
	DNQ	BESSY	BL14.2	0.17x0.23	5.6E+10	0.92	13.51	1.4	4.5
	Flu	BESSY	BL14.3	0.17x0.23	8.2E+10	0.92	13.51	1.4	5.1
	NA	BESSY	BL14.4	0.17x0.23	5.7E+10	0.92	13.51	1.4	4.4
	Native	BESSY	BL14.5	0.17x0.23	5.3E+10	0.92	13.51	1.4	7
	NF	BESSY	BL14.6	0.17x0.23	7.7E+10	0.92	13.51	1.4	5.8
	TU	BESSY	BL14.7	0.17x0.23	7.7E+10	0.92	13.51	1.4	6.2

Another problem caused by the need to use smaller crystals is the higher tendency for anisotropy as a result of a slight bending of the crystal. Many azurin crystals had to be discarded because their patterns showed elongated or double spots, which was not the case for big crystals. The problem of anisotropy was more important in myoglobin crystals which were in general longer and thinner with a tendency to stick onto each other and to bend very easily. They demanded even more effort in screening for useful crystals.

CuZnSOD, however, was much easier to work with. The long rods were not too thin and quite robust, aligning themselves mostly along the longest axis of the loop as we wanted them to. Insulin also was good to work with, although rhombohedra fitting completely in

the 100 µm beam at the ESRF were often diffracting sub-optimally. The larger crystals that could be used at BESSY, however, diffracted well.

Oscillation range

An oscillation range of 1 degree was used for all the proteins for data collection at synchrotron sources. At the home source the optimal oscillation range is 0.5 degrees and was used for the collection of the myoglobin data sets.

Doses

Doses applied in the experiments range between 4.5 MGy for insulin and 11 MGy for azurin and are found in Table 4.2.

The lowest dose a crystal of one protein species had absorbed was chosen as the target dose for all the other crystals. We removed all the images from the end of the collected data sets that had had a higher dose. It would have been preferable to get data sets with the same number of images for all the crystals, but this is not possible using a beam with non-constant flux.

Data collection at the home source

Collection of one data set with comparable dose at the rotating anode instrument takes at least four days. We decided to focus on myoglobin at the home source and spend the limited beamtime at the ESRF on azurin and insulin. Myoglobin mostly demanded elaborate screening of dozens of crystals until a useful one was found, in which case it was safer to just keep it on the goniometer and collect it at home.

We were able to collect only five data sets. Among them, one was of very bad quality and another one seemed to have suffered two disturbances that can be located clearly by looking at the change in spot and spindle deviations during integration. We could, however, not explain them and the problem is not solved by integrating the parts before and after the disturbances separately. Finally, our yield are two native and one salicylate data sets, but the quality of the latter one is not as good as we would wish for.

4.6 Integration and scaling – data statistics and global signs of radiation damage

The total collected images were integrated in space group P1 and then a second time in the correct space group to be able to extract unit cell dimensions from the file INTEGRATE.LP. Then five consecutive subsets were integrated separately. They have been treated as five completely unconnected data sets. Apart from unit cell volumes, other factors that have been proposed to be relevant to radiation damage were taken from the XSCALE output, such as R_{meas}, R_{mrgd}, I/sig, mosaicity, completeness and Wilson B factors. A detailed analysis of these for azurin (general comments would be the same for all proteins) is made in the next few paragraphs. Further graphs also for the other proteins and the corresponding tables can be found in Appendix 7.2.

Table 4.3: Crystal space groups, unit cells and cell contents.

	Space group	Unit cell [Å] (a b c α β γ)	Molecules/ asymm. unit	Protein concentration [mM]	[mg/ml]
Azurin	$P2_12_12_1$	57 80 109.5 90 90 90	4	53	744
CuZnSOD	$P2_12_12_1$ $P2_12_12_1$	51 101 199 90 90 90 47 51 147 90 90 90	4	25.9	421
Insulin	H3	82.5 82.5 34 90 90 120	2	158	918
Myoglobin	$P2_1$ (P1 2_1 1)	64 29 36 90 107.19 90	1	52	881
NiSOD	$P2_12_12_1$	111 113 128 90 90 90	12	49.6	652

Azurin

We have collected nine azurin data sets at two different dates, among them two natives. Data was good to about 1.9 Å for all the crystals, so this was used as resolution cutoff. Statistics for the complete data sets are displayed in Table 4.4. Those factors have also been determined for the five subsets and shall be investigated in regard to their correlation with radiation damage and their reliability as possible indicators of radiation damage.

Table 4.4: Statistics for complete azurin data sets, calculated over all resolution shells above 1.9 Å.

Azurin	Completeness	I/Sigma	Rmeas	Rmrgd	WilsonB [Å²]	Mosaicity [°]	Volume [Å³]
native	99.00%	30.84	11.0%	4.3%	25.30	0.20	512380
native2	99.60%	32.67	11.0%	4.0%	22.66	0.25	503388
hepes	99.80%	31.85	10.2%	3.0%	21.55	0.11	497163
nitroimidazole	99.70%	31.66	11.4%	3.2%	20.60	0.11	495252
ascorbate	97.30%	30.19	11.2%	3.7%	22.70	0.14	493578
thiourea	99.30%	28.09	12.1%	4.3%	22.41	0.24	497633
cysteine	99.50%	36.63	8.4%	3.2%	27.08	0.20	499019
DNQ	99.90%	30.01	10.8%	4.0%	27.51	0.40	501674
nicotinic acid	99.90%	24.36	13.5%	5.0%	26.26	0.49	502150

R_{mrgd}, R_{meas}

Rmrgd and Rmeas, the R factors output by XDS and XSCALE, have been suggested by Diederichs and Karplus[81] as improvements for previously used R factors. R_{meas} indicates how well the crystallographic model fits the unmerged experimental data and is not dependent on redundancy. R_{mrgd} is a measure of the accuracy of data when symmetry-related reflections have been averaged.

The general features of the curves are the same for R_{mrgd} and R_{meas}, therefore only one of them is shown as an example below (Fig. 4.11). Even if in total an upwards trend can be observed in most of the data sets, there are a lot of fluctuations and with only five time points it is therefore problematic to quantify a trend. The ups and downs are most prominent in NA, NI and cysteine. The first native's R factors are smaller and smaller from subset 2 to 5. We tested if this inconsistent behavior can be reduced by always using 180° subsets (overlapping, to reach the same kind of dose), but this was not the case. Ascorbate and hepes are the data sets with the least fluctuations and the smallest increase over all subsets.

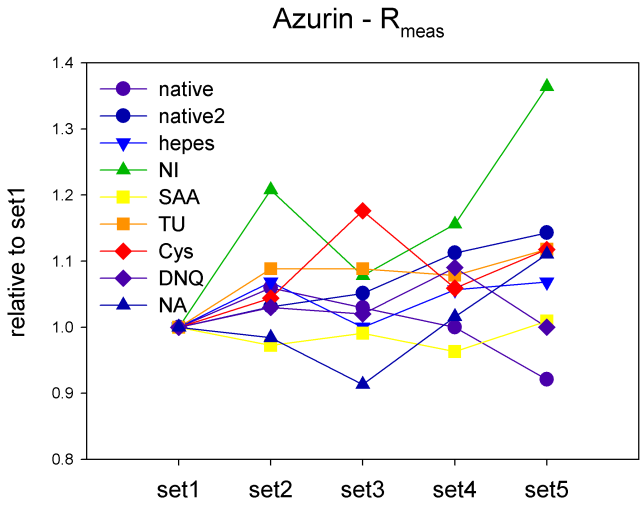


Fig. 4.11: Rmeas as it evolves over the five subsets from each collected azurin crystal.

I/Sigma

As is to be expected, signal intensities follow in principle the trends displayed by R factors, only in opposite directions.

Wilson B factors

The overall Wilson B factors show in general a more regular behavior than the two factors mentioned above. In most cases the increase from set1 to set5 is close to linear. Ascorbate and DNQ show the slowest rise, thiourea is increasing the most.

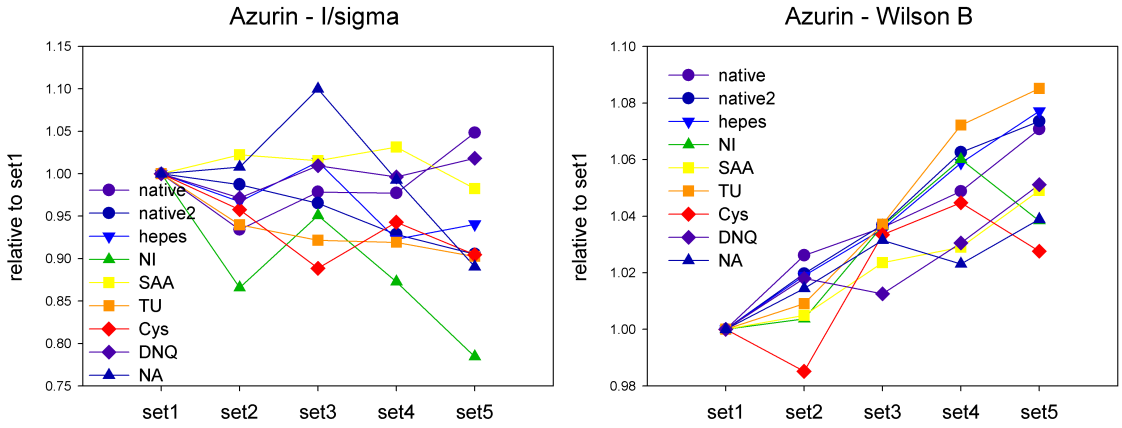


Fig. 4.12: Signal intensities and Wilson B factors as they evolve over the five subsets.

Completeness of data sets, cell volume

Both completeness and cell volume are changing only very slightly from one subset to the next. The unusual drop in completeness in subsets 2 and 5 of the cysteine data set has of course made us examine the data very closely, but no explanation could be found for this behavior. We suspected it might be because one subset for cysteine comprises only 120 degrees, which is already relatively far from 180. Processing the data in overlapping sets of 180 degrees, however, gave the same results.

Mosaicity

A decrease in mosaicity from the first to the last subset is unexpected and difficult to explain, but was observed for many of the data sets.

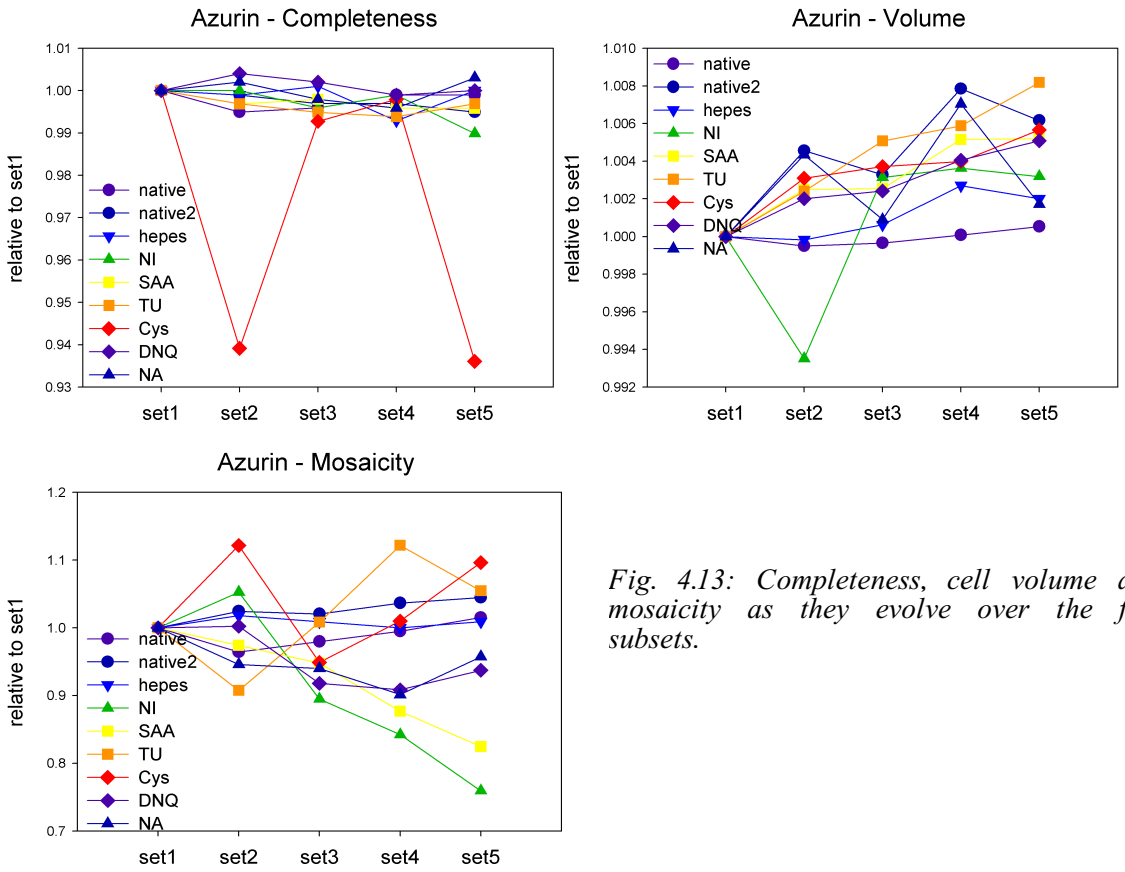


Fig. 4.13: Completeness, cell volume and mosaicity as they evolve over the five subsets.

For the data sets we have collected, it is difficult to draw definite conclusions from the global factors discussed above. The changes observed are small in general and indicate that we are working with rather resistant proteins. More sensitive proteins that display more marked changes could yield more conclusive data.

Insulin

Insulin was scaled to a resolution limit of 1.6 Å. The data set from the crystal with nicotinic acid had a problem close to image 245 that prevented integration of this segment of data. Integrating the parts before and after image 245 separately and merging them only afterwards in the scaling step solved the problem. Although the thiourea and the fluorescein data sets displayed rather high R_{meas} and R_{mrgd} in the lowest resolution shell (R_{meas}: 7 and 10%, R_{mrgd}: 5 and 7%), in total the data looked inconspicuous.

The statistics for the full data sets are summarized in Table 4.5.

Table 4.5: Statistics for complete insulin data sets, calculated over all resolution shells above 1.6 and 1.8 Å for data collected at BESSY and the ESRF (*), respectively.

Insulin	**Completeness**	**I/Sigma**	**Rmeas**	**Rmrgd**	**WilsonB [Å^2]**	**Mosaicity [°]**	**Volume [Å^3]**
native	99.50%	48.56	6.40%	1.50%	29.93	0.27	189381
native*	99.90%	41.24	9.70%	1.80%	20.15	0.20	192123
cysteine	99.90%	44.88	7.50%	1.60%	21.10	0.19	191344
DNQ	99.90%	38.15	8.10%	2.20%	20.74	0.35	191416
Flu	99.90%	29.62	11.40%	2.70%	18.87	0.18	191396
NA	99.90%	60.34	6.40%	1.30%	25.72	0.19	191574
NF	99.90%	41.24	9.70%	1.80%	20.15	0.20	192123
TU	99.90%	60.34	6.40%	1.30%	25.72	0.19	191574
NSA*	99.90%	60.34	6.40%	1.30%	25.72	0.19	191574

Myoglobin and the analysis of data collected on the home source

Unfortunately we have not been able to collect enough myoglobin crystals to do statistically significant comparisons. Of the two data sets collected at the ESRF, the one with salicylate is quite good but the native crystal showed clear signs of anisotropy.

The software delivered with the Proteum X8 home source is designed for integrating and scaling data collected with the machine. Furthermore it includes xprep[82], a program for data analysis and modification, which outputs reflections in different file formats to be used by other programs (for usage with CCP4, scalepack format was chosen by us). It is not possible to output the data into the same format as XDS_ASCII.HKL, which is required for input into XDSSTAT. Therefore we decided to integrate also the home source data in XDS. This proved to be non-trivial, as several parameters have to be taken care of when preparing the input file. The most important is that the images have to be "unwarped" (possible with the Proteum software) to remove spatial distortions before they can be used by XDS (these corrections are taken care of by the Proteum software automatically). Furthermore, the keyword AXIS had to be deleted in the REFINE(INTEGRATE) step.

In the end we were able to process the data in XDS with reasonable statistics, but noticed that refinement statistics using REFMAC5 were significantly (1 to 2% in R and Rfree) worse than for data processed by the Proteum software. We thereafter used diffraction data processed by the Proteum software to calculate and refine structures, but extracted data statistics and obtained R_ds from XDS/XSCALE output.

Statistics for the processed myoglobin data sets can be found in Table 4.6.

Table 4.6: Statistics for complete myoglobin data sets, calculated over all resolution shells above 1.8 and 2.2 Å respectively for data collected at the ESRF and at home (*).

Myoglobin	**Completeness**	**I/Sigma**	**Rmeas**	**Rmrgd**	**WilsonB [Å^2]**	**Mosaicity [°]**	**Volume [Å^3]**
native	99.80%	14.81	15.20%	6.00%	19.34	0.82	60700
native1*	95.50%	30.08	15.60%	3.30%	20.92	0.44	64702
native2*	96.00%	60.57	10.10%	2.00%	21.85	0.47	64570
NSA	99.30%	35.39	6.30%	2.60%	22.56	0.23	60932
NSA1*	98.90%	38.44	14.40%	2.50%	27.5	1.02	64695
NSA2*	97.70%	15.1	38.00%	7.10%	14.12	0.8	66824

4.7 The R_d

The decay R factor proposed by Diederichs[14] (see also the chapter on global damage in the introduction) has been calculated for all of our data sets using XDSSTAT[14]. This program outputs for each frame difference the number of frames that contribute to it and an R_d value. The number of contributing frames is naturally smaller for the largest frame differences, which made the curves very noisy towards the end. This in turn impaired the linear fit of the curves in some cases significantly. Instead of removing a certain number of points towards the end we decided to weight the fit by the number of contributing frames (fitting was done with Sigmaplot, Systat Software Inc.). Fig. 4.14 shows example R_d plots for azurin and insulin, as well as an example for the linear fit. R_d plots for all the crystals can be found in Appendix 7.3.

The periodic features of the curves are more prominent in azurin due to the flat shape of the crystals, but present in all proteins. The wiggles occur around multiples of 180 degrees and are due to the inherent symmetry of the X-ray diffraction pattern. It has to be noted that at exactly 180 degrees (and its multiples) the number of contributing reflections is 10 to 20 times higher than in the regions in between, which leads to a down-shift of the weighted linear fit compared to an unweighted one.

The slopes of the linear fit of the R_d plots are indicators of the level of radiation damage a crystal has suffered. They can be found – along with some statistical analysis of the fits – in Appendix 7.3. Although a linear fit seems appropriate when looking at the graphs, the 180 degree fragments by themselves are not linear at all and sometimes even tend to lean downwards (see Fig. 4.15). This fact as well as the very small magnitude of the slopes (10^{-7} to 10^{-9}) leads to weak correlation coefficients.

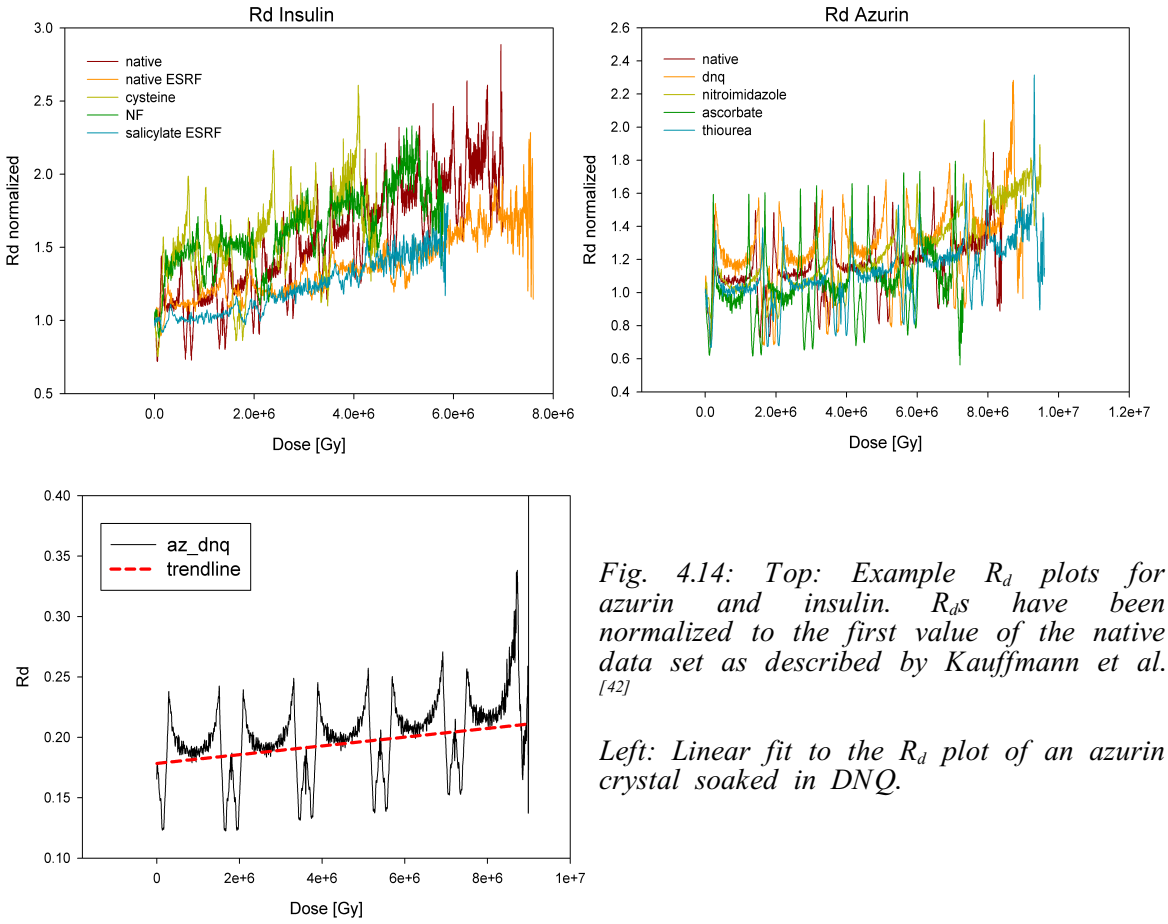


Fig. 4.14: Top: Example R_d plots for azurin and insulin. R_ds have been normalized to the first value of the native data set as described by Kauffmann et al. [42]

Left: Linear fit to the R_d plot of an azurin crystal soaked in DNQ.

R_d plots of two of our data sets – myoglobin with salicylate and insulin with DNQ – are not linear, with a steep increase towards the end. We noticed that this was due to a refill of the storage ring during the data collection. Considerable difference in intensity before and after the refill is naturally interfering with the determination of the R_d, even though

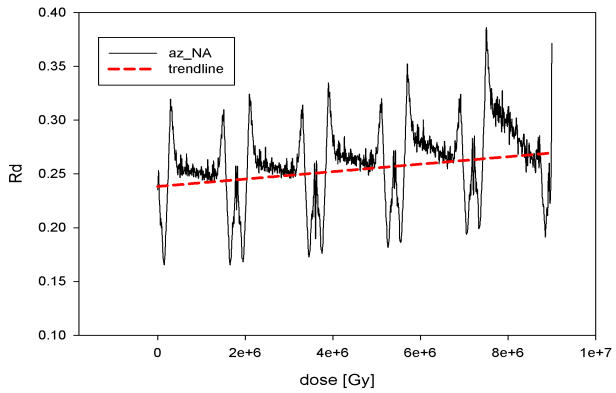


Fig. 4.15: R_d slope of the NA-soaked azurin data set that displays a downwards trend of the curve inside the 180 degree segments.

collection was paused during the refill. We can therefore not compare those two data sets to the others. Fig. 4.16 depicts the magnitude of the slopes of all our crystals along with those obtained for human carbonic anhydrase II (HCAII, data kindly supplied by Bjoern Sjoeblom) native and soaked in fluorescein. For comparison, slopes observed by Kauffmann *et al.*[42] are included in the picture. The slopes reported in the paper are for Rd/frame number and were converted by us to Rd/dose to be directly comparable to our values.

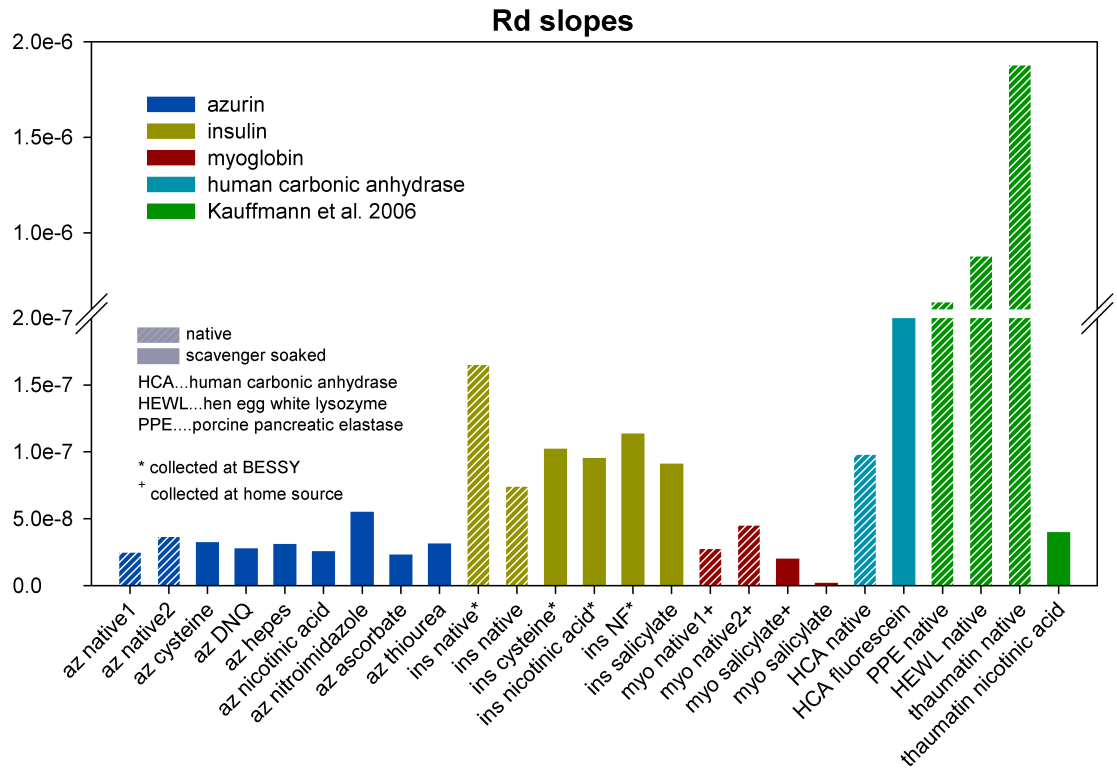


Fig. 4.16: Rd-slopes for azurin, insulin and myoglobin, compared to HCAII and proteins used in the paper by Kauffmann et al.[42]

It can not be missed that our slopes are smaller in comparison to those obtained by Kauffmann *et al*. Among all their scavenger soaked crystals, the thaumatin sample soaked in nicotinic acid (Fig. 4.16) has the lowest Rd slope. All of our azurin or myoglobin samples have comparably low slopes, only for insulin they are a little higher, but still far from the native thaumatin, pancreatic elastase or lysozyme samples reported by Kauffmann *et al*. This is especially striking because our experimental procedure was very similar and dose of the same order of magnitude (ours was higher in azurin and myoglobin).

Azurin and myoglobin crystallization conditions both contain relatively high concentrations of ammonium sulfate (3.8 and 3.2 M) and it is tempting to suspect a

radiation protective effect of ammonium sulfate. Insulin, however, does not fit into the scheme and its slopes, although a little higher, are still lower than the native values reported by Kauffmann *et al*.

The surprisingly low slopes we observe, together with the low correlation coefficients make a comparison and evaluation of the data on the basis of R_d slopes difficult. Overall there are few cases in which the differences scavenger-native are bigger than those observed between two native datasets of the same protein. An extraordinary small slope is observed in the salicylate sample of myoglobin. The same effect is not observed for salicylate-soaked insulin.

4.8 Structure refinement

The individual refinement and analysis of the structures is one of the most time consuming steps in a study like this, even though the refinement is already reduced to a minimum by the availability of models of exactly the same protein for molecular replacement.

The ingredients for structure determination by molecular replacement in MOLREP[73] (CCP4 suite) are

- a *model* of a previously solved homologous structure (in our case of the same protein) and
- the *observed structure factor amplitudes* (plus estimated standard deviations) derived from diffraction spot intensities, brought into mtz file format using XDSCONV or xprep in the case of data collected at the home source

The structures from the PDB (Protein Data Bank) were usually edited to remove solvent molecules and metals from the file. Furthermore, the atoms of the model were subjected to small random shifts (using PDBset[68] and shifts of 0.3 Å) to reduce model bias.

Molecular replacement (even only rigid body refinement) turned out to be unnecessary for insulin and azurin. The model structures deposited in the PDB (2A3G and 4AZU, respectively) were fitting our data well enough so that we could start refinement in REFMAC5[74] immediately. Myoglobin data sets, however, had to be subjected to a molecular replacement search prior to refinement. The starting model used is found in the PDB under the accession number 1WLA.

For every crystal only the first (low dose) subset was refined to obtain calculated phases for the calculation of Fo-Fo difference maps. A short summary of the refinement procedure leading to the final models used for the analysis of local specific damage

effects is described below. The weight w_a mentioned for the REFMAC5 runs is a parameter determining how much influence the experimental X-ray data should have over theoretical geometric optimal restraints. For high resolution data, more weight can be given to the X-ray term and the value of w_a can be higher.

Azurin

- 10 cycles of restrained refinement in REFMAC5 (weight 0.2)
- manual model building in COOT and placement of the 4 copper ions (attributed to their respective chain by editing the pdb file)
- 10 cycles of restrained refinement in REFMAC5 (weight 0.3) with one cycle of ARP/wARP to add waters
- manual model building in COOT, validation of water moleculess and placement of additional solvent molecules: sulfates, nitrates, acetates and one sodium ion
- 10 cycles of restrained refinement in REFMAC5 (weight 0.3), again adding waters by ARP/wARP

The final number of water molecules with B factors lower than 50 $Å^2$ (~ two times of Wilson B factor) for azurin was between 400 and 450 molecules. Statistics of the refinement for all azurin data sets are shown in Table 4.7.

Table 4.7: Final statistics for the refinement of azurin structures.

Azurin	native	native2	cysteine	DNQ	Hepes	NA	NI	SAA	TU
R	**0.183**	**0.178**	**0.190**	**0.188**	**0.171**	**0.191**	**0.177**	**0.175**	**0.182**
Rfree	**0.236**	**0.225**	**0.238**	**0.237**	**0.218**	**0.256**	**0.215**	**0.218**	**0.227**
RMSbond [Å]	0.018	0.015	0.017	0.017	0.014	0.018	0.015	0.014	0.016
RMSangle [°]	1.546	1.465	1.595	1.658	1.412	1.795	1.422	1.466	1.553
RMSchiral [$Å^3$]	0.179	0.188	0.177	0.179	0.189	0.204	0.192	0.196	0.197

Myoglobin

- molecular replacement using MOLREP
- 10 cycles of restrained refinement in REFMAC5 (weight 0.2)
- manual model building in COOT and placement of the iron, the heme group and two sulfates
- 10 cycles of restrained refinement in REFMAC5 (weight 0.3) and modelling of water molecules through ARP/wARP

For the data collected using the home source it was necessary to use a weight of 0.1 in both REFMAC5 runs to obtain acceptable geometry. In the data sets collected at the ESRF a little over 170 water molecules could be found, while in the data collected at the home source only very few (< 10 mostly) could be placed. The data statistics in Table 4.8 are in general worse than those for azurin. With the exception of the NSA data set collected at the ESRF, they all have relatively high differences between R and Rfree, which points to a possible problem of over-refinement which it has not been possible to resolve. In the case of the NSA set from the home source, the collected data might also have been of too low quality (see Table 4.6 above).

Table 4.8: Final statistics for the refinement of myoglobin structures.

Myoglobin	native	native1*	native2*	NSA	NSA*
R	**0.167**	**0.187**	**0.188**	**0.162**	**0.198**
Rfree	**0.240**	**0.268**	**0.258**	**0.214**	**0.287**
RMSbond [Å]	0.017	0.012	0.012	0.012	0.016
RMSangle [°]	1.569	1.290	1.331	1.244	1.526
RMSchiral [Å^3]	0.107	0.082	0.095	0.079	0.097

Myoglobin's space group $P2_1$ can be indexed in more than one way. Given a reference mtz file, the new program POINTLESS[69], which will be part of the next CCP4 release, can propose the reindexing scheme that will adjust a test mtz to the reference. The reindexing itself then has to be performed in a separate step in REINDEX[68].

Insulin and crystal twinning

- 10 cycles of restrained refinement in REFMAC5 (weight 0.3) and modelling of waters through ARP/wARP
- manual model building in COOT and placement of the metals, validation of water molecules
- 10 cycles of restrained refinement in REFMAC5 (weight 0.3)

Insulin's asymmetric unit contains no more than 100 amino acids. Already after the first REFMAC5 run there were hardly any residual difference peaks in the maps. The refinement statistics, however, showed unacceptably high R factors for many of the data sets (for a resolution of 1.6 Å and 1.8 Å for ESRF data).

Table 4.9: Final refinement statistics of insulin structures.
** collected at the ESRF*

Insulin	native	native*	cysteine	DNQ	fluorescein	NA	NF	TU	NSA*
R	**0.191**	**0.252**	**0.188**	**0.190**	**0.234**	**0.262**	**0.244**	**0.284**	**0.192**
Rfree	**0.247**	**0.299**	**0.225**	**0.227**	**0.283**	**0.336**	**0.284**	**0.351**	**0.241**
RMSbond [Å]	0.012	0.018	0.012	0.012	0.015	0.014	0.017	0.020	0.018
RMSangle [°]	1.286	1.659	1.256	1.219	1.460	1.493	1.709	1.900	1.634
RMSchiral [Å^3]	0.086	0.117	0.092	0.078	0.096	0.094	0.109	0.112	0.112

We had noticed that POINTLESS (space group H3 also allows for alternative indexing schemes) was very ambiguous about the correct indexing scheme. A suspicion grew that the data might be twinned, which was unfortunately supported by twinning indicators obtained from running SFCHECK[68] and TRUNCATE[68]. The partial twinning test performed by SFCHECK and cumulative intensity distributions of acentric reflections listed by TRUNCATE both indicate merohedral twinning. Twinning fractions (α) – indicating the fractional volume the smaller twin occupies – are output by SFCHECK (Fig. 4.17b).

When merohedral twinning is present, reflections of two lattices overlap and since they will not, in most cases, both be strong or weak at the same time. The relative number of reflections with extreme intensities is reduced. This leads to the observed deviation of the cumulative intensity distribution from the theoretical untwinned shape (see Fig. 4.17a).

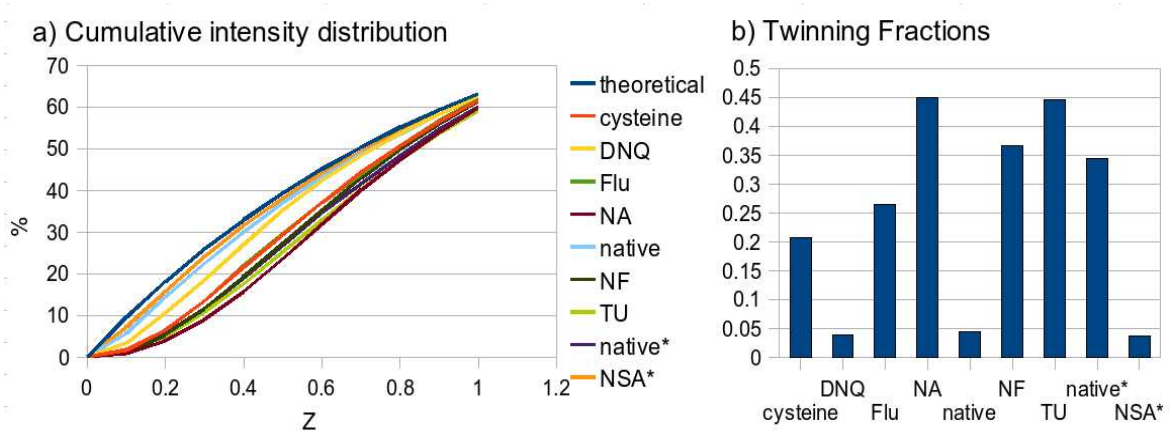


Fig. 4.17: Cumulative intensity distributions and twinning fractions as determined for the insulin data sets. a) plots the percentage of acentric reflections with a Z value lower than the value indicated on the abscissa ($Z = I'/\langle I' \rangle$).

An example for the result of the partial twinning test by TRUNCATE is printed in Fig. 4.18. It calculates theoretical H values for different twinning fractions ($\langle H \rangle = \langle |I_h - I_{h'}| / (I_h - I_{h'}) \rangle$) and reflection subsets of different size. For an untwinned crystal the observed values should lie along the diagonal from top left to bottom right. The example in Fig. 4.18 would indicate a twinning fraction between 0.3 and 0.4.

There are programs for detwinning merohedrally twinned data sets. In general, the closer the twinning fraction is to 0.5, the more difficult this will be. We tested DETWIN[72] (CCP4 suite) on a data set of intermediate twinning fraction (fluorescein). While it did improve the Rfree a little (Table 4.10), data quality is still not as good as it would be necessary for the subsequent analysis.

Table 4.10: Refinement statistics before and after detwinning (insulin soaked in fluorescein).

	Rfact	Rfree	FOM	RMSbond [Å]	RMSangle [°]	RMSchiral [Å^2]
before DETWIN	0.234	0.283	0.814	0.015	1.460	0.096
after DETWIN	0.233	0.273	0.799	0.016	1.603	0.099

We therefore had to exclude most of the insulin data sets collected from the analysis. Only native and DNQ collected at BESSY and NSA collected at the ESRF were used, since their twinning fraction is very small indicating that they might not be twinned at all.

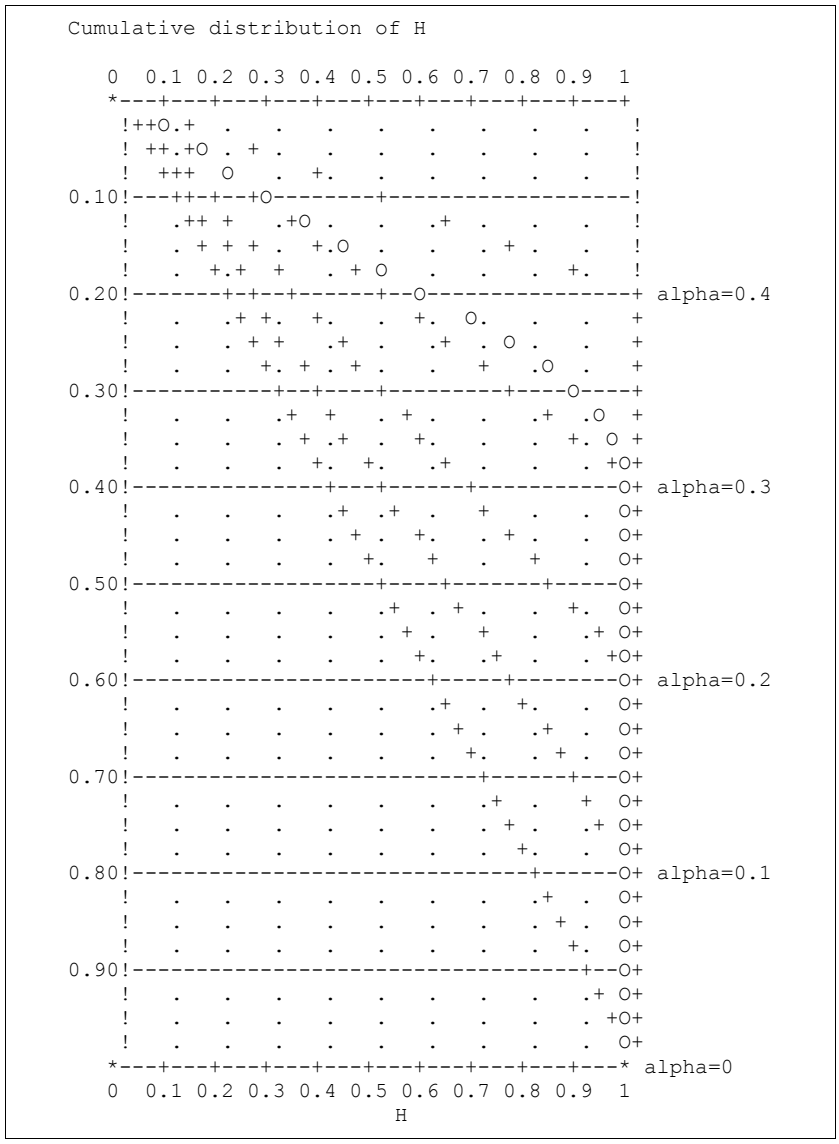


Fig. 4.18: Result of the partial twinning test performed by TRUNCATE for the example of the native data set collected at the ESRF. The ordinate on the left specifies the fraction of reflections for which H was calculated.
The o's are the observed values, +'s are theoretical curves at different twinning fractions (alpha, on the right).

4.9 Analysis of local changes

Fo-Fo maps were calculated (using FFT, CCP4 suite) between subsets 1 to 5, using the calculated phases obtained from refinement of the first subset (the one with the lowest dose). By using only observed structure factors the model bias should be kept at a minimum. For the further analysis, the coordinates of all the peaks higher than 3 sigma levels in those difference maps were output. By using the program CONTACT[72] (CCP4 suite), peaks closer than 1 Å to a protein or solvent atom of the model can be identified. This allows to assign peak heights to specific parts of the structure and evaluate local changes.

4.9.1 General features of the Fo-Fo difference maps

In general, difference peaks in Fo-Fo maps are negative, meaning that electron density is lost in the course of radiation. This is the expected scenario. Higher peaks are always close to either a carboxylate, a thio group or a backbone oxygen atom. This is also reflected in average peak heights if a comparison is made between the values for all atoms and for only carboxyl oxygens and sulfurs.

In azurin the most affected groups are the carboxylates of Asp 6, 55, 62, 71 and 93 and Glu 91, the methylthio group of Met 13 and the carboxylate group of the C-terminus (Lys 128). Furthermore, relatively high negative peaks developed on specific nitrates from the crystallization solution and waters. The only disulfide bridge between Cys 3 and Cys 26 on the other hand is little affected. As for the copper ions, there is either no difference density close to the atom positions at all or the density seems randomly placed with pieces of negative or positive density.

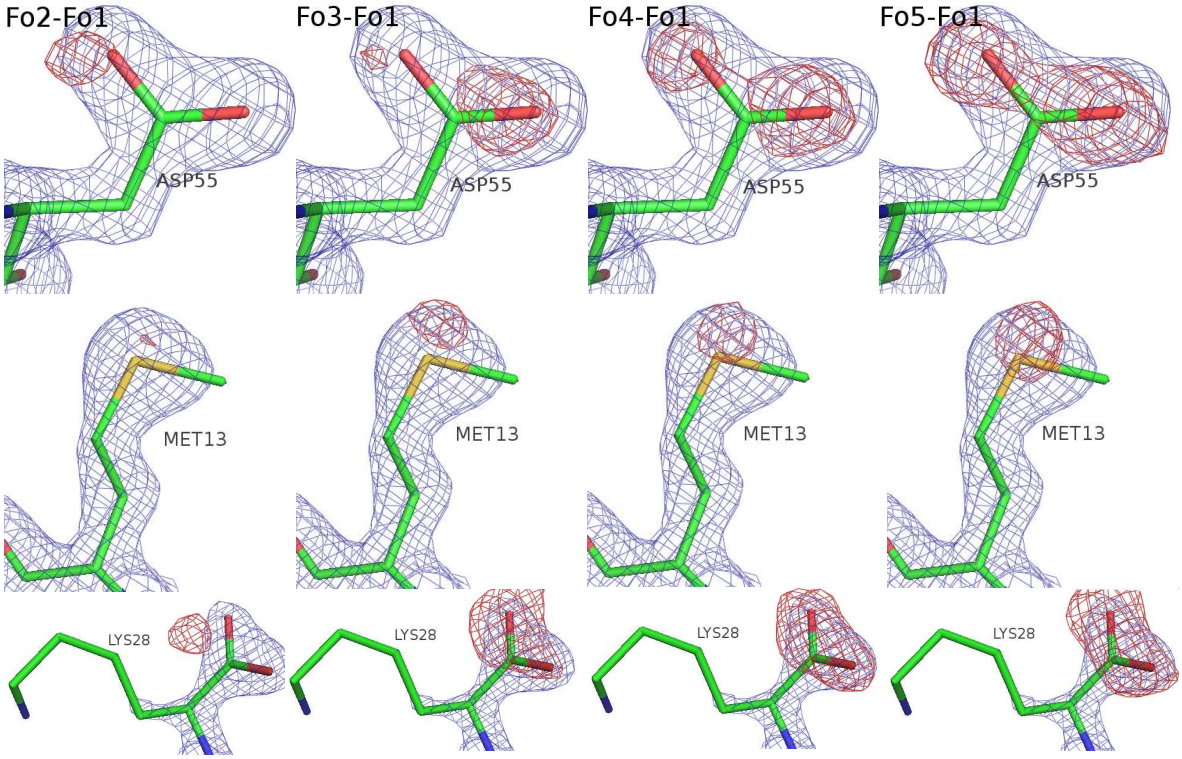


Fig. 4.19: Difference peaks are growing from map Fo2-Fo1 (left-most picture) to Fo5-Fo1 (right-most picture).
The 2Fc-Fo map (blue) is at 2 σ, the Fo-Fo map (red) at -3 σ. The figure was generated from the native2 data set and shows Asp55, Met13 and the C-terminal Lys128 (figure created with PyMOL).

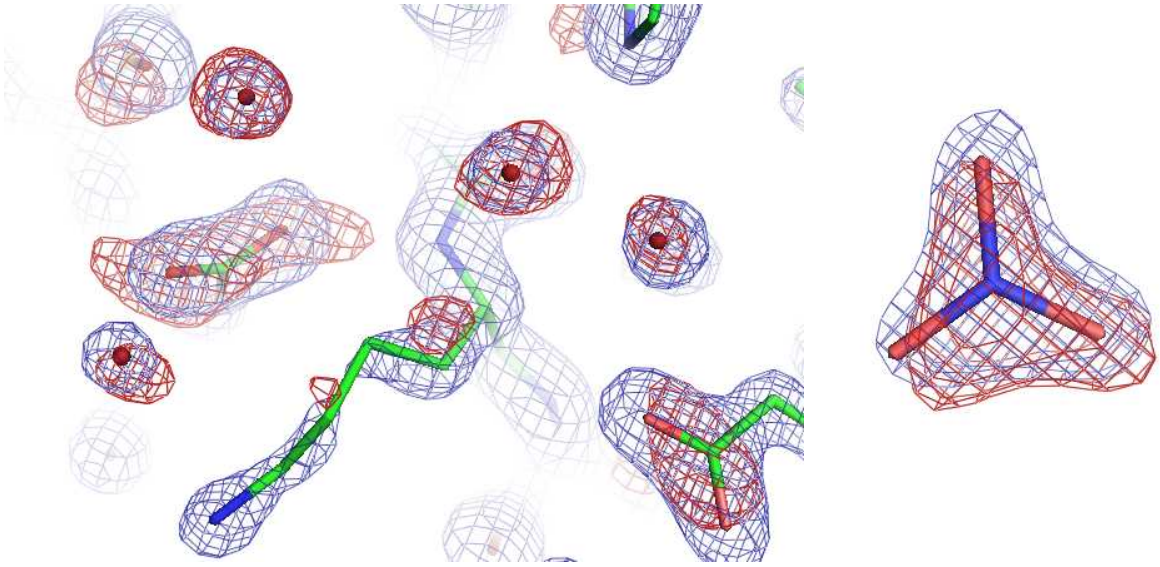


Fig. 4.20: Specific loss of density at solvent molecules (picture from azurin native2, subset 5, Fo-Fo difference map at -3 σ; figure created with PyMOL).

In myoglobin, the most affected residues are Asp 20 and 109, Glu 27, 83 and 105. One methionine (Met 131) is also affected, but to a lesser extent than most of the acidic residues. There are no disulfide bridges in myoglobin. The iron ions show the same random behavior as azurin's copper. It seems, however, that the coordinated water molecule is partly lost (Fig. 4.21).

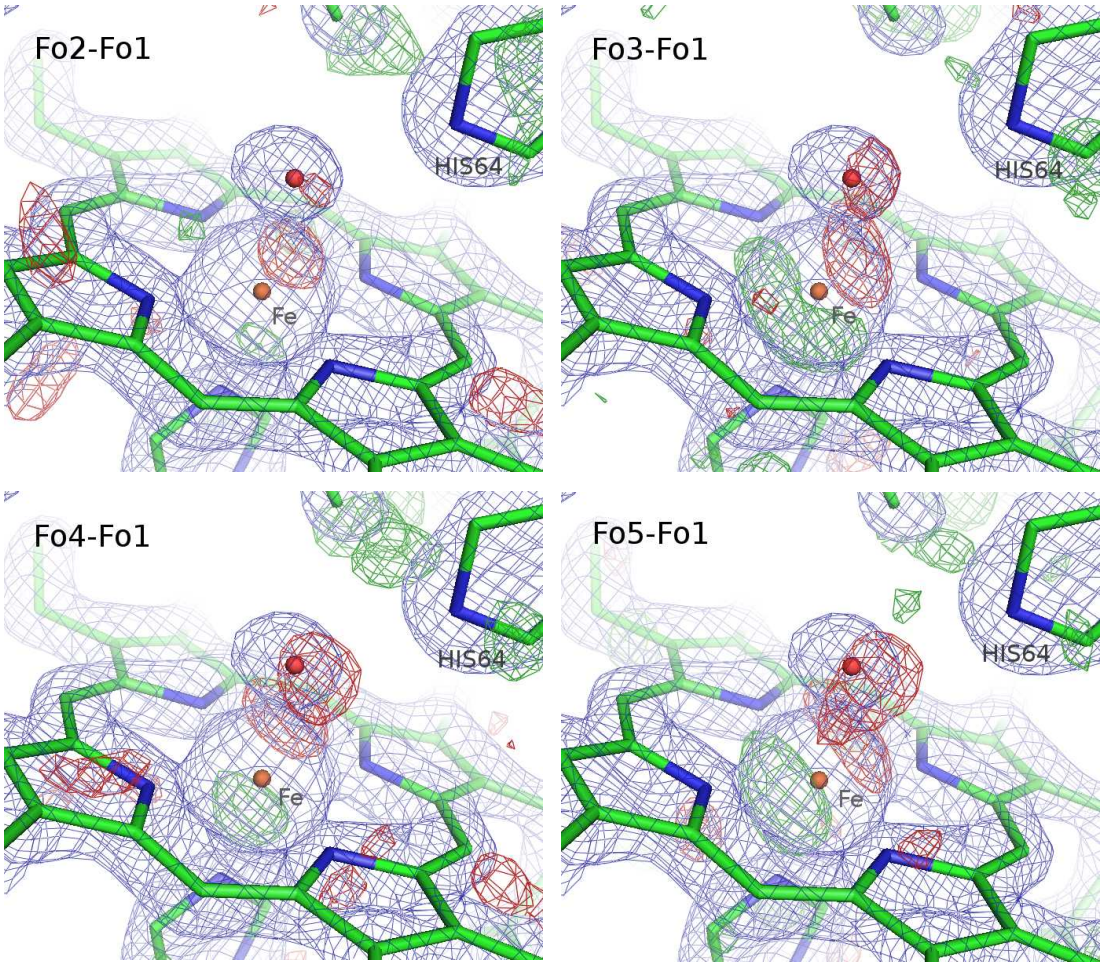


Fig. 4.21: Loss of density at the water molecule coordinated to myoglobin's iron (NSA data set collected at the ESRF). Fo-Fo at -3 σ (figure created with PyMOL).

In insulin the most affected residues are disulfide bridges. There are two inter- and one intramolecular cystines (chain A can be exchanged by C, B by D):

chain	residue		chain	residue
A	Cys 6	-	A	Cys 11
A	Cys 7	-	B	Cys 7
A	Cys 20	-	B	Cys 19

In the three untwinned insulin data sets, there is a relatively consistent ranking of the cysteines in terms of their loss of electron density: first are always Cys 6 (A) and Cys 7 (B). The corresponding residues in chains C and D are always less affected. Also among the first are Cys 7 (A,C) and Cys 20 (A,C), while Cys 19 (B,D) is mostly less affected, after Glu 13 (B,D) and the carboxyl terminii. We have never observed significant positive density close to the disulfide bridges which would indicate defined new coordinates after irradiation. There is only a loss of definition, but not a movement to a specific new position.

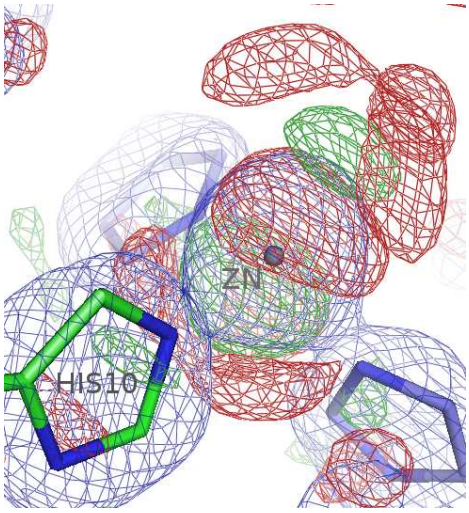


Fig. 4.22: Difference density around the zinc (insulin native). Fo-Fo map at 3 σ (figure created with PyMOL).

The zinc moieties of insulin are positioned exactly at the 3-fold symmetry axis. In the Fo-Fo difference maps, very high negative AND positive peaks (-18 and 30 sigma, respectively) are present close to the metals. An example can be found in Fig. 4.22. The special position at the 3-fold axis might cause this in some way, but it might also be a sign of changes in coordination distances.

4.9.2 Comparison and evaluation of electron difference map peaks

The absolute measure of electron density is e/Å^3 and should therefore be used when peaks from different maps are compared. The unit that we standardly work with when manipulating maps is the sigma level (the *rms deviation from mean intensity*). 2Fo-Fc maps are inspected at a level of 1-1.5 sigma, Fo-Fc maps at 3 sigma. The conversion factor is different for every individual map and can be found in the FFT log file.

Table 4.11: Rms deviations from mean intensity for the Fo-Fo difference maps of all azurin data sets.

	cysteine	DNQ	Hepes	NA	native2	native	NI	SAA	TU
Fo2-Fo1	0.044	0.027	0.024	0.039	0.024	0.027	0.022	0.023	0.032
Fo3-Fo1	0.039	0.028	0.025	0.039	0.026	0.027	0.025	0.023	0.033
Fo4-Fo1	0.026	0.030	0.030	0.048	0.031	0.030	0.028	0.028	0.035
Fo5-Fo1	0.052	0.031	0.032	0.048	0.034	0.032	0.034	0.030	0.039

Not surprisingly, the values grow with growing map distance for most of the data sets. Worrying, however, are data for cysteine or NA, where the values fluctuate and/or are a lot

higher than in the other samples. This data has a higher error level and should be evaluated with the appropriate reserves.

For the three insulin data sets in Table 4.12, the native has a higher deviation from mean intensity than the other two.

Table 4.12: Rms deviations from mean intensity for the Fo-Fo difference maps of insulin data sets that are not impaired by twinning.

	DNQ	native	NSA*
Fo2-Fo1	0.019	0.025	0.015
Fo3-Fo1	0.016	0.031	0.014
Fo4-Fo1	0.024	0.057	0.021
Fo5-Fo1	0.026	0.047	0.024

The comparison at the level of e/Å^3 is not possible at all between data collected at the home source and at the synchrotron. Rms deviations from mean intensities are one order of magnitude smaller in the former, while peaks at the sigma level are of comparable size.

Overall peak statistics

Assessment of scavenger effectiveness by analysis of average peak heights close to more susceptible residues (namely the carboxylates and sulfurs):

- **Azurin**

In Fig. 4.23 the average peak heights over all the peaks were added to demonstrate that this leads to a more random behavior of the curves, since increasing differences between two data sets further apart from each other than two closer data sets is typically expected.

Peak heights at carboxylates and sulfurs as shown in Fig. 4.23b are mostly increasing smoothly from difference map (Fo_2-Fo_1) to (Fo_5-Fo_1), excluding cysteine and NA because of doubts regarding data quality (see map noise levels above and refinement statistics. We have two native crystals that give us an – of course insufficient – estimate of the variance of the data.

Hepes, SAA and also DNQ might have slightly lower peaks than the natives, TU and NI slightly higher, but the differences are certainly not significant when compared with the natives' variance.

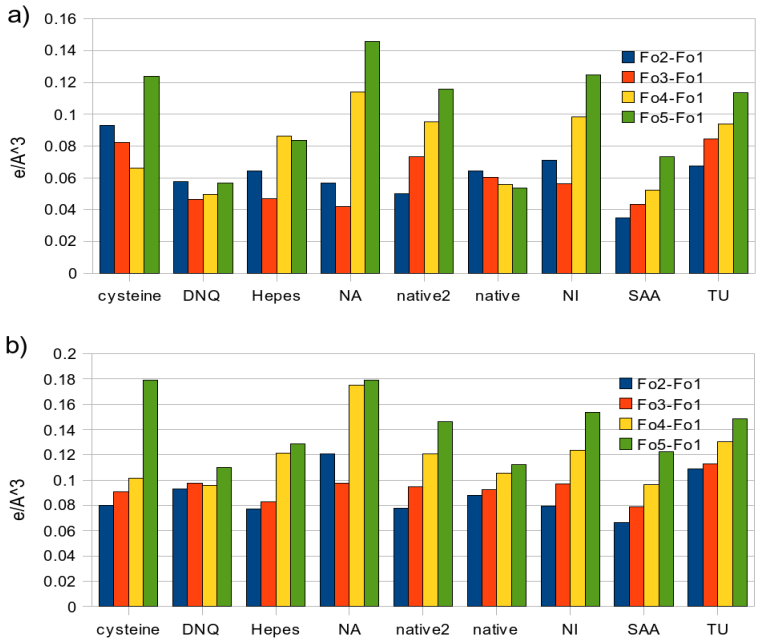


Fig. 4.23: Average negative peak heights for azurin data sets.
(The sign of the values has been removed because the graphs seem easier to evaluate if columns are pointing upwards.)

a) average height of all the peaks close to any protein or solvent atom

b) average height of peaks close to carboxylates or sulfurs

- **Insulin**

Only native and DNQ data sets collected at BESSY and the NSA data set from the ESRF have very low twinning fractions and were included in the analyses.

Rms deviations from mean intensity for the Fo-Fo maps for the three data sets are shown in Fig. 4.24a). The noticeable difference between the native set and the two others is also reflected in the peak heights when they are expressed in e/Å^3 (Fig. 4.24b), while peak heights in sigma levels are very similar in all the sets (Fig. 4.24c).

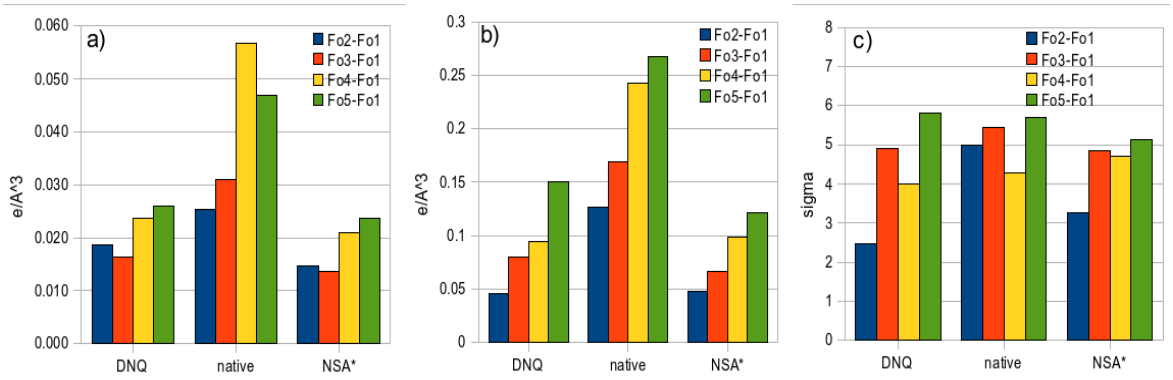


Fig. 4.24: Peak statistics for insulin. Rms deviations from mean intensity (a), average negative peak heights in e/Å^3 (b) and in sigma levels (c)

- **Myoglobin**

Unfortunately we were not able to collect more data for myoglobin. The average peak heights would indicate that NSA is an efficient scavenger, but further tests are absolutely necessary if meaningful conclusions are to be drawn.

Average peak heights are depicted in Fig. 4.25.

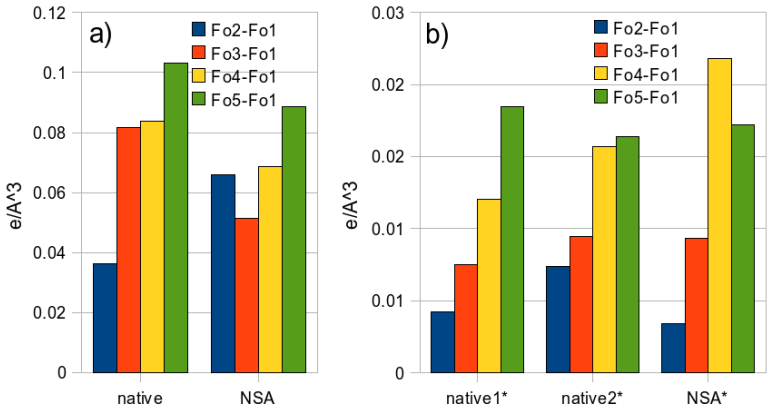


Fig. 4.25: Average peak heights for myoglobin collected at the ESRF (a) and at the home source (b).

4.10 Microspectrophotometry

4.10.1 Preparation

In preparation to the experiments with the microspectrophotometer at the synchrotron, we measured solution spectra of the proteins and scavengers to see where peaks were to be expected and what would be their relative heights.

Solution spectra taken by a standard spectrophotometer at the institute are shown in Fig. 4.26 and Fig. 4.27.

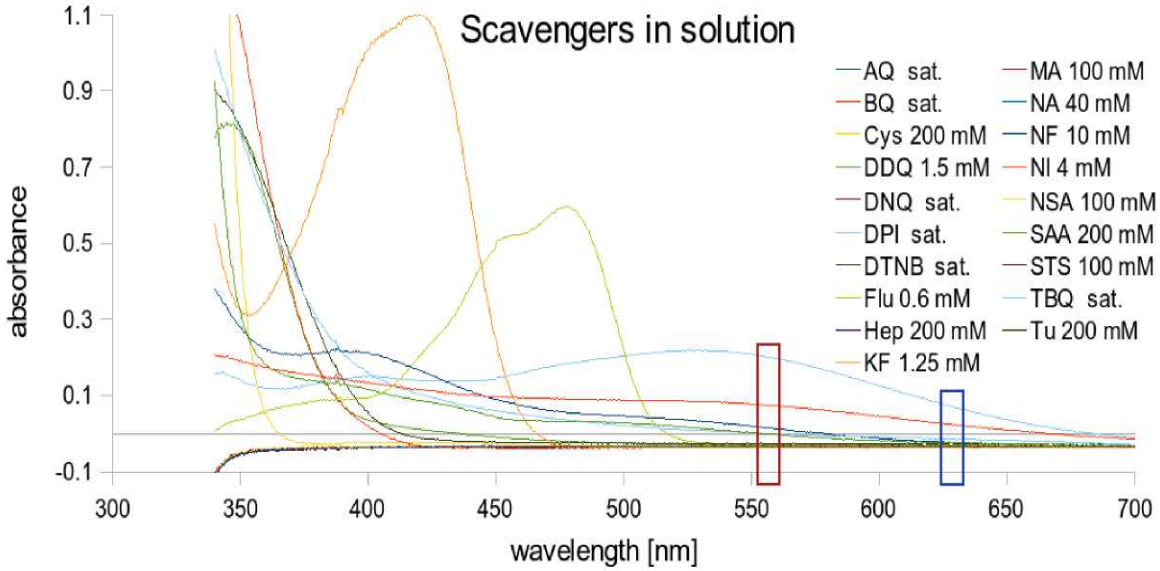


Fig. 4.26: Scavenger solution spectra. KF, Flu, NF, NA, NSA and Ni have been diluted as much as necessary to stay inside the detectors dynamic range, the other scavengers are either at 200 mM as indicated or in a saturated solution. The red and blue rectangles mark the interesting regions for myoglobin and azurin respectively.

For those scavengers that do absorb noticeably at wavelengths beyond 600 nm, the extent of absorption is smaller than that by the proteins in solution. When we consider that the concentration of the protein in the crystal is about 60 times higher than in solution we do not expect scavenger absorption to interfere significantly with the shape of the protein spectrum. Additionally, since we are looking only at changes to the spectrum during irradiation, absorbing scavengers should not be a problem as long as their absorption does not change significantly in the course of irradiation (see below the section on scavenger absorbance during the microspectrophotometry experiment).

If the absorptions of scavengers and proteins in solution are extrapolated to the concentrations that would be present in the cryoloop and with the path length in the

cryosolution and the protein crystal both about the same, it can be seen that at 632 nm Cu,Zn SOD absorbs about an order of magnitude, azurin three orders of magnitude higher than any of the scavengers.

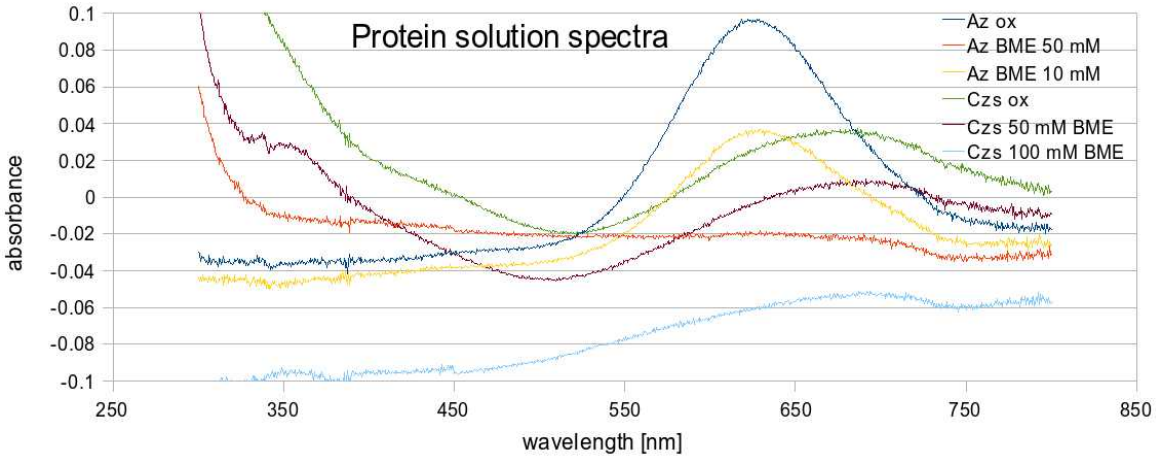


Fig. 4.27: Protein solution spectra for azurin (Az, 0.5 mg/ml) and Cu,Zn SOD (Czs, 15 mg/ml) in oxidized state (ox) and reduced by β-mercaptoethanol (BME).

4.10.2 General

Spectra with well defined characteristics could be collected for both azurin and myoglobin. Both proteins contain a metal whose spectral features change markedly upon reduction[83],[84]. Furthermore the crystals are plate-shaped, which is an advantage when working with this method. Since the crystals are relatively strong-colored, absorbance would otherwise soon be too high. As zinc is not spectroscopically active insulin could not be used in these experiments.

CuZnSOD comes in relatively thin crystals and contains a copper. However, the spectrum of the protein in solution was already very weak (solution is almost colorless with only the slightest touch of blue), and no discernible peak could be observed from the crystal before irradiation. As soon as the X-ray shutter was opened a large peak developed at around 600 nm. This peak can be attributed to solvated electrons stemming from the polyethylene glycol and glycerol in the cryosolution (Fig. 4.28). The peak reaches a maximum and then decreases again slowly over time. We were aware of this problem which is a limitation of single crystal microspectrophotometry. Fortunately, we did not see this effect in the spectra of azurin. For myoglobin, since it has only a small concentration of glycerol in its cryosolution, we were hoping to be able to correct for the effect of solvated electrons (see below in the chapter on myoglobin).

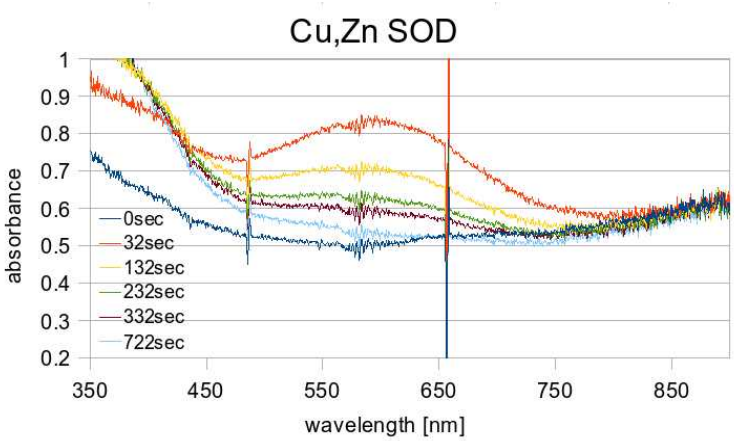


Fig. 4.28: The peak produced by solvated electrons from the Cu,Zn SOD cryosolution in spectra taken with the microspectrophotometer from a frozen crystalline Cu,Zn SOD sample.

Spectra from the following samples have been recorded:

Table 4.13: Samples measured in two beamtime dates at ID14-2 of the ESRF.
** three native azurin crystals were measured, a fourth time course has been recorded on one of the previous three crystals, but on a distinct part of the crystal.*
+ myoglobin was soaked in 100 mM sodium thiosulfate for 20 h, but the spectrum does not show signs of reduction.

	azurin	myoglobin
AQ	x	x
BQ	x	
cysteine	x	
DDQ	x	
DNQ	x	x
DPI	x	x
DTNB	x	x
fluorescein	x	
HEPES	x	x
KF	x	x
maleic acid	x	x
nicotinic acid	x	x
native	xxx(x)*	xx
NF	x	x
nitroimidazole	x	x
salicylate	x	x
ascorbate	x	x
TBQ	x	x
thiourea	x	x
reduced	x	(x)+

For myoglobin, spectra could not be obtained for some of the scavengers. These measurements were aborted because in no crystal orientation could the characteristic features of the spectrum be observed.

4.10.3 Azurin

The spectrum of azurin shows a broad peak around 632 nm that can be attributed to the copper[85]. As soon as X-radiation hits the sample, this peak starts to decrease, rapidly in the beginning and then more slowly (Fig. 4.29). After 20 minutes, which was about the time needed to apply a dose of 6.5 MGy, changes become minimal. Comparison with a chemically reduced crystal (50 mM sodium dithionite, 1 min), where the peak at 632 nm is not present, however, shows that complete reduction by the X-ray beam does not occur. Fig. 4.30 illustrates the variability in the recorded spectra. Before- and after-irradiation graphs for each of the scavengers can be found in Appendix 7.4.

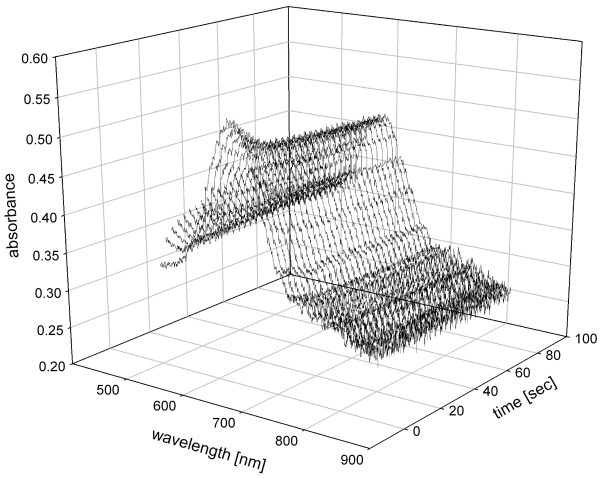


Fig. 4.29: Decrease of the 632 nm peak with time.

What information can be extracted from our measurements?

We are interested in the identification of a scavenger that is able to slow down the reduction of the metal and/or lessen the severity of reduction.

Two features of the experimental data will be investigated in the following:

A) the change of the main peak characteristic of the metal oxidation state

B) the time course of this change

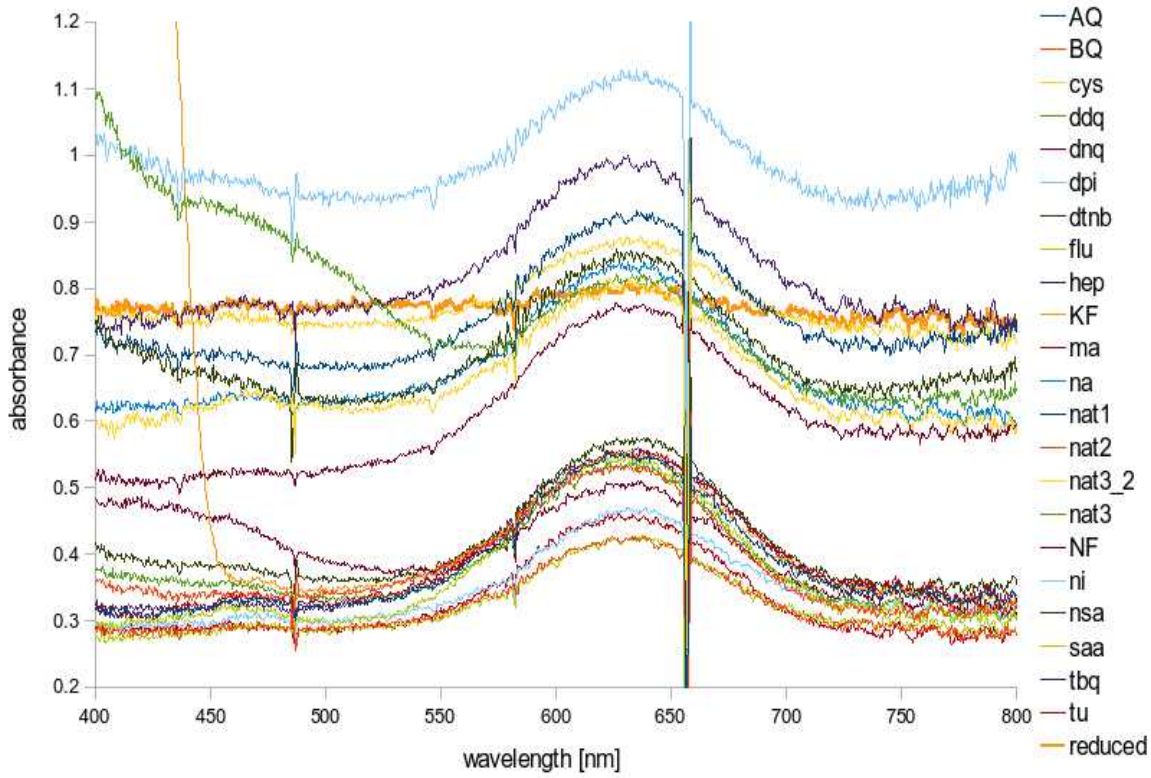


Fig. 4.30: Raw azurin spectra, before the onset of irradiation.

A) Total peak decrease

Differences in the path length and orientation of the crystal make it difficult to directly compare spectra from different crystals. Scaling of spectra was therefore performed to allow for comparison between crystals.

a) Scaling to basal absorbance

Short of a conventional blank that could be measured for each crystal and subtracted from the spectrum, an internal standard had to be found. The region of the spectrum beyond 750 nm does not exhibit characteristic changes in the course of the irradiation. In principle the absorption in this region should be proportional to the crystal dimension in the direction of the light beam (= path length), assuming that the cryosolution layer's contribution can be neglected. If crystal thickness is the determinant factor for the absorbance at 760 nm, the ratio Abs_{632nm} : Abs_{760nm} should be more or less constant throughout the different samples. The observed values vary between 1.14 and 1.75. This variation could indicate that some of the samples are already partly reduced before the start of the experiment.

Assuming that for fully oxidized crystals of whatever thickness, the ratio Abs_{632nm} : Abs_{760nm} should be a fixed value, we can calculate a peak height before and after the application of radiation dose normalized to the absorbance at 760 nm after radiation by simply dividing the whole spectrum by this value. Results are depicted in Fig. 4.31 and Fig. 4.32a).

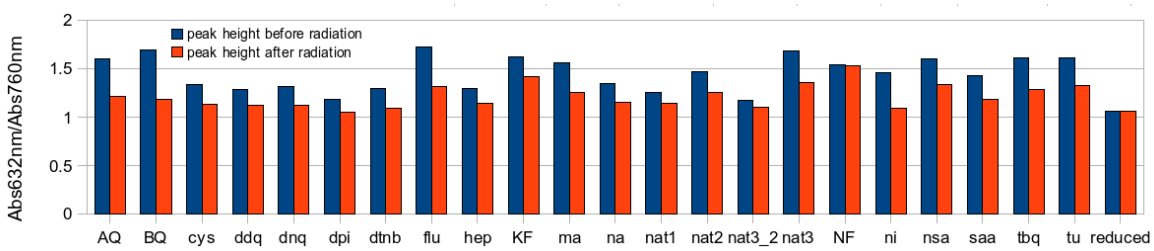


Fig. 4.31: Heights of azurin's 632 nm peak before and after radiation. Scaled to the absorption of the respective sample at 760 nm to correct for different crystal thicknesses.

However, this approximation does not take into account the absorption in the surrounding liquid and cryoloop. Measurements on samples containing only cryosolution show similar magnitude of absorption at 760 nm as the crystals (absorption coefficients are similar at this wavelength). It is therefore not possible to estimate the thickness of a crystal from the absorbance at this wavelength because we can not separate the contributions by the cryosolution layer and the crystal.

Table 4.14: Calculated crystal thicknesses for azurin.

	Abs632-Abs760	crystal thickness [µm]
AQ	0.55	20.3
BQ	0.53	19.6
cys	0.81	30.0
ddq	0.81	30.1
dnq	0.76	28.3
dpi	1.12	41.3
dtnb	0.85	31.4
flu	0.54	19.9
hep	0.99	36.7
KF	0.56	20.6
ma	0.55	20.2
na	0.83	30.6
nat1	0.9	33.3
nat2	0.42	15.6
nat3_2	0.87	32.0
nat3	0.53	19.8
NF	0.51	18.7
ni	0.47	17.3
nsa	0.57	20.9
saa	0.42	15.6
tbq	0.55	20.3
tu	0.46	16.8

Assuming, however, that no or only negligible reduction takes place in the cryocooled crystal before irradiation, the height of the peak can tell us something about crystal thickness and provide us with a means to get all the samples on the same scale.

b) Assuming completely oxidized samples

Knowing the protein concentration in the crystal and the absorption coefficient (5100 $M^{-1}cm^{-1}$ [85]), the thickness of the crystal can be calculated. Given that we can not provide a blank for the crystal spectrum to get an estimate of the

peak height. The value at 760 nm is chosen and subtracted from the peak absorbance again. This should be acceptable since the cryosolution absorbs at a constant value over all wavelengths higher than 500 nm. The unit cell volume plus the number of monomers it comprises allows the determination of a protein concentration in the crystal of 744 mg/ml (53 mM). The calculated crystal thicknesses are listed in Table 4.14. They match the rough values that can be measured under the microscope (around 20 µm) very well.

Following the assumption that crystals are fully oxidized at the start of the experiment and considering that the buffer adds a constant contribution throughout the relevant wavelengths, the difference between Abs_{760nm} and Abs_{632nm} can be taken as a 100% peak and a percental decrease of this peak can be determined. This procedure is consistent with the fact that in a chemically reduced crystal the peak region is completely flat. 'Peak height decreases' in percent for all the azurin samples can be found in Fig. 4.32b). The copper peak never completely disappears for any of the scavengers, which would tell us that photoreduction is not as complete at the delivered dose as in the chemical reduction of the crystal.

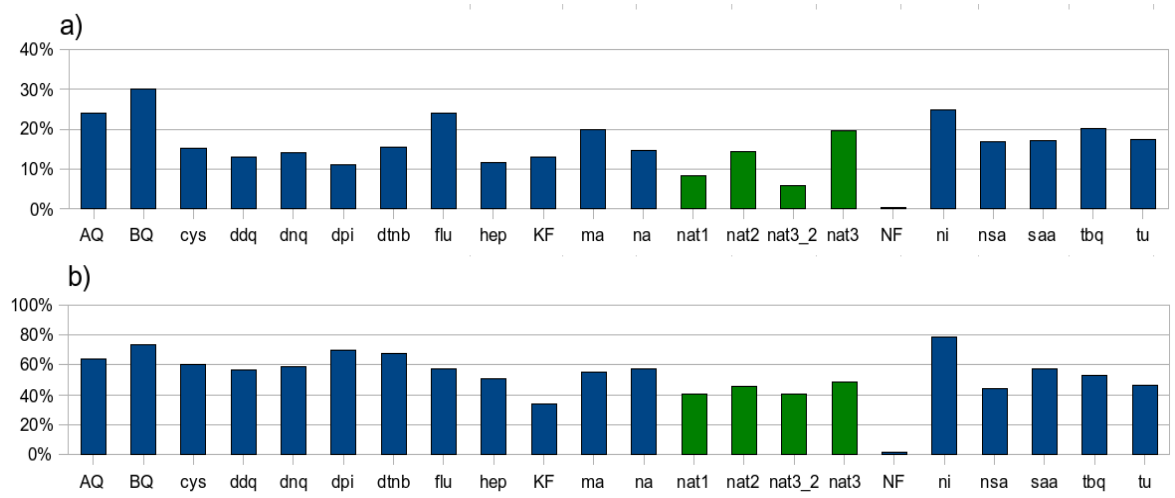


Fig. 4.32: Percental decrease of the 632 nm peak.
Decreases were calculated from data that had been scaled in two different ways: a) scaled only to the absorption at 760 nm (see section a) above), b) scaled assuming that the peak is 100% oxidized in the beginning and total reduction is reached when $Abs_{632nm} = Abs_{760nm}$ *(see section b) above).*

Two additional factors have to be taken into consideration for interpretation of the data:

- scavengers that contribute to absorbance and
- calibration effects

Absorbance by scavengers: Several of the scavengers, in particular the quinones, do absorb in the wavelength range we studied (see chapter 4.10.1). However, this alone should not concern us as long as we are only looking at changes of the spectra. In contrast, a scavenger whose spectral features change with dose will affect our results. For this reason, spectra of selected scavengers (due to lack of time in the beamline only the ones with strongest absorbance in solution) dissolved in the protein cryosolution and frozen in a loop were taken in the same way as for the crystals. Some of them do change with dose and Fig. 4.33 displays the magnitude of those changes in comparison with the changes of the azurin copper peak.

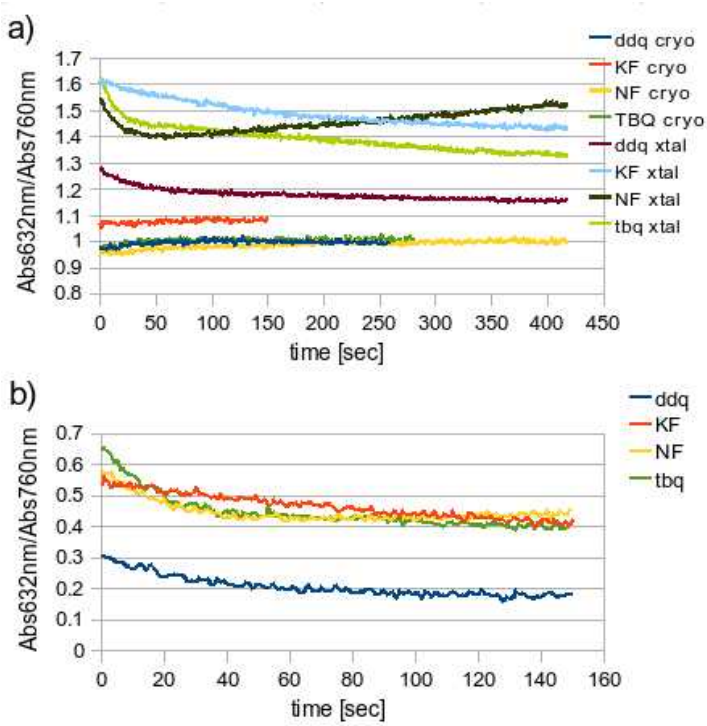


Fig. 4.33: The extent of scavenger absorbance.
a) time courses of the 632 nm-peak for the cryosolutions alone and the crystals, all scaled to Abs_{760nm}.
b) the difference between the cryosolution and crystal spectra.

To compare the graphs, crystal spectra as well as scavenger-cryosolution spectra were scaled to their values at 760 nm, i.e. the whole time course was divided by the absorbance at 760 nm, to account for the differences in path length. The changes in absorbance at 632 nm during 150 sec of irradiation were the following:

	cryosolution + scavenger	crystal
ddq	-0.027	-0.096
KF	-0.017	-0.113
NF	-0.020	-0.108
tbq	-0.033	-0.213

The relative contribution of the scavenger would therefore range between -15 and -30 % of the total change. Scaling everything to the values at 760 nm is questionable in this case. As concluded in the paragraph above (a) Scaling to basal absorbance), the absorbance at 760 nm results from a crystal- and a cryosolution/scavenger contribution and the two can not be separated. We have therefore no means to estimate the thickness of a cryosolution layer in the loop with or without crystal.

Basal absorbance for the cryosolution/scavenger measurements is not different from the average crystal's basal absorbance. This indicates that differences in absorbance between crystal and cryosolution are not as important as the total thickness of the sample. If we assumed that the cryosolution layer is of approximately the same thickness around the crystal as when it is alone in the loop, then the comparison of changes in the raw data would be more accurate. In this case, the relative contributions to absorbance by the scavengers would be very similar to those calculated above for the scaled version, also between -15 and -30% of the total change at 632 nm.

With the presently available data we have to conclude that no absolute estimation of the contribution of the scavenger to absorbance is possible. However, we expect it to be noticeable (15 to 30% of the observed effect) and results for the concerned scavengers have to be taken with reservation.

Calibration effects: Lack of the frequent enough recalibration of the microspectrophotometer between the measurements lead to a gradual upward slope of the spectrum towards longer wavelengths in data collected later in time. While with the well-calibrated device the spectrum is basically flat between 750 nm and 950 nm, it develops a clear positive slope for later data sets (e.g. for DPI as can be seen in Fig. 4.30). We were not aware of the fact that the heating up of the lamps with time can lead to spectral changes in the emitted light and make a recalibration necessary. The reference wavelength of 760 nm used above was chosen as close as possible to the peak in order to reduce this effect and at the same time not be influenced too much by the copper peak at 632 nm.

Peak decrease: Conclusions

In an attempt to estimate the extent of reduction of a metal center in a crystal during the delivery of 6.5 MGy radiation dose (Fig. 4.33), all the above mentioned factors/corrections that prevent analysis of the spectra on absolute scale were taken into account. Of the two scaling models presented, the premises of the second are more likely to be met. This approach results in peak decreases between 40 and 80% - excluding NF and KF spectra that seem clearly to be more heavily influenced by scavenger absorbance.

The time course of the peak from the NF sample (Fig. 4.34 left hand side) as well as the large shoulder growing next to the copper peak (Fig. 4.34 right hand side) for this scavenger suggest that no changes in copper oxidation can be determined here because they are overlaid by the changes produced by the scavenger. KF does not change the spectrum as strongly as NF, but still a large influence is to be expected when the time course of the spectra is inspected. It can, however, not be excluded that the copper oxidation state does benefit from the addition of those two scavengers. The large rise in absorption with a maximum around 450 nm is associated with the uptake of electrons by the iron ions thereby shielding the copper in the active site.

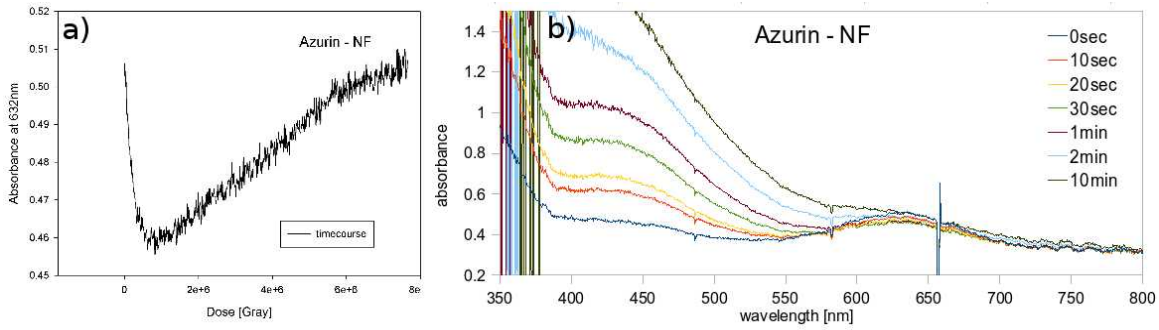


Fig. 4.34: Effects of NF on the spectrum.
a) the time course of the peak at 632 nm. b) an illustration of the growth of the peak at 450 nm that has to stem from the scavenger in comparison to the copper peak.

The decrease of the copper peak upon reduction is smaller for the native crystals (which are fairly homogeneous using the second scaling model). This seems to suggest that none of the scavengers is able to alleviate the problem of metal reduction in the X-ray beam.

B) Rate constants

The decrease with dose of azurin's peak at 632 nm can – with a few exceptions – be fitted very well by a double exponential function of the following form:

$$A = y_0 + a*\exp(-b*x) + c*\exp(-d*x) \qquad (6)$$

I will term b and d "rate constants". They are characteristic for the shape of the time courses and do *not* depend on any kind of scaling that was described in the previous section on peak decrease. The quality of the fits is reported in Appendix 7.4. where relevant statistics are also listed. Fig. 4.36 shows the values of the rate constants. Three (ascorbate, thiourea and hepes) out of twenty-two samples show a bump (see Fig. 4.35 right graph) in the time course curve making convergence dependent on a lot of iterations. Three other samples (native2, KF and salicylate) also do not fit well to the double exponential even though no bump is present. All of them are depicted in Fig. 4.36.

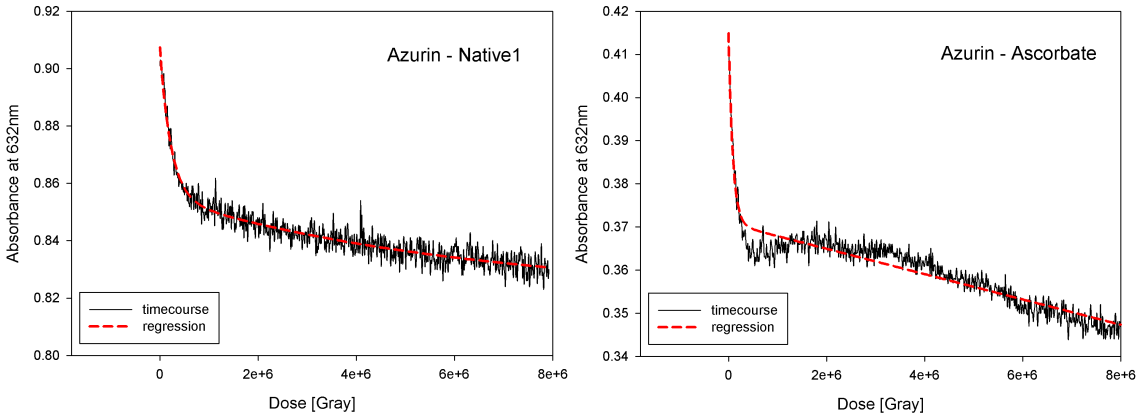


Fig. 4.35: Example graphs showing the decrease of azurin's 632 nm-peak with dose and the fitted trendlines.

The sample soaked in NF can not be fitted at all because after the first quick decrease its time course manifests a marked increase in absorbance (see Fig. 4.34). Probably what we see for KF is just a weaker example of what is observed for NF, both stemming from the same phenomenon discussed in the previous chapter (the scavengers' increase in absorption in the copper peak region with irradiation).

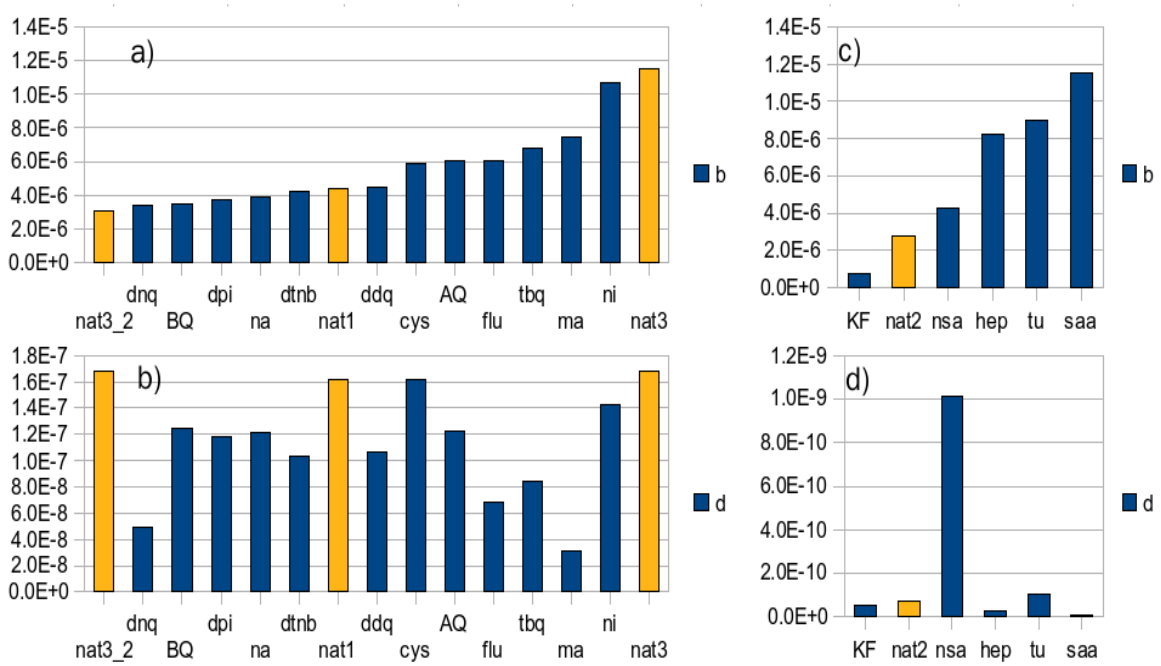


*Fig. 4.36: Rate constants b and d for the decay of azurin's copper peak at 632 nm. On the left hand side (a,b) those samples that are fitted well by A=y0+a*exp(-b*x)+c*exp(-d*x), on the right hand side (c,d) those who can not be fitted well. (Note the difference in scale!) Native crystals are highlighted in yellow, the data in all graphs is sorted by the magnitude of b.*

A larger rate constant leads to a steeper decrease in the beginning of the curve, but also to a flatter shape further to the right. By combining two exponentials the initial and later slopes can be modulated in a more fine-tuned way.

Being the bigger rate constant, b has a larger influence in the beginning, determining how fast the initial reduction is taking place. Looking at Fig. 4.36 we find native crystals with rate constants at the extreme as well as in the center of the observed distribution, there even is one among the curves that could not be fitted. Even if we exclude native3_2, which was collected from the same crystal as native3, only on a different spot and which might constitute an already fairly reduced sample, the spread of the native values is too large to allow any quantitative conclusions on the small differences that might be found between the different scavenger samples.

We therefore have spectroscopic evidence that the tested scavengers do not provide protection from photoreduction to the copper in the active site of azurin.

4.10.4 Myoglobin

Myoglobin has a more complex spectrum than azurin with its single broad peak (Fig. 4.29). Strong characteristic features are only appearing as the metal gets reduced, which makes it more difficult to find the best orientation for the crystal in the beam (see Fig. 4.37). The regions around 540 nm and 650 nm showed some weak characteristics that could be exploited in the initial positioning of the crystal. In some cases it was impossible to find an orientation where these features could be identified, which meant that there would also be no characteristic changes during the radiation procedure. This explains why some spectra are missing for myoglobin while present for azurin. According to Beitlich[83] we should expect the following changes upon reduction: a shift of the Soret band from 413 nm to 427 nm, two new double peaks at 528/536 and 556/566 nm and a slow decrease of the bands between 500 and 700 nm. The shift of the Soret band could, in most cases, not be observed because at these wavelengths absorbance was beyond the detector's dynamic range (only for TBQ it is nice). In contrast, the two double peaks can be distinguished well in the example spectrum with anthraquinone (Fig. 4.37).

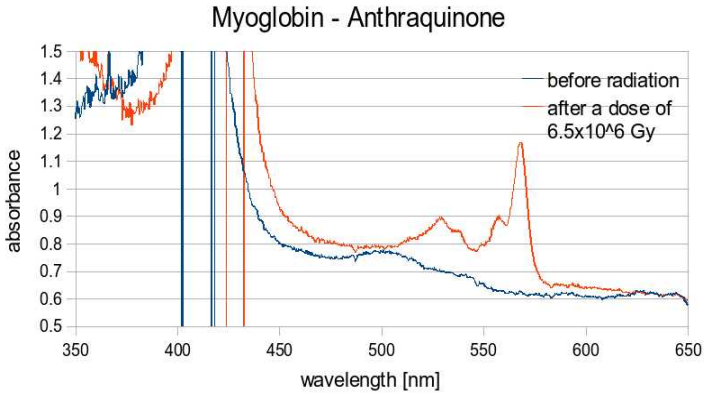


Fig. 4.37: The myoglobin spectrum before and after applying a dose of 6.5 MGy.

Some of the measured samples hardly develop any sign of these characteristic features (DTNB, NI, NSA, native3). Reasons for that might be a suboptimal positioning in the beam despite our efforts, these samples were therefore excluded from analysis.

Time course fitting

For their analysis, Beitlich *et al.*[83] used only the absorbance difference between 566 nm and 555 nm “in order to eliminate any uncorrelated underlying absorbance changes originating for example from solvated electrons”, which we decided to do, too.

In analogy to the disappearing azurin peak, also this time course can be fit by the same

double exponential. An example can be found in Fig. 4.38a), graphs for all the tested scavengers are found in the Appendix 7.4. Unfortunately only a minor subset of the collected samples behaves in this way: AQ, Hepes, native1. MA, SAA and TU can be added to this small group, the fit is of reasonable quality even though the time course of Abs_{566nm}-Abs_{555nm} is decreasing slowly after the initial strong increase.

Others, displaying a steeper decrease later on can not be fitted at all by equation (6). Those are DNQ, NA, TBQ, the chemically reduced crystal, NF and DTNB. A shift of the absorbance difference in this direction can be explained by growing underlying absorption which affects mainly Abs_{555nm}.

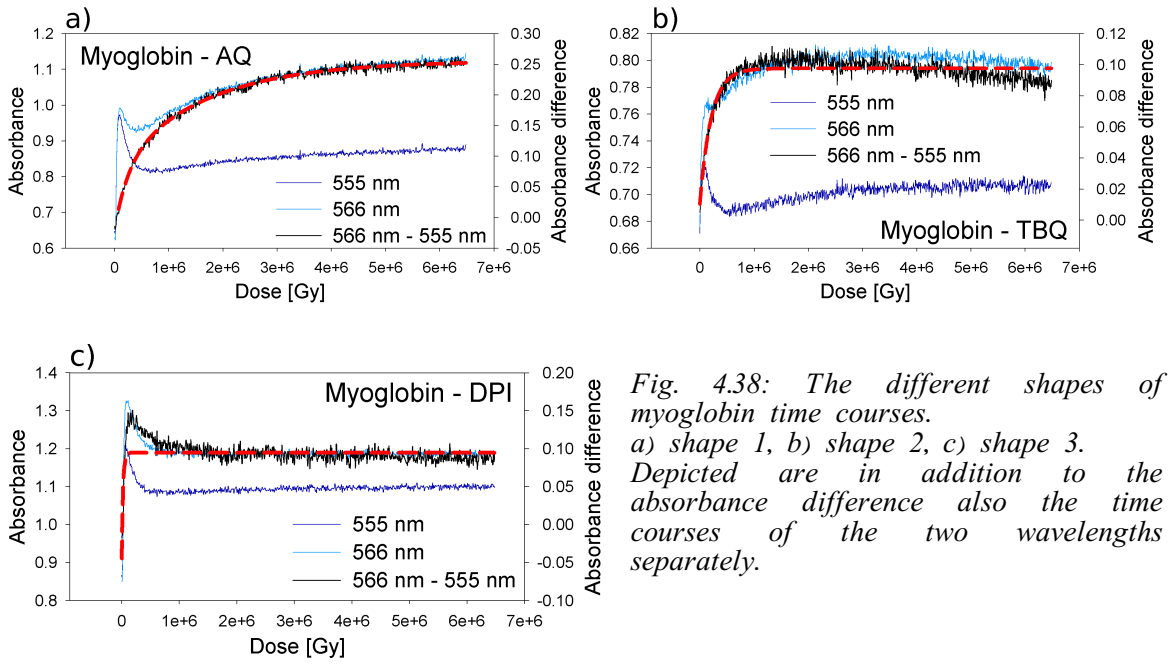


Fig. 4.38: The different shapes of myoglobin time courses. a) shape 1, b) shape 2, c) shape 3. Depicted are in addition to the absorbance difference also the time courses of the two wavelengths separately.

Terming the two curve shapes described above shape 1 and shape 2, respectively (see Fig. 4.38a) and b)), we can observe yet another one, that is present in the time courses of DPI, KF and NI. In contrast to the convex shapes 1 and 2, shape 3 (Fig. 4.38c)) becomes concave after the initial rise that is observed in all cases. For NSA and native3 there is no rise at all, the absorbance difference shows only a slight linear decrease with time. Fig. 4.38a), b), c) summarizes the observed time course shapes. The rapid increase at the very start of the separate time courses of 555 nm and 566 nm is due to the already mentioned solvated electrons, which cause a peak that rises quickly and then disappears slowly over a longer period of time (see Fig. 4.28).

Why do we have so few samples for myoglobin that behave as expected (as observed by Beitlich *et al.*[83])? The different shapes of timecourses can be interpreted as the result of spectral changes that do not stem from the protein and are superimposed on its spectrum. Those changes might, in part, be due to the solvated electrons also reported by

Beitlich *et al.*, but the best candidates at hand are the scavengers. Depending on how strong the changes caused by scavengers are, the curve will either only show a slow decrease when the reduction of the heme has been exhausted or even bend the curve to a concave form. Since we are looking at the difference Abs_{566nm}-Abs_{555nm} this behavior is to be expected when a peak is growing at a little lower wavelength, affecting 555 nm more than 566 nm.

That influences by scavengers should be larger for myoglobin than for azurin becomes plausible, when considering the following: a main difference to azurin is that here we are looking at an absorbance difference between two wavelengths. The same effects would have to be stronger in azurin to be noticeable. Additionally, the scavenger peaks in the solution spectra are predominantly residing in the lower wavelength region of the spectrum, thus stronger at the wavelength of the myoglobin peaks (555 nm) than at 632 nm.

The difficulties with myoglobin spectra were in part caused by the fact that it was probably not possible to find a good orientation for many crystals. Looking at the spectra (→ Appendix 7.4) we see that only a few show the really sharp peak at 566 nm like the AQ sample (Fig. 4.37). Less defined, smaller peaks might be due to a suboptimal orientation. The general rise of the spectra for all scavengers (in both proteins) towards lower wavelengths, becomes a problem then. If all the crystals were oriented in the same way, with the peak at 566 nm as prominent as possible, this influence would be the same for every sample and therefore irrelevant when comparing them.

With only 3 (6 if one is very indulgent) out of 17 samples appropriate to fit a double exponential curve, it does not seem to be an adequate means of analysis for myoglobin. It therefore seems doubtful that any conclusions can be drawn concerning the influence of a specific scavenger on the reduction of the metal.

5 Conclusions and outlook

A methodological project like this is faced with several important considerations that need to be met in order to set up an experimental basis for the planned experiments. It is of crucial importance to produce a large number of crystals with comparable high quality, which is a time-consuming challenge that might delay X-ray experiments.

Beamtime at a synchrotron is limited and long data collection times – 3 hours per data set – consume an average shift of 12 hours quickly. Additionally, inevitably problems can occur – like a thunderstorm that shuts off ring current or a malfunctioning diode – that will reduce effective data collection machine time. At the rotating anode home source a data collection of 4 days needed to reach a comparable dose showed to be very susceptible to any kind of disturbance and time consuming. Due to hardware problems we were able to collect only a limited number of datasets on our home instrument.

Another limitation was the maximum number of samples that can be shipped to the synchrotron, a problem encountered mostly during the microspectrophotometry experiments.

With the data we have been able to collect we can nonetheless draw some conclusions:

5.1 Diffraction data

- A reliable metric of radiation damage is still a subject of research and development in the community. None of the factors we were investigating (R factors, I/sig, mosaicity, unit cell volume, completeness and Wilson B factors) was conclusive at the dose we applied and for the model proteins and scavenger substances we investigated. Cell volume and completeness changes are very small. R factors and signal intensities often show significant fluctuations from one subset to the next inside one data set, which makes a comparison between crystals difficult.

  Crystals soaked with ascorbate show very little of the above mentioned fluctuations and in addition R factors and signal intensities indicate a smaller decrease in quality than the other data sets. Wilson B factors – that are the most conclusively increasing factors measured – are also rising comparably moderately for the ascorbate-soaked data set.

- The R_d is not comparable between different proteins. For our crystals, we obtained consistently very low slopes compared to another publication[42]. Considering that both azurin and myoglobin crystallization conditions include high concentrations of ammonium sulfate, we suspect that this is a favorable factor for preventing radiation damage.

  No scavenger-soaked azurin sample has a significantly lower R_d slope than a native. Salicylate indicates a protective effect for myoglobin collected, which is not present for the insulin data set.

- Comparison of Fo-Fo difference map peaks relies on data of good and comparable quality. Nevertheless, a recommendation to further investigate the effects of hepes and DNQ might be extracted from those maps.

5.2 Microspectrophotometry data

Based on spectroscopic data analysis we can see that none of the tested scavengers prevents the metal photo-reduction. The extracted data might not be very accurate given the specific experimental requirements. Experiments do require a large number of different samples to be tested, evoking limitations in time and storage space. Several control measurements were thus not possible and there was neither time nor space for backup crystals or repeat determinations.

A common feature observed in all of the over 40 measurements is the remarkable change in the spectrum taking place in the very initial stage of the exposure, followed by relatively small spectral changes throughout the rest of the irradiation. The largest leap of reduction takes place in the first 30-60 seconds, corresponding to a dose of 0.2-0.4 MGy. This observation clearly indicated that it is not possible to collect a data set where the metal is not already significantly reduced. It seems that scavengers, even if they can be used for protection from general radiation damage effects, cannot be exploited for protection of protein metal centres from photo-reduction.

We have therefore reasons to believe that structures deposited in the Protein Data Bank (PDB) in the past might not reflect the protein's native state as it was before the X-ray experiment.

Retrospectively, it is to be regretted that we did not have more beamtime with the microspectrophotometer which would have allowed us to test all the scavengers for inherent changes of absorbance during radiation. Then the influence of many of these

scavengers on the appearance of our spectra could eventually have been elucidated.

In the course of the project we had to establish a series of correct and reproducible protocols for the model systems under investigation. As important advancements have been made – crystals of adequate quality can be grown, problems like twinning of insulin crystals are identified, protocols to generate reproducible data have been developed – it should be possible to add further data sets very efficiently that will extend the range of scavengers and allow for a more accurate evaluation of their protective capacity.

6 Bibliography

1 Nave, C. (1995). Radiation damage in protein crystallography. *Radiation Physics and Chemistry*, 45, 483-490.

2 Teplyakov, A., Oliva, G. & Polikarpov, I. (1998). On the choice of an optimal wavelength in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, 54, 610-614.

3 von Sonntag C. (1987). The chemical basis of radiation biology. *Taylor & Francis, London*.

4 Kiefer J. (1990). Biological radiation effects. *Springer-Verlag, Berlin*.

5 Rao, P. S. and Hayon, E. (1974). Interaction of hydrated electrons with the peptide linkage. *Journal of Physical Chemistry*, 78, 1193-1196.

6 Dick, L.A. and Malfant, I. and Kuila, D. and Nebolsky, S. and Nocek, J.M. and Hoffman, B.M. and Ratner, M.A. (1998). Cryogenic Electron Tunneling within Mixed-Metal Hemoglobin Hybrids: Protein Glassing and Electron-Transfer Energetics. *Journal of the American Chemical Society*, 120, 11401-11407.

7 O'Neill, P., Stevens, D. L. & Garman, E. F. (2002). Physical and chemical considerations of damage induced in protein crystals by synchrotron radiation: a radiation chemical perspective. *J Synchrotron Radiat*, 9, 329-332.

8 Jones, G. D., Lea, J. S., Symons, M. C. & Taiwo, F. A. (1987). Structure and mobility of electron gain and loss centres in proteins. *Nature*, 330, 772-773.

9 Sevilla, Michael D. and D'Arcy, James B. and Morehouse, Kim M. (1979). An electron spin resonance study of gamma-irradiated frozen aqueous solutions containing dipeptides. Mechanisms of radical reaction. *Journal of Physical Chemistry*, 83, 2887-2892.

10 Ravelli, R. B. & McSweeney, S. M. (2000). The 'fingerprint' that X-rays can leave on structures. *Structure*, 8, 315-328.

11 Ravelli, R. B. G. & Garman, E. F. (2006). Radiation damage in macromolecular cryocrystallography. *Curr. Opin. Struct. Biol.*, 16, 624-629.

12 Burmeister, W. P. (2000). Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr. D Biol. Crystallogr.*, 56, 328-341.

13 Murray, J. & Garman, E. (2002). Investigation of possible free-radical scavengers and metrics for radiation damage in protein cryocrystallography. *J Synchrotron Radiat*, 9, 347-354.

14 Diederichs, K. (2006). Some aspects of quantitative analysis and correction of radiation damage. *Acta Crystallogr. D Biol. Crystallogr.*, 62, 96-101.

15 Mhaisekar, A., Kazmierczak, M. J. & Banerjee, R. (2005). Three-dimensional numerical analysis of convection and conduction cooling of spherical biocrystals with localized heating from synchrotron X-ray beams. *J Synchrotron Radiat*, 12, 318-328.

16 Weik, M., Ravelli, R. B., Silman, I., Sussman, J. L., Gros, P. & Kroon, J. (2001). Specific protein dynamics near the solvent glass transition assayed by radiation-induced structural changes. *Protein Sci*, 10, 1953-1961.

17 Weik, M., Ravelli, R. B., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman,

I., Kroon, J. & Sussman, J. L. (2000). Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc. Natl. Acad. Sci. U.S.A.*, 97, 623-628.

18 Fioravanti, E., Vellieux, F. M. D., Amara, P., Madern, D. & Weik, M. (2007). Specific radiation damage to acidic residues and its relation to their chemical and structural environment. *J Synchrotron Radiat*, 14, 84-91.

19 Adam, V., Royant, A., Nivière, V., Molina-Heredia, F. P. & Bourgeois, D. (2004). Structure of superoxide reductase bound to ferrocyanide and active site expansion upon X-ray-induced photo-reduction. *Structure*, 12, 1729-1740.

20 Baxter, R. H. G., Seagle, B., Ponomarenko, N. & Norris, J. R. (2004). Specific radiation damage illustrates light-induced structural changes in the photosynthetic reaction center. *J Am Chem Soc*, 126, 16728-16729.

21 Carugo, O. & Djinovic Carugo, K. (2005). When X-rays modify the protein structure: radiation damage at work. *Trends Biochem Sci*, 30, 213-219.

22 Yano, J., Kern, J., Irrgang, K., Latimer, M. J., Bergmann, U., Glatzel, P., Pushkar, Y., Biesiadka, J., Loll, B., Sauer, K., Messinger, J., Zouni, A. & Yachandra, V. K. (2005). X-ray damage to the Mn4Ca complex in single crystals of photosystem II: a case study for metalloprotein crystallography. *Proc. Natl. Acad. Sci. U.S.A.*, 102, 12047-12052.

23 Ascone, I., Meyer-Klaucke, W. & Murphy, L. (2003). Experimental aspects of biological X-ray absorption spectroscopy. *J Synchrotron Radiat*, 10, 16-22.

24 Wilmot, C. M., Sjögren, T., Carlsson, G. H., Berglund, G. I. & Hajdu, J. (2002). Defining redox state of X-ray crystal structures by single-crystal ultraviolet-visible microspectrophotometry. *Meth. Enzymol.*, 353, 301-318.

25 Henderson, R. (1990). Cryoprotection of protein crystals against radiation damage in electron and X-ray diffraction. *Proc R Soc Lond*, 241, 6-8.

26 Sliz, P., Harrison, S. C. & Rosenbaum, G. (2003). How does radiation damage in protein crystals depend on X-ray dose?. *Structure*, 11, 13-19.

27 Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proc. Natl. Acad. Sci. U.S.A.*, 103, 4912-4917.

28 Leiros, H. K. S., Timmins, J., Ravelli, R. B. G. & McSweeney, S. M. (2006). Is radiation damage dependent on the dose rate used during macromolecular crystallography data collection?. *Acta Crystallogr. D Biol. Crystallogr.*, 62, 125-132.

29 Leiros, H. K. S., Timmins, J., Ravelli, R. B. G. & McSweeney, S. M. (2006). Is radiation damage dependent on the dose rate used during macromolecular crystallography data collection?. *Acta Crystallogr D Biol Crystallogr*, 62, 125-132.

30 Weiss, M. S., Panjikar, S., Mueller-Dieckmann, C. & Tucker, P. A. (2005). On the influence of the incident photon energy on the radiation damage in crystalline biological samples. *J Synchrotron Radiat*, 12, 304-309.

31 Shimizu, N., Hirata, K., Hasegawa, K., Ueno, G. & Yamamoto, M. (2007). Dose dependence of radiation damage for protein crystals studied at various X-ray energies. *J Synchrotron Radiat*, 14, 4-10.

32 Teng, T. & Moffat, K. (2002). Radiation damage of protein crystals at cryogenic temperatures between 40 K and 150 K. *J Synchrotron Radiat*, 9, 198-201.

33 Hanson, B. L., Schall, C. A. & Bunick, G. J. (2003). New techniques in macromolecular cryocrystallography: macromolecular crystal annealing and cryogenic helium. *J. Struct. Biol.*, 142, 77-87.

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

35 Garman, E. F. & McSweeney, S. M. (2007). Progress in research into radiation damage in cryo-cooled macromolecular crystals. *J Synchrotron Radiat*, 14, 1-3.

36 Nave, C. & Hill, M. A. (2005). Will reduced radiation damage occur with very small crystals?. *J Synchrotron Radiat*, 12, 299-303.

37 Fulford, J., Nikjoo, H., Goodhead, D. T. & O'Neill, P. (2001). Yields of SSB and DSB induced in DNA by Al(K) ultrasoft X-rays and alpha-particles: comparison of experimental and simulated yields. *Int J Radiat Biol*, 77, 1053-1066.

38 Zaloga, G. & Sarma, R. (1974). New method for extending the diffraction pattern from protein crystals and preventing their radiation damage. *Nature*, 251, 551-552.

39 Cascio, D., Williams, R. & McPherson, A. (1984). The reduction of radiation damage in protein crystals by polyethylene glycol. *Journal of Applied Crystallography*, 17, 209-210.

40 Southworth-Davies, R. J. & Garman, E. F. (2007). Radioprotectant screening for cryocrystallography. *J Synchrotron Radiat*, 14, 73-83.

41 Blundell T. L. & Johnson L. N. (1976). Protein Crystallography. *Academic Press*.

42 Kauffmann, B., Weiss, M. S., Lamzin, V. S. & Schmidt, A. (2006). How to avoid premature decay of your macromolecular crystal: a quick soak for long life. *Structure*, 14, 1099-1105.

43 Bourenkov, G. P. & Popov, A. N. (2006). A quantitative approach to data-collection strategies. *Acta Crystallogr D Biol Crystallogr*, 62, 58-64.

44 Diederichs, K., McSweeney, S. & Ravelli, R. B. G. (2003). Zero-dose extrapolation as part of macromolecular synchrotron data reduction. *Acta Crystallogr D Biol Crystallogr*, 59, 903-909.

45 Berglund, G. I., Carlsson, G. H., Smith, A. T., Szöke, H., Henriksen, A. & Hajdu, J. (2002). The catalytic pathway of horseradish peroxidase at high resolution. *Nature*, 417, 463-468.

46 Wuerges, J., Lee, J., Yim, Y., Yim, H., Kang, S. & Djinovic Carugo, K. (2004). Crystal structure of nickel-containing superoxide dismutase reveals another type of active site. *Proc Natl Acad Sci U S A*, 101, 8569-8574.

47 Ravelli, R. B. G., Leiros, H. S., Pan, B., Caffrey, M. & McSweeney, S. (2003). Specific radiation damage can be used to solve macromolecular crystal structures. *Structure*, 11, 217-224.

48 Nanao, M. H., Sheldrick, G. M. & Ravelli, R. B. G. (2005). Improving radiation-damage substructures for RIP. *Acta Crystallogr. D Biol. Crystallogr.*, 61, 1227-1237.

49 Grininger, M., Ravelli, R. B. G., Heider, U. & Zeth, K. (2004). Expression, crystallization and crystallographic analysis of DegS, a stress sensor of the bacterial periplasm. *Acta Crystallogr D Biol Crystallogr*, 60, 1429-1431.

50 Gigant, B., Wang, C., Ravelli, R. B. G., Roussi, F., Steinmetz, M. O., Curmi, P. A., Sobel, A. & Knossow, M. (2005). Structural basis for the regulation of tubulin by vinblastine. *Nature*, 435, 519-522.

51 Bailey, S., Evans, R. W., Garratt, R. C., Gorinsky, B., Hasnain, S., Horsburgh, C., Jhoti, H., Lindley, P. F., Mydin, A., Sarra, R. & et al. (1988). Molecular structure of serum transferrin at 3.3-A resolution. *Biochemistry*, 27, 5804-5812.

52 Wang, Y., Chen, J., Luo, Y., Funk, W. D., Mason, A. B., Woodworth, R. C., MacGillivray, R. T. & Brayer, G. D. (1992). Preliminary crystallographic analyses of the N-terminal lobe of recombinant human serum transferrin. *J. Mol. Biol.*, 227, 575-576.

59 Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K. & Noble, M. (2002). The CCP4 molecular-graphics project. *Acta Crystallographica Section D*, 58, 1955-1957.

53 Mirkin, N., Jaconcic, J., Stojanoff, V. & Moreno, A. (2008). High resolution X-ray crystallographic structure of bovine heart cytochrome c and its application to the design of an electron transfer biosensor. *Proteins*, 70, 83-92.

54 Wuerges, J., Lee, J., Yim, Y., Yim, H., Kang, S. & Djinovic Carugo, K. (2004). Crystal structure of nickel-containing superoxide dismutase reveals another type of active site. *Proc. Natl. Acad. Sci. U.S.A.*, 101, 8569-8574.

55 Chillemi, G., Falconi, M., Amadei, A., Zimatore, G., Desideri, A. & Di Nola, A. (1997). The essential dynamics of Cu, Zn superoxide dismutase: suggestion of intersubunit communication. *Biophys. J.*, 73, 1007-1018.

56 Hough, M. A. & Hasnain, S. S. (2003). Structure of fully reduced bovine copper zinc superoxide dismutase at 1.15 A. *Structure*, 11, 937-946.

57 Youn, H. D., Kim, E. J., Roe, J. H., Hah, Y. C. & Kang, S. O. (1996). A novel nickel-containing superoxide dismutase from Streptomyces spp. *Biochem. J.*, 318 (Pt 3), 889-896.

58 Smith, G. D., Pangborn, W. A. & Blessing, R. H. (2005). The structure of T6 bovine insulin. *Acta Crystallogr. D Biol. Crystallogr.*, 61, 1476-1482.

60 Dennison, C. (2005). Ligand and loop variations at type 1 copper sites: influence on structure and reactivity. *Dalton Trans*, , 3436-3442.

61 Crane, B. R., Di Bilio, A. J., Winkler, J. R. & Gray, H. B. (2001). Electron tunneling in single crystals of Pseudomonas aeruginosa azurins. *J. Am. Chem. Soc.*, 123, 11623-11631.

62 Cheung, K., Strange, R. W. & Hasnain, S. S. (2000). 3D EXAFS refinement of the Cu site of azurin sheds light on the nature of structural change at the metal centre in an oxidation-reduction process: an integrated approach combining EXAFS and crystallography. *Acta Crystallographica Section D*, 56, 697-704.

63 Adman, E. T., Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1978). A crystallographic model for azurin a 3 A resolution. *J. Mol. Biol.*, 123, 35-47.

64 Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M. & Canters, G. W. (1991). Crystal structure analysis of oxidized Pseudomonas aeruginosa azurin at pH 5.5 and pH 9.0. A pH-induced conformational transition involves a peptide bond flip. *J. Mol. Biol.*, 221, 765-772.

65 Hersleth, H., Dalhus, B., Görbitz, C. H. & Andersson, K. K. (2002). An iron hydroxide moiety in

the 1.35 A resolution structure of hydrogen peroxide derived myoglobin compound II at pH 5.2. *J. Biol. Inorg. Chem.*, 7, 299-304.

66 Murray, J. W., Garman, E. F. & Ravelli, R. B. G. (2004). X-ray absorption by macromolecular crystals: the effects of wavelength and crystal composition on absorbed dose. *J Applied Crystallography*, 37, 513-522.

67 Kabsch, W. (1988). Automatic indexing of rotation diffraction patterns. *Journal of Applied Crystallography*, 21, 67-72.

68 CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, 50, 760-763.

69 Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallographica Section D*, 62, 72-82.

70 French, S. & Wilson, K. (1978). On the treatment of negative intensity observations. *Acta Crystallographica Section A*, 34, 517-525.

71 Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallographica Section D*, 55, 191-205.

72 Yeates, T. O. (1997). Detecting and overcoming crystal twinning. *Meth. Enzymol.*, 276, 344-358.

73 Vagin, A. & Teplyakov, A. (1997). MOLREP: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography*, 30, 1022-1025.

74 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.*, 53, 240-255.

75 Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, 60, 2126-2132.

76 Read, R. J. & Schierbeek, A. J. (1988). A phased translation function. *Journal of Applied Crystallography*, 21, 490-495.

77 Bourgeois, Dominique,, Vernede, Xavier,, Adam, Virgile,, Fioravanti, Emanuela, & Ursby, Thomas, (2002). A microspectrophotometer for UV-visible absorption and fluorescence studies of protein crystals. *J Appl Crystallography*, 35, 319-326.

78 Adman, E. T., Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1978). A crystallographic model for azurin a 3 A resolution. *J Mol Biol*, 123, 35-47.

79 Smith, G. D., Pangborn, W. A. & Blessing, R. H. (2005). The structure of T6 bovine insulin. *Acta Crystallogr D Biol Crystallogr*, 61, 1476-1482.

80 Newman, J. (2004). Novel buffer systems for macromolecular crystallization. *Acta Crystallogr D Biol Crystallogr*, 60, 610-612.

81 Diederichs, K. & Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.*, 4, 269-275.

82 Bruker (2004). XPREP (Version 7.12). *Bruker AXS Inc., Madison, Wisconsin, USA*, , .

83 Beitlich, T., Kühnel, K., Schulze-Briese, C., Shoeman, R. L. & Schlichting, I. (2007). Cryoradiolytic reduction of crystalline heme proteins: analysis by UV-Vis spectroscopy and X-ray

crystallography. *J Synchrotron Radiat*, 14, 11-23.

84 Karlsson, B. G., Tsai, L., Nar, H., Sanders-Loehr, J., Bonander, N., Langer, V. & Sjolin, L. (1997). X-ray Structure Determination and Characterization of the Pseudomonas aeruginosa Azurin Mutant Met121Glu. *Biochemistry*, 36, 4089-4095.

85 Lu, Y. (2003). Electron Transfer: Cupredoxins. In *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology*,

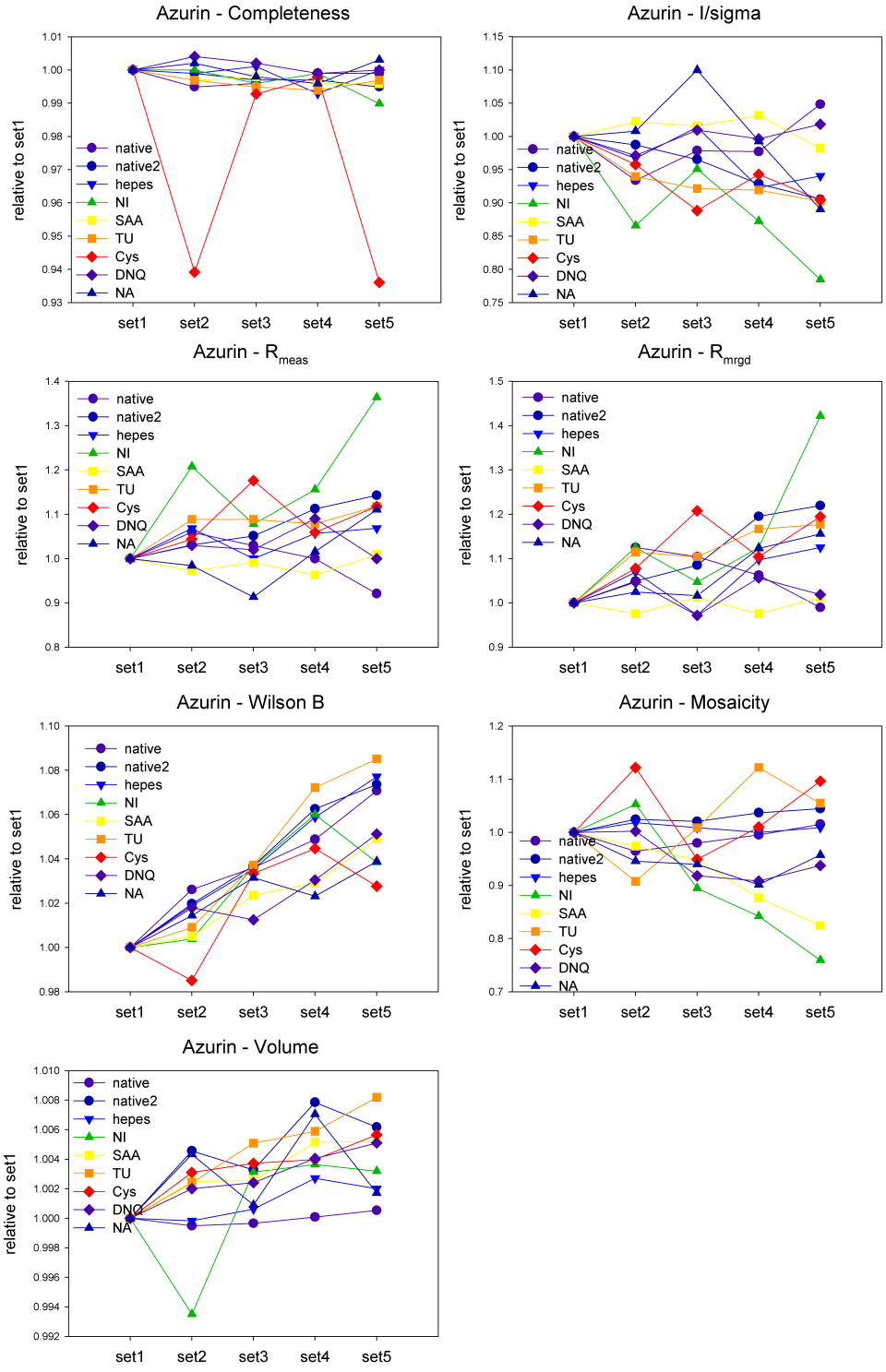
7 Appendix

7.1 Data collection specifics

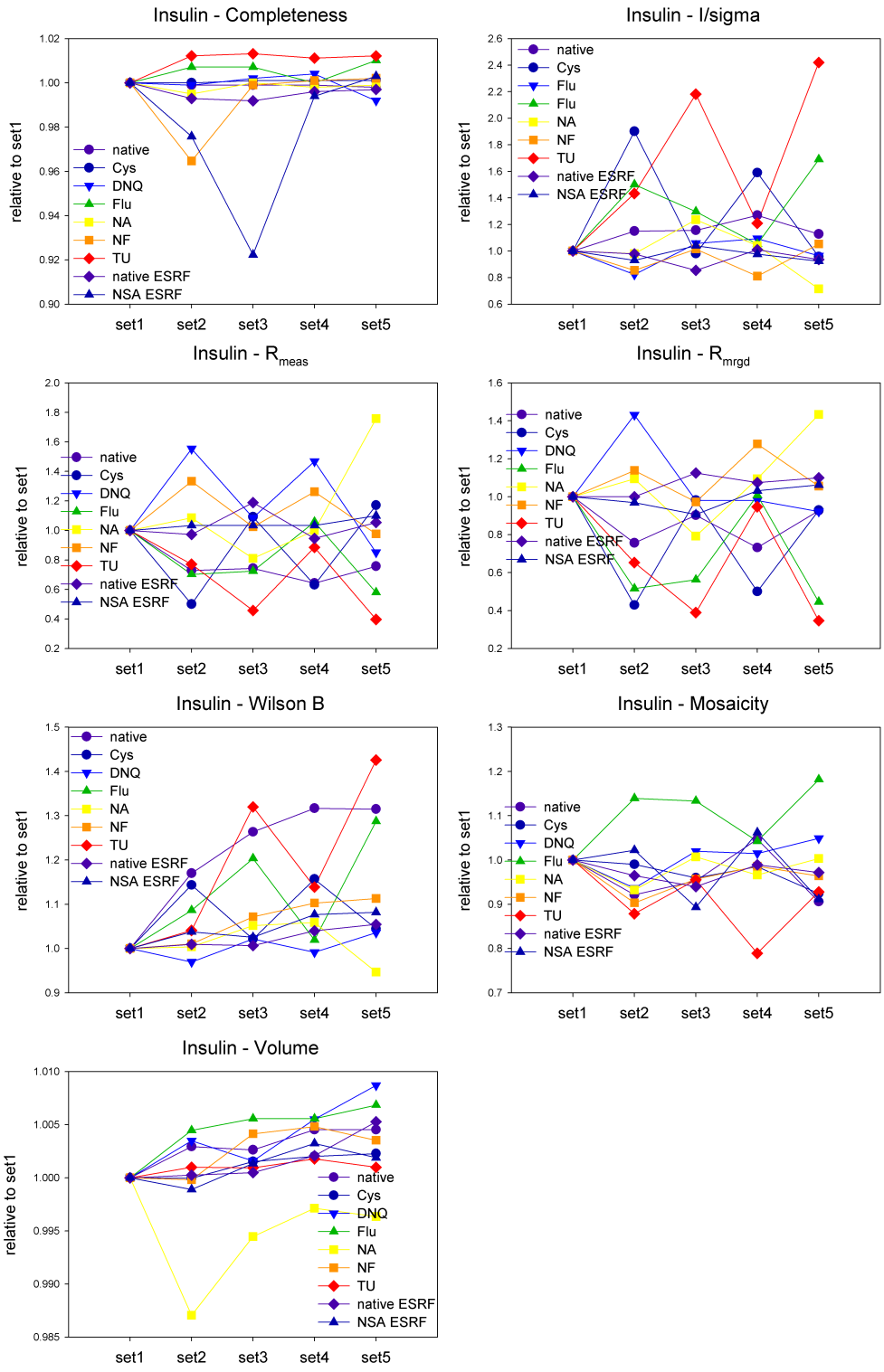
		Flux [ph/sec]	Crystal size	Oscillation [deg]	Exposure [sec]	# img
Azurin	HEP	7.70E+10	0.12x0.1x0.02	1	3	900
	Native	6.47E+10	0.1x0.1x0.02	1	3	900
	NI	7.28E+10	0.15x0.1x0.02	1	3	900
	SAA	5.58E+10	0.1x0.1x0.02	1	3	900
	TU	7.24E+10	0.1x0.08x0.02	1	3	900
	Native	3.22E+10	0.19x0.1x0.01	1	3	900
	Cys	7.10E+10	0.12x0.1x0.01	1	3.5	900
	DNQ	5.85E+10	0.12x0.1x0.02	1	3.5	900
	NA	5.97E+10	0.12x0.08x0.02	1	3.5	900
Myo	Native	2.59E+10	0.16x0.1x0.02	1	4	900
	NSA	3.94E+10	0.1x0.1x0.01	1	4	900
	Native	6.48E+08	0.2x0.15x0.02	0.5	30	9999
	Native	6.48E+08	0.22x0.1x0.02	0.5	30	9999
	NSA	6.48E+08	0.2x0.12x0.02	0.5	30	9999
CuZnSOD	Native	6.29E+10	0.73x0.08x0.06	1	5	900
Insulin	Native	2.42E+10	0.08x0.08x0.02	1	8	834
	NSA	2.94E+10	0.08x0.08x0.02	1	8	539
	Cys	6.37E+10	0.1x0.1x0.04	1	5	950
	DNQ	5.61E+10	0.1x0.1x0.04	1	5	1080
	Flu	8.17E+10	0.12x0.12x0.04	1	5	850
	NA	5.66E+10	0.14x0.14x0.04	1	5	1050
	Native	5.32E+10	0.12x0.12x0.04	1	5	1851
	NF	7.70E+10	0.12x0.12x0.04	1	5	950
	TU	7.69E+10	0.1x0.1x0.04	1	5	1100

7.2 Scaling statistics

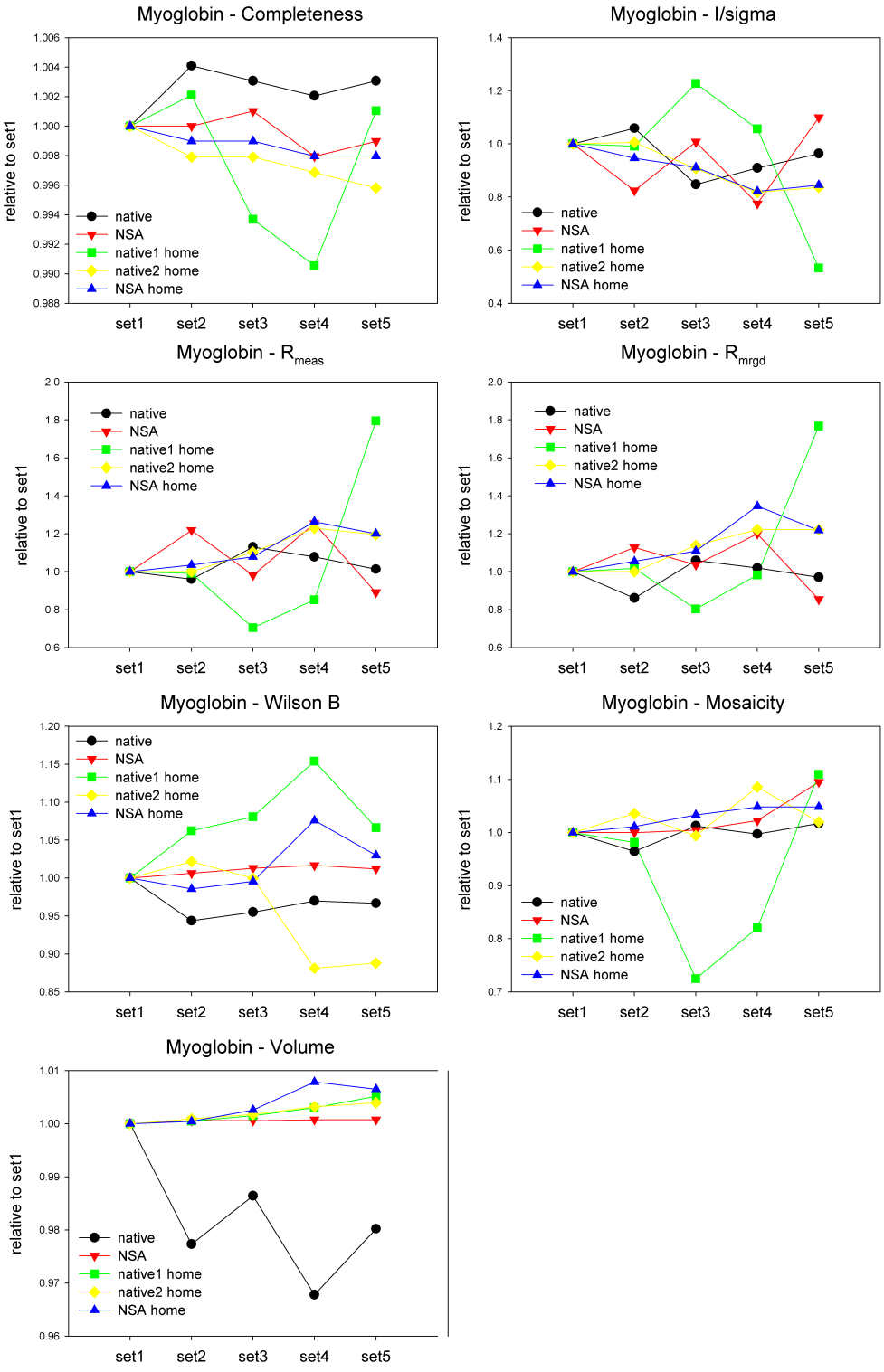
7.2.1 Azurin



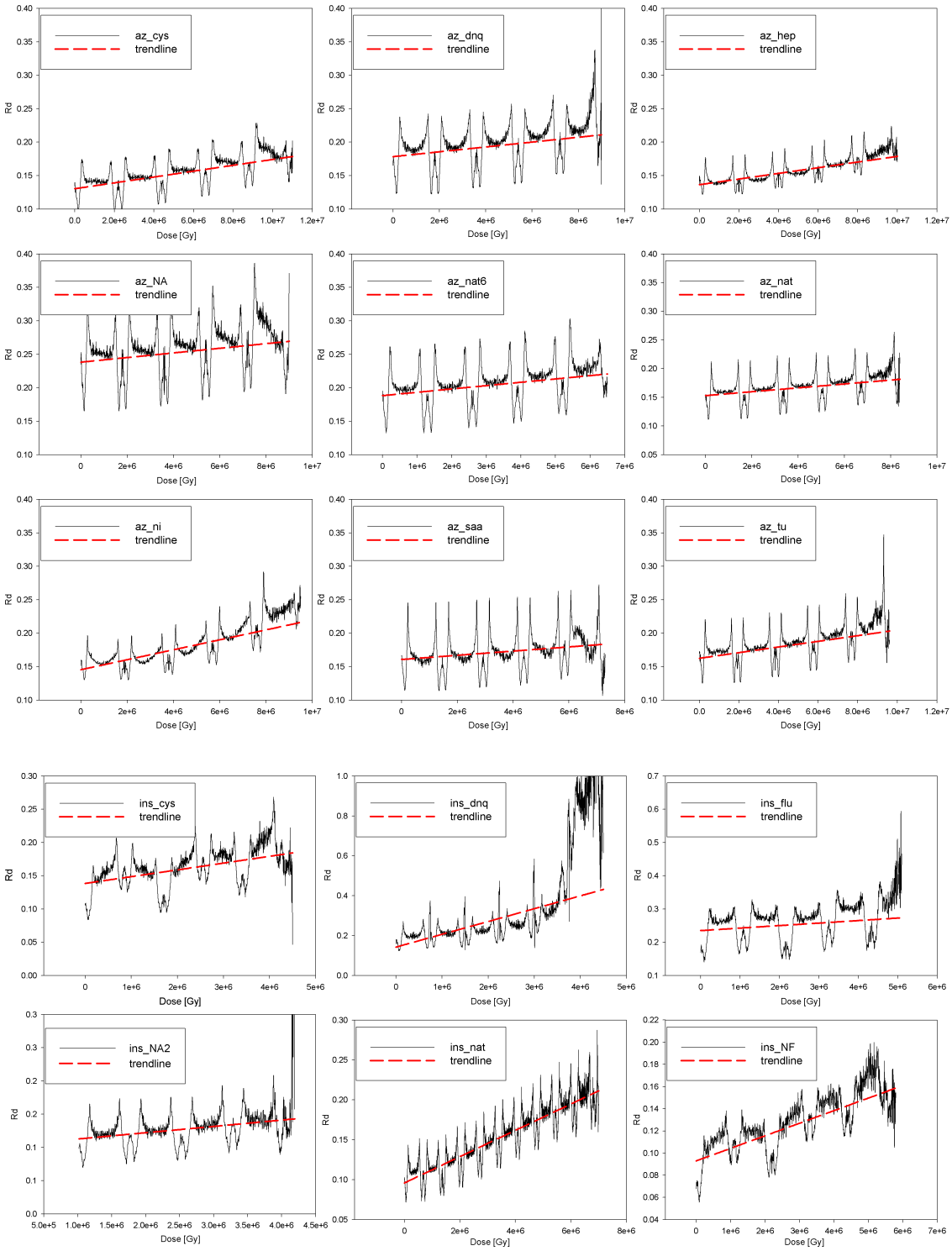
7.2.2 Insulin

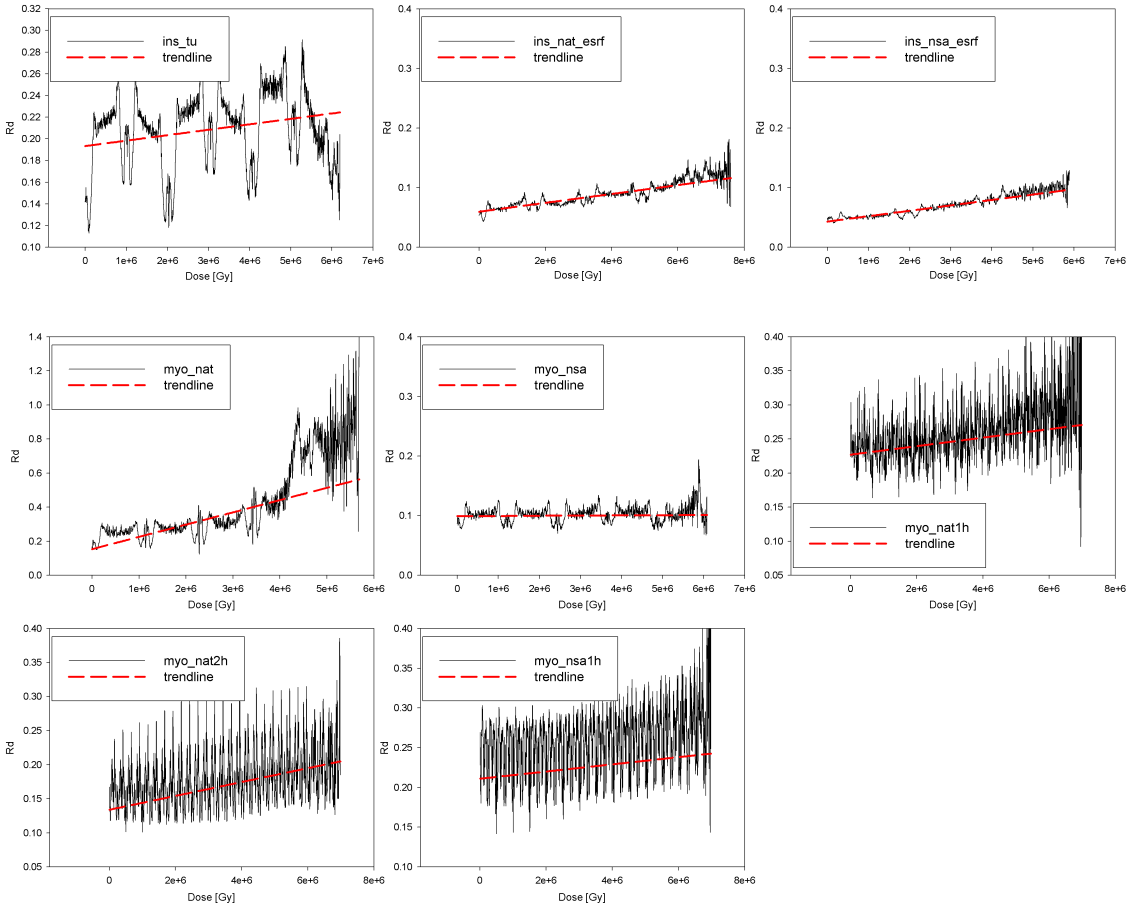

Insulin - Completeness
relative to set1
native
Cys
DNQ
Flu
NA
NF
TU
native ESRF
NSA ESRF
set1
set2
set3
set4
set5
Insulin - I/sigma
Insulin - R meas
Insulin - R mrgd
Insulin - Wilson B
Insulin - Mosaicity
Insulin - Volume

7.2.3 Myoglobin



7.3 Rd plots


az_cys
trendline
az_dnq
trendline
az_hep
trendline
az_NA
trendline
az_nat6
trendline
az_nat
trendline
az_ni
trendline
az_saa
trendline
az_tu
trendline
ins_cys
trendline
ins_dnq
trendline
ins_flu
trendline
ins_NA2
trendline
ins_nat
trendline
ins_NF
trendline
Rd
Dose [Gy]



7.3.1 Rd regression statistics

Myoglobin

f=y0+a*x		native1*	native2*	native	NSA1*	NSA
Coefficient	y0	0.23	0.13	0.15	0.21	0.1
	a	6.24E-009	1.02E-008	7.21E-008	4.52E-009	3.36E-010
Std. Error	y0	0.0020	0.0017	0.0061	0.0024	0.0008
	a	6.91E-010	6.32E-010	2.82E-009	8.26E-010	2.96E-010
t	y0	112.94	77.55	25.08	86.43	132.04
	a	9.03	16.09	25.54	5.48	1.13
P	y0	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	a	<0.0001	<0.0001	<0.0001	<0.0001	0.26
R		0.26	0.44	0.65	0.16	0.04
Rsqr		0.07	0.19	0.42	0.03	0
Adj Rsqr		0.07	0.19	0.42	0.03	0
Std. Err. of Estimate		1.2	2.16	4.28	3.27	0.6

* collected at home

Azurin

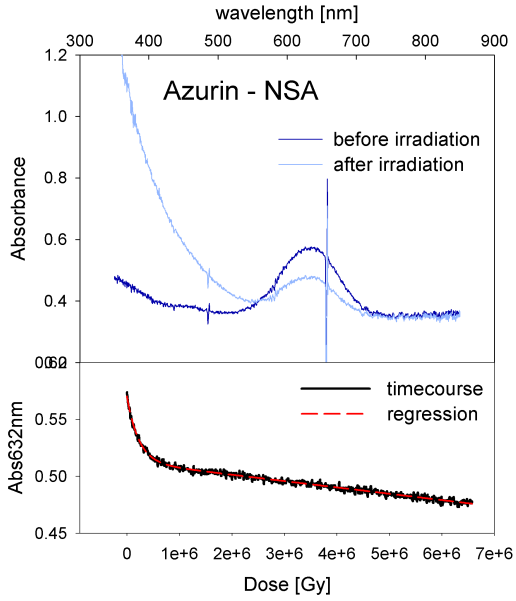
f=y0+a*x		cysteine	DNQ	hepes	NA	native6	native	NI	SAA	TU
Coefficient	y0	0.13	0.18	0.14	0.24	0.19	0.15	0.15	0.16	0.16
	a	4.37E-009	3.62E-009	4.23E-009	3.46E-009	4.98E-009	3.38E-009	7.43E-009	3.13E-009	4.23E-009
Std. Error	y0	0.0011	0.0017	0.0007	0.0022	0.0018	0.0012	0.0010	0.0016	0.0012
	a	2.52E-010	4.56E-010	1.74E-010	5.92E-010	6.63E-010	3.56E-010	2.43E-010	5.18E-010	3.03E-010
t	y0	114.18	105.74	190.02	108.11	105.77	124.63	152.49	103.5	136.24
	a	17.36	7.95	24.29	5.84	7.51	9.51	30.6	6.04	13.97
P	y0	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	a	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R		0.5	0.26	0.63	0.19	0.24	0.3	0.71	0.2	0.42
Rsqr		0.25	0.07	0.4	0.04	0.06	0.09	0.51	0.04	0.18
Adj Rsqr		0.25	0.06	0.4	0.04	0.06	0.09	0.51	0.04	0.18
Std. Err. of Estimate		3.63	5.47	2.09	6.87	5.89	3.62	2.77	4.58	3.46

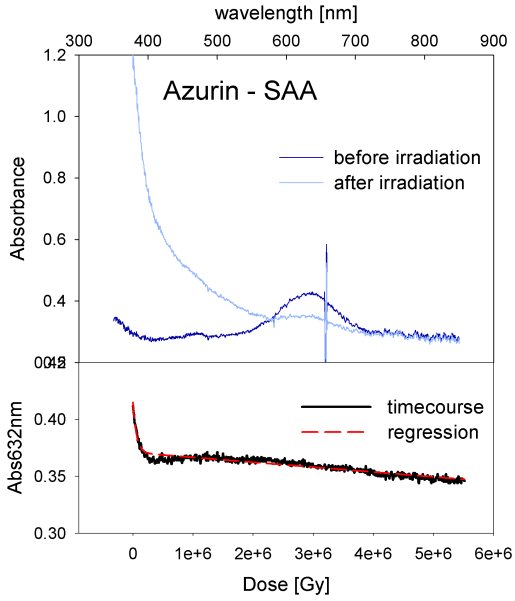
Insulin

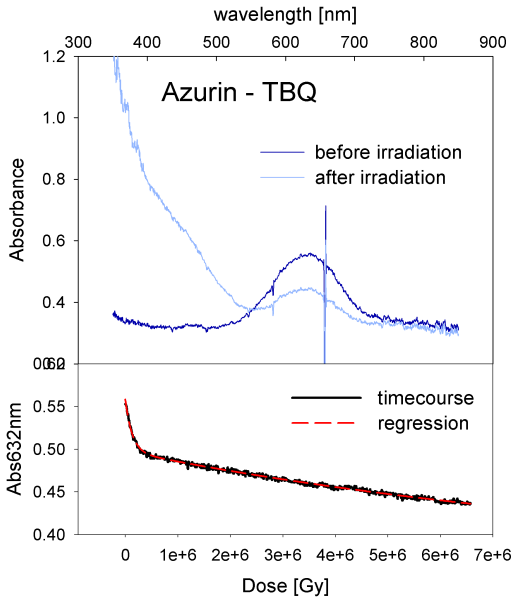
f=y0+a*x		cysteine	DNQ	fluorescein	NA	native	native*	NF	NSA*	TU
Coefficient	y0	0.14	0.14	0.24	0.1	0.1	0.06	0.09	0.04	0.19
	a	1.02E-008	6.42E-008	7.56E-009	9.50E-009	1.65E-008	7.42E-009	1.13E-008	9.07E-009	5.01E-009
Std. Error	y0	0.0015	0.0049	0.0030	0.0024	0.0007	0.0005	0.0009	0.0004	0.0019
	a	8.24E-010	2.78E-009	1.39E-009	1.06E-009	2.54E-010	1.56E-010	3.91E-010	1.73E-010	7.41E-010
t	y0	90.1	28.91	79.68	43.44	132.59	121.58	98.86	101.63	101.71
	a	12.37	23.06	5.44	8.94	64.95	47.64	29	52.5	6.76
P	y0	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	a	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R		0.37	0.58	0.18	0.31	0.83	0.86	0.69	0.91	0.2
Rsqr		0.14	0.33	0.03	0.1	0.7	0.73	0.47	0.84	0.04
Adj Rsqr		0.14	0.33	0.03	0.09	0.7	0.73	0.47	0.84	0.04
Std. Err. of Estimate		2.37	8.29	4	1.72	2.21	0.44	1.45	0.24	3.41

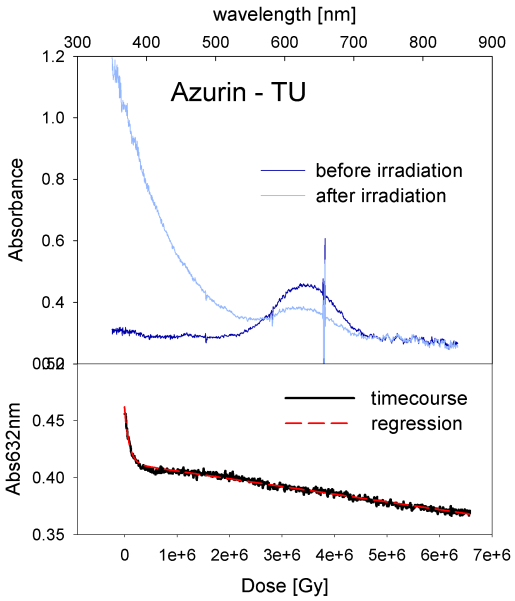
* collected at the ESRF

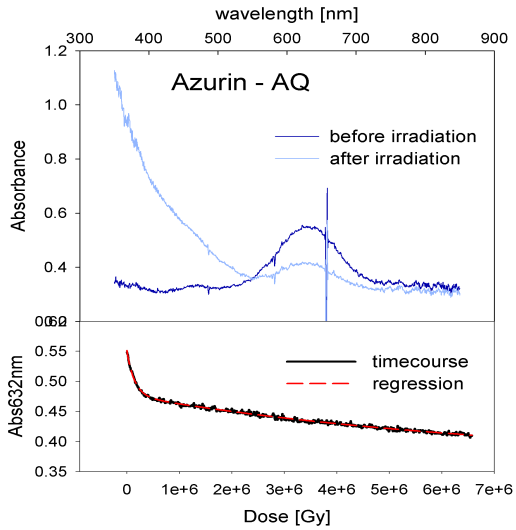
7.4 Microspectrophotometry - Spectra and timecourses

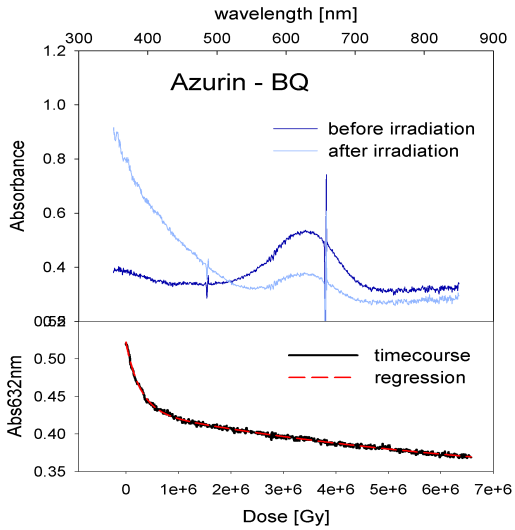


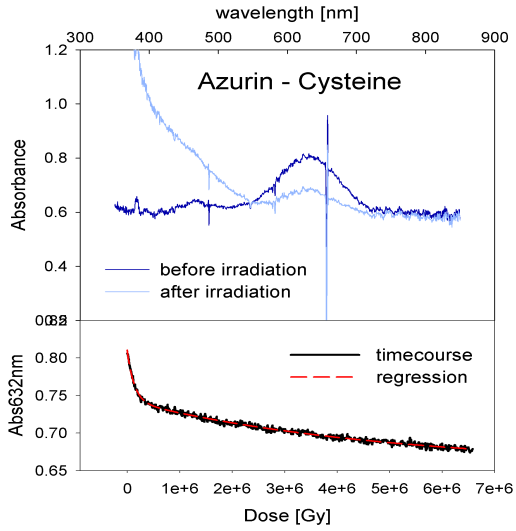


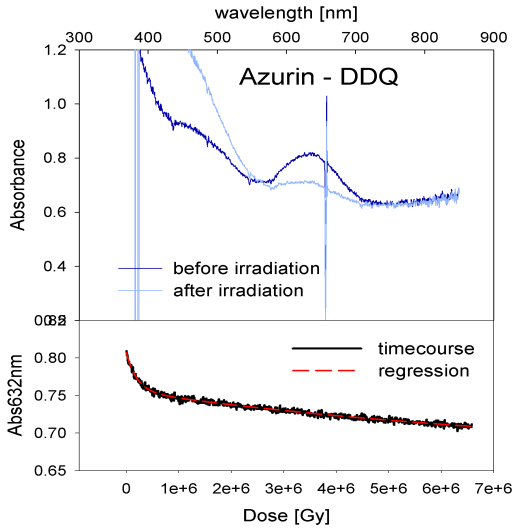


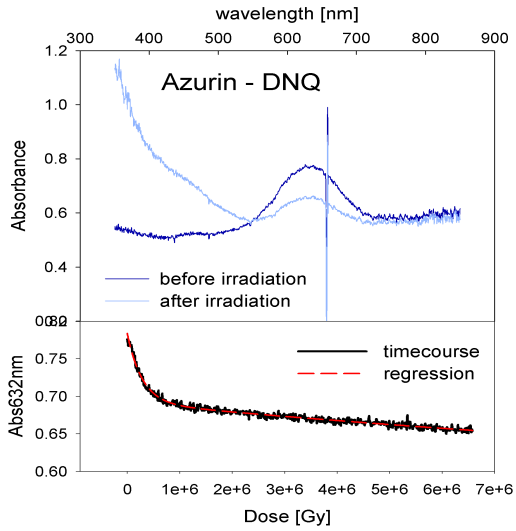


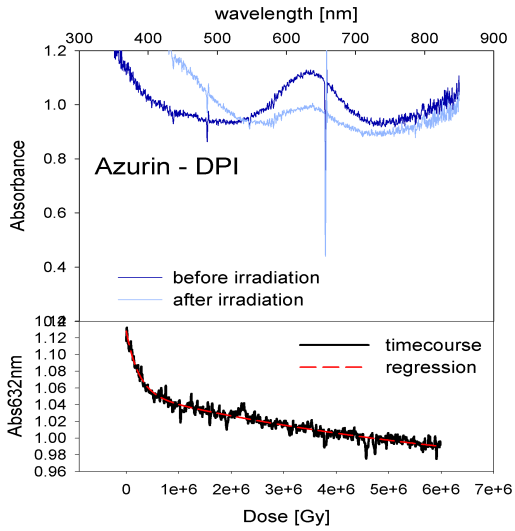

wavelength [nm]
Azurin - AQ
before irradiation
after irradiation
Absorbance
timecourse
regression
Abs632nm
Dose [Gy]

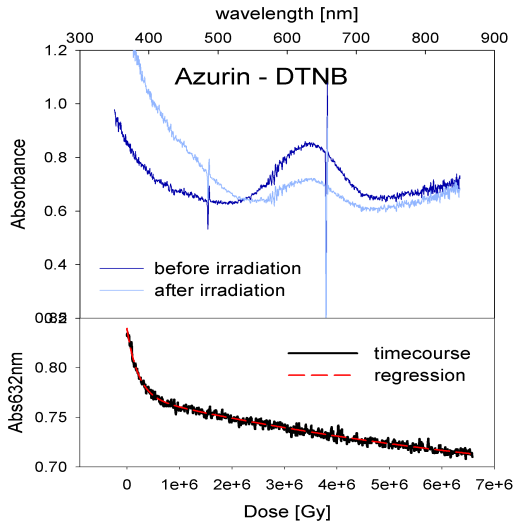

wavelength [nm]
Azurin - BQ
before irradiation
after irradiation
Absorbance
timecourse
regression
Abs632nm
Dose [Gy]

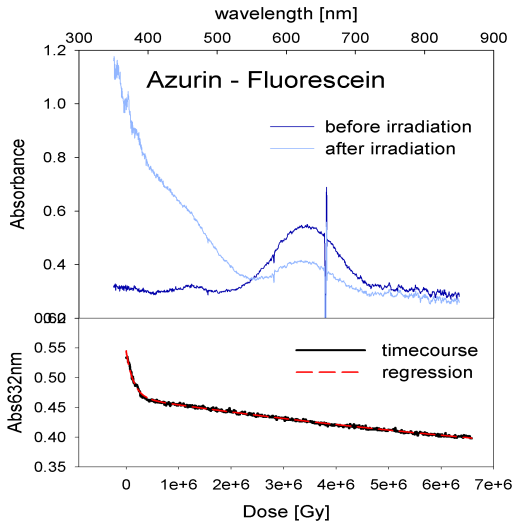

wavelength [nm]
Azurin - Cysteine
before irradiation
after irradiation
Absorbance
timecourse
regression
Abs632nm
Dose [Gy]

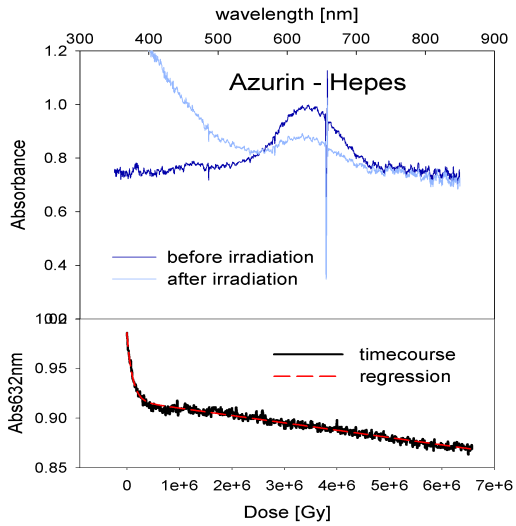

wavelength [nm]
Azurin - DDQ
before irradiation
after irradiation
Absorbance
timecourse
regression
Abs632nm
Dose [Gy]

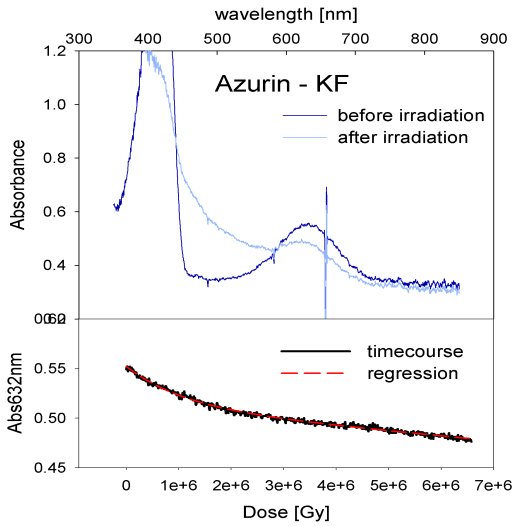

wavelength [nm]
Azurin - DNQ
before irradiation
after irradiation
Absorbance
timecourse
regression
Abs632nm
Dose [Gy]

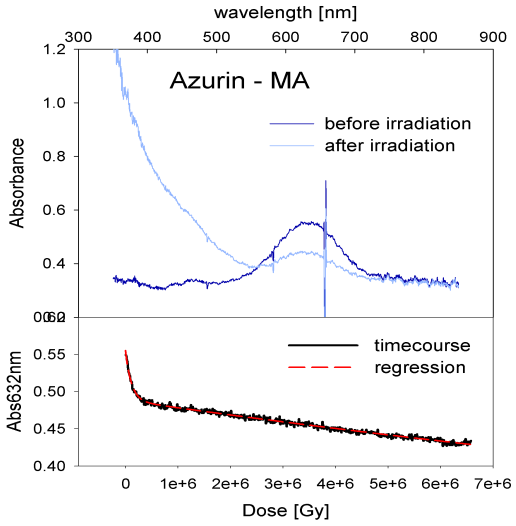

wavelength [nm]
Azurin - DPI
before irradiation
after irradiation
Absorbance
timecourse
regression
Abs632nm
Dose [Gy]

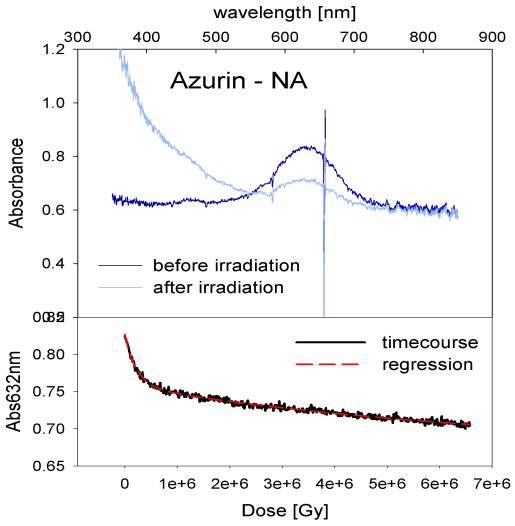

wavelength [nm]
Azurin - DTNB
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

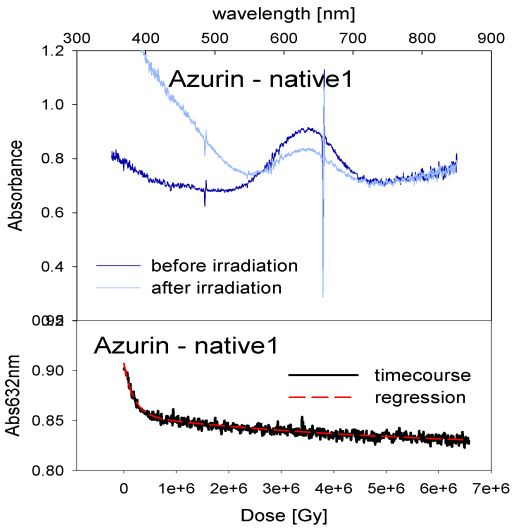

wavelength [nm]
Azurin - Fluorescein
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

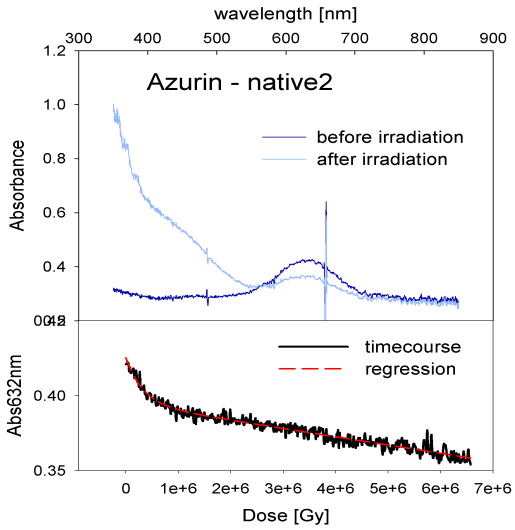

wavelength [nm]
Azurin - Hepes
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

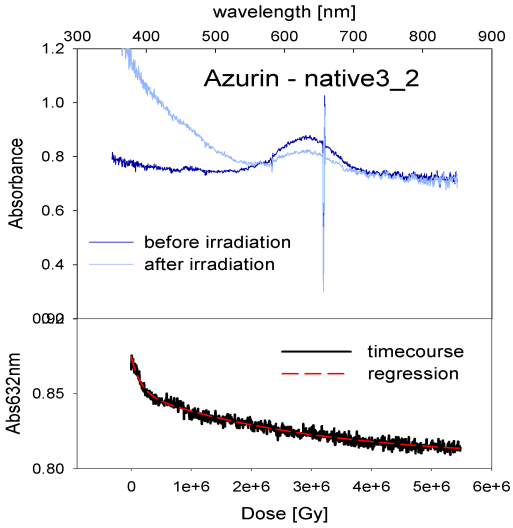

wavelength [nm]
Azurin - KF
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

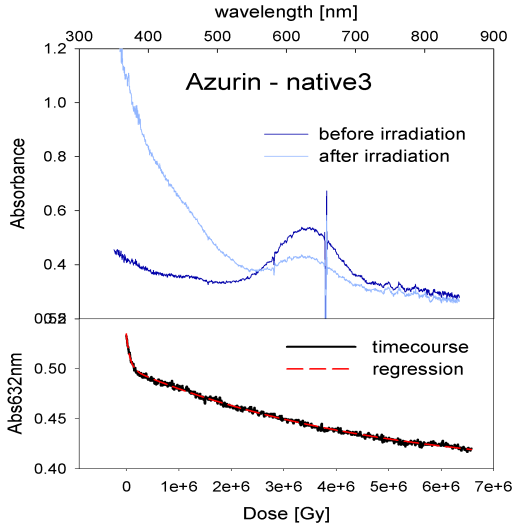

wavelength [nm]
Azurin - MA
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

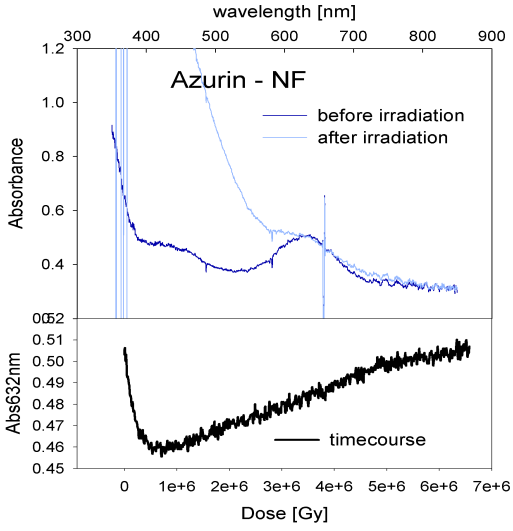

wavelength [nm]
Azurin - NA
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

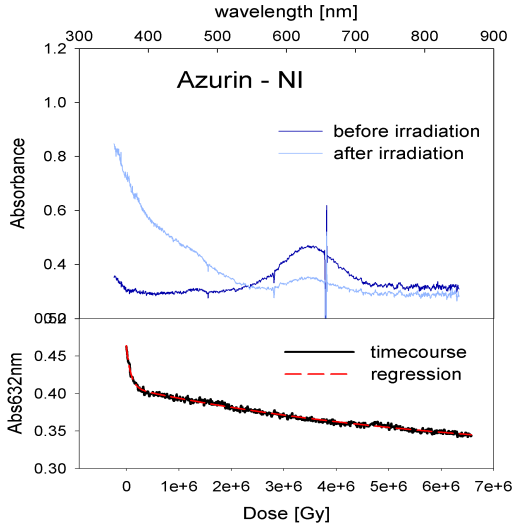

wavelength [nm]
Azurin - native1
Absorbance
before irradiation
after irradiation
Azurin - native1
Abs632nm
timecourse
regression
Dose [Gy]

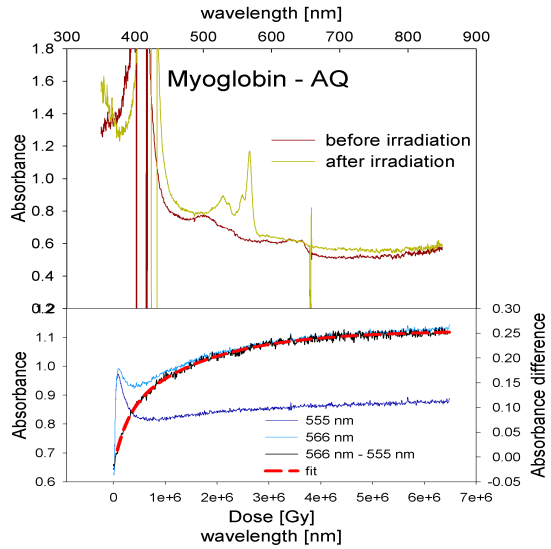

wavelength [nm]
Azurin - native2
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

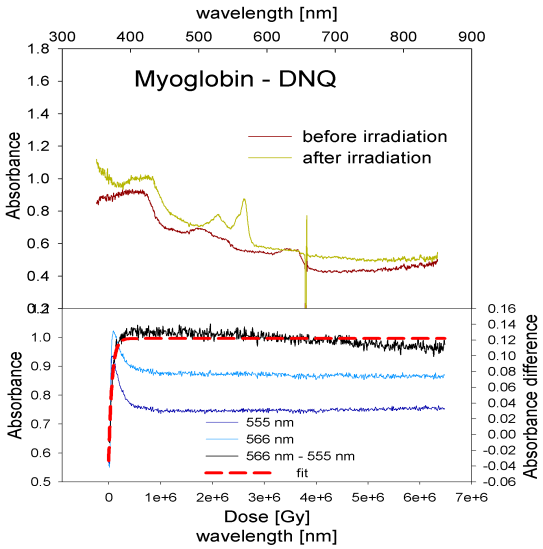

wavelength [nm]
Azurin - native3_2
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

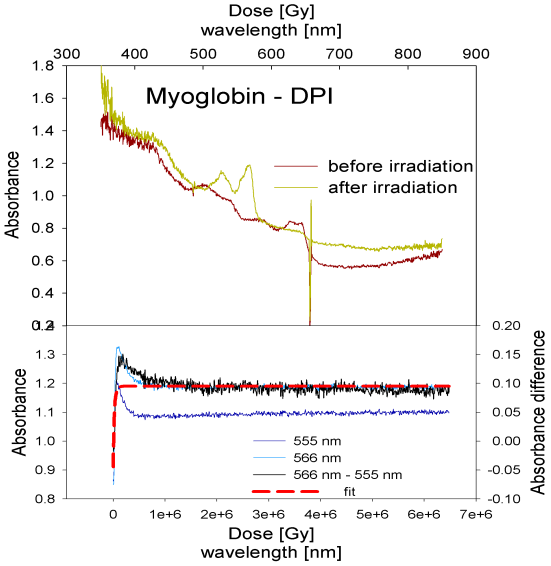

wavelength [nm]
Azurin - native3
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

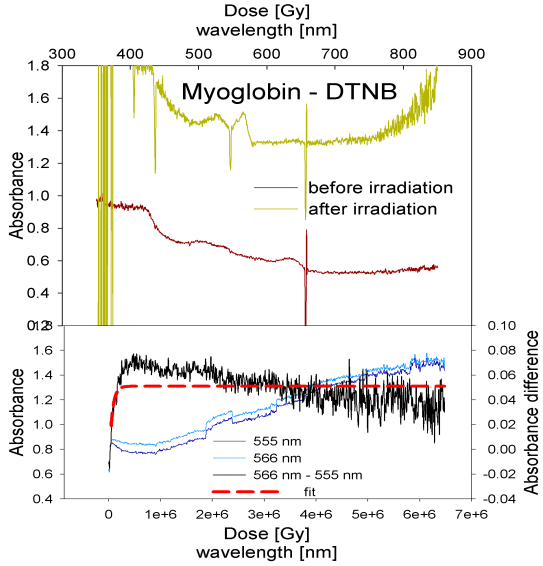

wavelength [nm]
Azurin - NF
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
Dose [Gy]

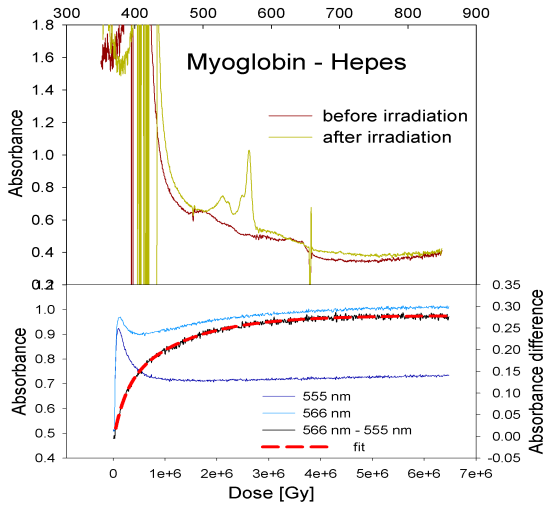

wavelength [nm]
Azurin - NI
Absorbance
before irradiation
after irradiation
Abs632nm
timecourse
regression
Dose [Gy]

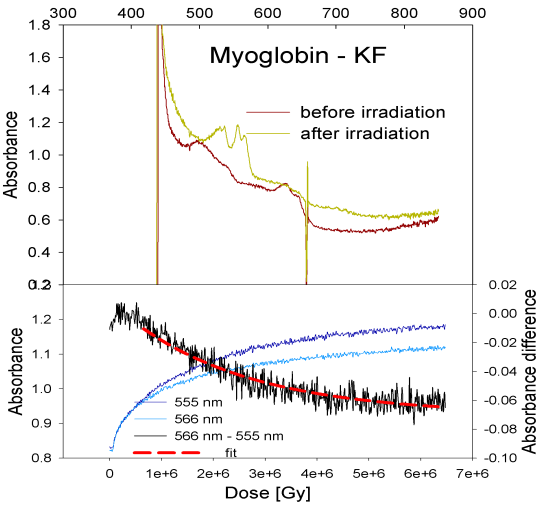

wavelength [nm]
Myoglobin - AQ
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

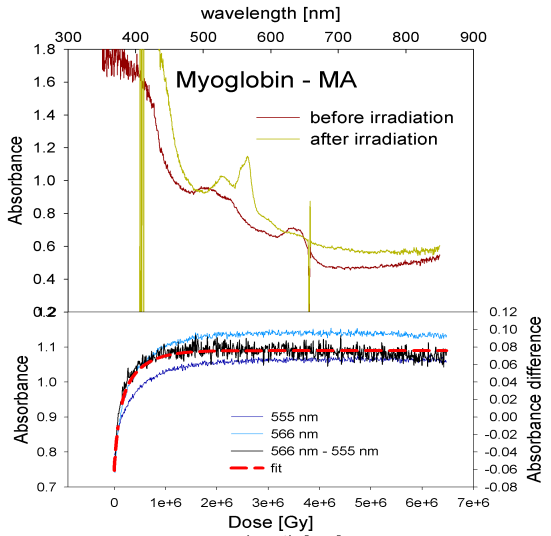

wavelength [nm]
Myoglobin - DNQ
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

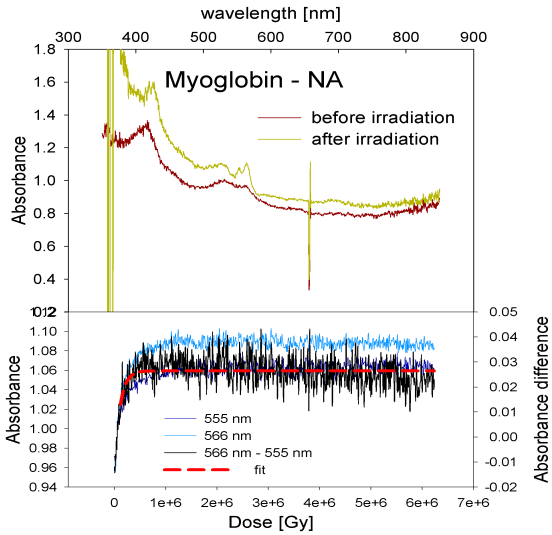

wavelength [nm]
Myoglobin - DPI
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

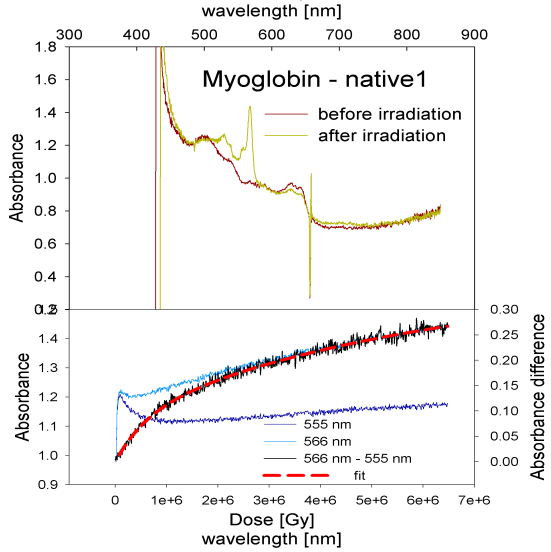

wavelength [nm]
Myoglobin - DTNB
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

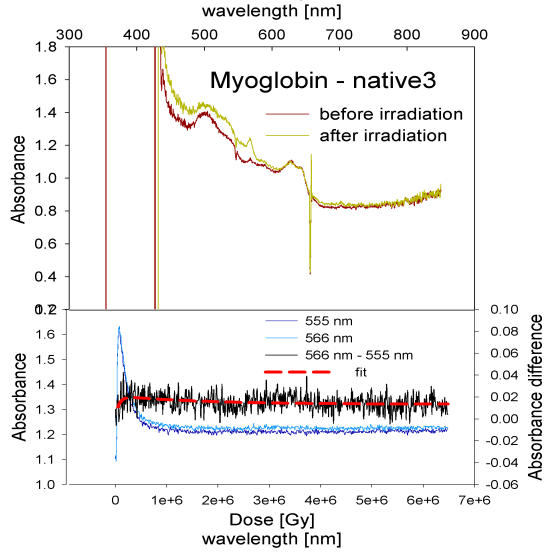

wavelength [nm]
Myoglobin - Hepes
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

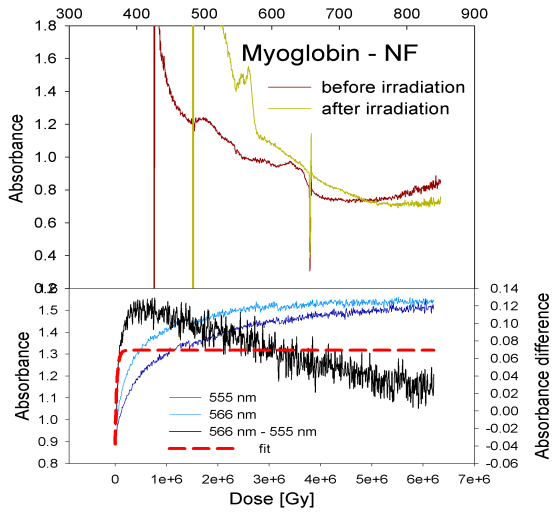

wavelength [nm]
Myoglobin - KF
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

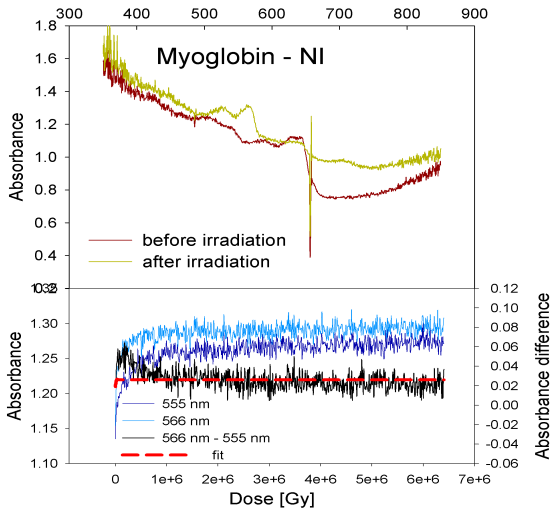

wavelength [nm]
Myoglobin - MA
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

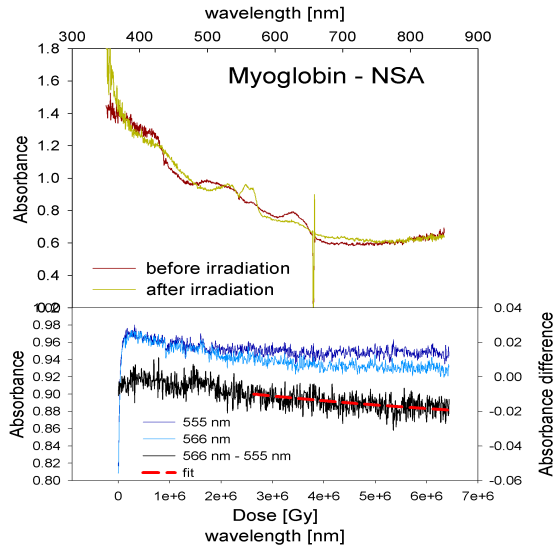

wavelength [nm]
Myoglobin - NA
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

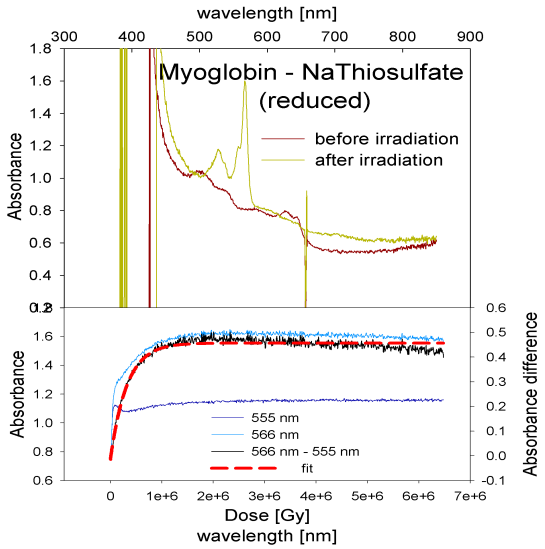

wavelength [nm]
Myoglobin - native1
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

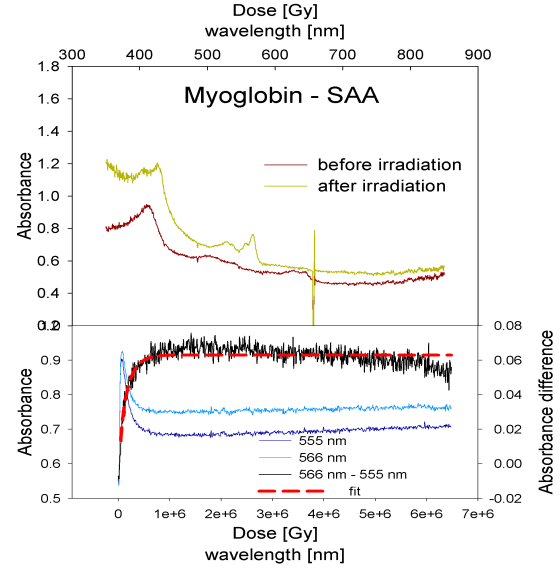

wavelength [nm]
Myoglobin - native3
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

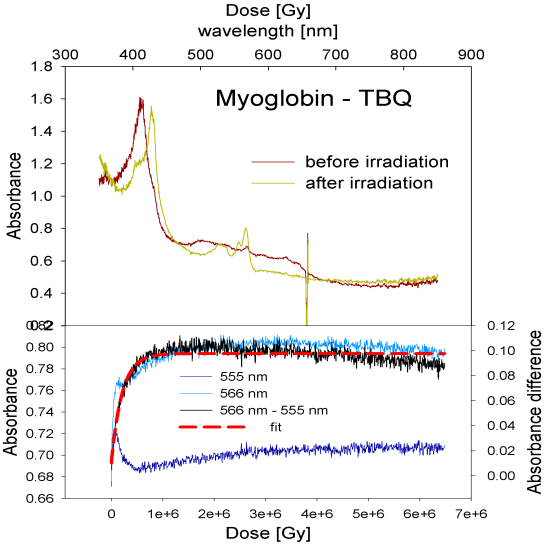

wavelength [nm]
Myoglobin - NF
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

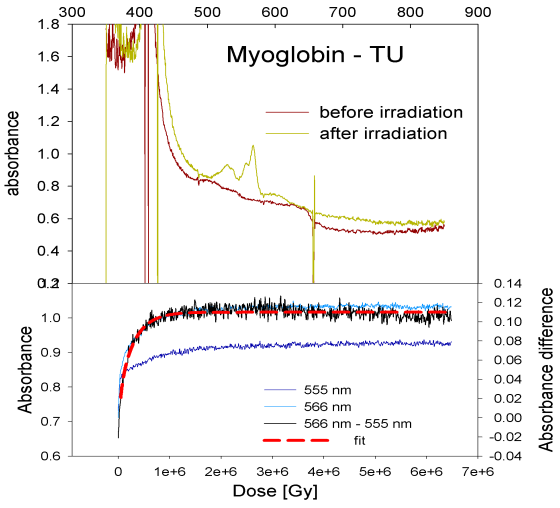

wavelength [nm]
Myoglobin - NI
before irradiation
after irradiation
Absorbance
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]


wavelength [nm]
Myoglobin - NSA
Absorbance
before irradiation
after irradiation
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]


wavelength [nm]
Myoglobin - NaThiosulfate (reduced)
Absorbance
before irradiation
after irradiation
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]


wavelength [nm]
Myoglobin - SAA
Absorbance
before irradiation
after irradiation
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]


wavelength [nm]
Myoglobin - TBQ
Absorbance
before irradiation
after irradiation
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]


wavelength [nm]
Myoglobin - TU
absorbance
before irradiation
after irradiation
555 nm
566 nm
566 nm - 555 nm
fit
Absorbance difference
Dose [Gy]

7.5 Timecourse regression statistics

7.5.1 Azurin

f=y0+a*exp(-b*x)+c*exp(-d*x)		AQ	BQ	cys	ddq	dnq	dpi	dtnb	flu	ma	na	nat1	nat3_2	nat3	ni	tbq	hep	KF	nat2	nsa	saa	tu
coefficients	y0	0.3665	0.3250	0.6567	0.6688	0.5612	0.9380	0.6608	0.2981	0.2074	0.6705	0.8212	0.8063	0.3929	0.3103	0.3662	-261.58	-78.92	-68.56	-4.75	-256.84	-52.34
	a	0.0748	0.0911	0.0676	0.0512	0.0925	0.0746	0.0680	0.0783	0.0672	0.0691	0.0522	0.0238	0.0342	0.0560	0.0611	0.0687	0.0406	0.0308	0.0580	0.0442	0.0498
	b	6.06E-6	3.51E-6	5.92E-6	4.47E-6	3.42E-6	3.71E-6	4.25E-6	6.07E-6	7.49E-6	3.94E-6	4.41E-6	3.10E-6	1.15E-5	1.07E-5	6.77E-6	8.20E-6	7.60E-7	2.79E-6	4.27E-6	1.15E-5	9.01E-6
	c	0.1103	0.1064	0.0863	0.0873	0.1305	0.1156	0.1111	0.1689	0.2806	0.0874	0.0340	0.0446	0.1079	0.0976	0.1312	262.4940	79.4340	68.9594	5.2646	257.2070	52.7537
	d	1.22E-7	1.25E-7	1.61E-7	1.07E-7	4.89E-8	1.18E-7	1.03E-7	6.82E-8	3.06E-8	1.21E-7	1.62E-7	1.68E-7	1.68E-7	1.43E-7	8.38E-8	2.47E-11	5.24E-11	7.40E-11	1.01E-9	1.14E-11	1.02E-10
Std. Error	y0	0.0036	0.0041	0.0017	0.0050	0.0465	0.0143	0.0074	0.0091	0.0492	0.0055	0.0028	0.0010	0.0011	0.0028	0.0061	8204.13	6613.57	7680.75	59.62	38482.38	3167.43
	a	0.0011	0.0008	0.0012	0.0009	0.0013	0.0024	0.0012	0.0009	0.0012	0.0013	0.0014	0.0010	0.0012	0.0015	0.0010	0.0017	0.0008	0.0011	0.0008	0.0017	0.0012
	b	1.70E-7	6.31E-8	2.01E-7	1.59E-7	9.65E-8	2.44E-7	1.54E-7	1.33E-7	2.37E-7	1.48E-7	2.41E-7	2.63E-7	6.86E-7	4.99E-7	1.93E-7	3.22E-7	2.86E-8	1.90E-7	1.17E-7	7.17E-7	3.52E-7
	c	0.0032	0.0036	0.0013	0.0046	0.0457	0.0129	0.0068	0.0088	0.0489	0.0048	0.0022	0.0006	0.0009	0.0025	0.0058	8204.13	6613.57	7680.75	59.62	38482.38	3167.43
	d	6.39E-9	7.79E-9	6.24E-9	9.35E-9	2.11E-8	2.28E-8	1.04E-8	4.83E-9	6.08E-9	1.26E-8	2.58E-8	9.23E-9	3.28E-9	6.86E-9	5.38E-9	7.72E-10	4.36E-9	8.24E-9	1.15E-8	1.71E-9	6.10E-9
t	y0	102.3	79.8	389.9	133.4	12.1	65.5	88.9	32.8	4.2	121.8	295.2	840.4	366.0	110.3	60.4	-0.0319	-0.0119	-0.0089	-0.0797	-0.0067	-0.0165
	a	66.2159	115.0860	54.4453	54.9399	73.3227	31.3963	54.6935	82.9843	55.0064	53.5961	36.0481	23.7231	28.8776	37.4305	63.0965	40.5305	49.6605	27.4657	71.1856	25.4496	43.0562
	b	35.6561	55.6155	29.4654	28.0053	35.4101	15.2279	27.5493	45.7026	31.6511	26.5420	18.2728	11.7981	16.7746	21.4841	35.0219	25.4868	26.5291	14.6878	36.4038	16.0732	25.6118
	c	34.3315	29.8551	65.0162	19.0343	2.8567	8.9901	16.2796	19.2116	5.7440	18.0943	15.7684	74.8769	122.3931	38.9303	22.6915	0.0320	0.0120	0.0090	0.0883	0.0067	0.0167
	d	19.1209	16.0395	25.8320	11.4258	2.3144	5.1584	9.9175	14.1217	5.0371	9.6770	6.2631	18.2330	51.2594	20.8185	15.5844	0.0320	0.0120	0.0090	0.0880	0.0067	0.0166
p	y0	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.9746	0.9905	0.9929	0.9365	0.9947	0.9868
	a	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	b	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
	c	<0.0001	<0.0001	<0.0001	<0.0001	0.0044	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.9745	0.9904	0.9928	0.9297	0.9947	0.9867
	d	<0.0001	<0.0001	<0.0001	<0.0001	0.0210	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.9745	0.9904	0.9928	0.9299	0.9947	0.9867
R		0.9960	0.9976	0.9942	0.9910	0.9895	0.9771	0.9918	0.9969	0.9938	0.9920	0.9633	0.9815	0.9974	0.9936	0.9964	0.9858	0.9952	0.9790	0.9929	0.9555	0.9911
Rsqr		0.9920	0.9952	0.9885	0.9822	0.9792	0.9548	0.9837	0.9939	0.9876	0.9841	0.9280	0.9634	0.9949	0.9872	0.9927	0.9718	0.9905	0.9585	0.9859	0.9130	0.9823
Adj Rsqr		0.9920	0.9952	0.9884	0.9821	0.9790	0.9545	0.9836	0.9939	0.9875	0.9840	0.9276	0.9633	0.9948	0.9871	0.9927	0.9716	0.9904	0.9582	0.9858	0.9124	0.9822
Std Err of Estimate		0.0021	0.0019	0.0024	0.0024	0.0030	0.0055	0.0029	0.0019	0.0023	0.0027	0.0032	0.0024	0.0017	0.0023	0.0018	0.0029	0.0018	0.0028	0.0018	0.0026	0.0020

7.5.2 Myoglobin

f=y0+a*exp(-b*x)+c*exp(-d*x)		AQ	DNQ	DPI	DTNB	hepes	KF	MA	NA	native1	native3	NF	NI	NSA	reduced	SAA	TBQ	TU
coefficients	y0	0.26	0.12	0.09	0.05	0.28	-0.07	0.08	0.03	0.36	0.01	0.07	0.03	-0.03	0.46	0.06	0.1	0.11
	a	-0.1	-0.09	-0.08	-0.05	-0.11	0.04	-0.07	-0.02	-0.08	-0.02	-0.06	-0.01	0.02	-0.25	-0.02	-0.04	-0.04
	b	4.11E-6	1.61E-5	4.21E-5	1.72E-5	4.84E-6	4.19E-7	1.15E-5	9.33E-6	1.73E-6	1.39E-5	2.94E-5	7.60E+2	1.36E-7	3.15E-6	1.92E-5	4.83E-6	2.23E-5
	c	-0.18	-0.07	-0.06	-0.03	-0.18	0.04	-0.07	-0.02	-0.27	0.01	-0.05	0	0.01	-0.22	-0.06	-0.04	-0.09
	d	6.16E-7	1.61E-5	4.21E-5	1.72E-5	8.39E-7	4.19E-7	2.49E-6	9.33E-6	1.72E-7	5.95E-7	2.94E-5	5.21E-17	1.36E-7	3.15E-6	5.84E-6	4.83E-6	3.53E-6
Std. Error	y0	0	0	0	0	0	0	0	0	0.02	0	0	(NAN)	0.08	0	0	0	0
	a	0	92355.37	215610.3	144233.41	0	50994.39	0.01	59742.87	0.01	0	430510.7	(NAN)	11085.57	76473.2	0.02	19646.43	0.01
	b	2.22E-7	6.73E-2	6.33E-1	1.76E-1	2.34E-7	4.50E-3	2.51E-6	1.97E-2	1.81E-7	6.18E-6	6.14E-1	(NAN)	4.70E-3	1.86E-2	1.75E-5	3.00E-2	7.10E-6
	c	0	92355.37	215610.3	144233.41	0	50994.39	0.01	59742.87	0.01	0	430510.7	(NAN)	11085.57	76473.2	0.02	19646.43	0.01
	d	1.33E-8	8.38E-2	8.21E-1	2.64E-1	1.47E-8	5.60E-3	2.63E-7	2.54E-2	2.19E-8	2.87E-7	8.42E-1	(NAN)	5.60E-3	2.06E-2	1.17E-6	2.96E-2	2.04E-7
t	y0	437.59	386.74	194.71	97.5	824.17	-14.64	247.54	114.73	21.49	16.53	58.47	(+inf)	-0.41	602.84	320.02	458.49	493.55
	a	-32.83	-9.32E-7	-3.65E-7	-3.17E-7	-34.52	8.48E-7	-6.52	-3.91E-007	-12.75	-4.12	-1.44E-7	(-inf)	1.61E-6	-3.24E-6	-1.04	-2.20E-6	-5.63
	b	18.58	0	6.64E-5	9.76E-5	20.66	9.23E-5	4.58	0	9.53	2.24	4.79E-5	(+inf)	2.88E-5	0	1.1	0	3.14
	c	-63.89	-7.48E-7	-2.81E-7	-2.12E-7	-58.57	6.89E-7	-6.69	-3.07E-007	-27.1	4.79	-1.05E-7	(-inf)	1.34E-6	-2.92E-6	-2.67	-2.23E-6	-14.48
	d	46.2	0	5.12E-5	6.52E-5	57.08	7.50E-5	9.44	0	7.83	2.07	3.49E-5	(+inf)	2.41E-5	0	4.97	0	17.32
p	y0	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.68	<0.0001	<0.0001	<0.0001	<0.0001
	a	<0.0001	1	1	1	<0.0001	1	<0.0001	1	<0.0001	<0.0001	1	<0.0001	1	1	0.3	1	<0.0001
	b	<0.0001	1	1	1	<0.0001	1	<0.0001	1	<0.0001	0.03	1	<0.0001	1	1	0.27	1	0
	c	<0.0001	1	1	1	<0.0001	1	<0.0001	1	<0.0001	<0.0001	1	<0.0001	1	1	0.01	1	<0.0001
	d	<0.0001	1	1	1	<0.0001	1	<0.0001	1	<0.0001	0.04	1	<0.0001	1	1	<0.0001	1	<0.0001
R		0.9965	0.8252	0.5026	0.3856	0.9975	0.9576	0.9438	0.5850	0.9940	0.2590	0.2039	0.0285	0.7691	0.9753	0.8464	0.8944	0.9447
Rsqr		0.9930	0.6809	0.2527	0.1487	0.9950	0.9170	0.8907	0.3422	0.9880	0.0671	0.0416	0.0008	0.5916	0.9512	0.7163	0.7999	0.8924
Adj Rsqr		0.9930	0.6790	0.2481	0.1436	0.9950	0.9164	0.8899	0.3381	0.9879	0.0613	0.0359	0.0000	0.5893	0.9509	0.7147	0.7989	0.8918
Std Err. of Estimate		0.0046	0.0079	0.0124	0.0132	0.0039	0.0061	0.0059	0.0055	0.0072	0.0073	0.0305	0.0098	0.0046	0.0162	0.0048	0.0054	0.0052

Zusammenfassung

Die makromolekulare Kristallographie ist eine leistungsfähige Methode zur Strukturaufklärung und basiert auf der Wechselwirkung intensiver Röntgenstrahlung mit kristallinem biologischem Material. Diese Wechselwirkung führt unter anderem auch zur Freisetzung von Elektronen aus den Atomhüllen und damit zu Ionisierungen, Entstehung von Radikalen und der Lösung chemischer Bindungen. Während sich so Schäden in den einzelnen Molekülen immer mehr anhäufen kommt es gleichzeitig zu einer fortlaufenden Zersetzung des Kristallgitters.

Das Ausmaß der Schädigung wird drastisch vermindert, wenn Experimente bei Temperaturen um 100 K durchgeführt werden, was daher heute zur Standardprozedur gehört und das Problem viele Jahre in Schach halten konnte. Mit den neuen starken Undulator-Strahlen in Synchrotronen der dritten Generation, entwickeln sich Strahlenschäden jedoch wieder zu einer bedeutenden Hürde, wenn es darum geht maximale Strukturinformation aus einem einzelnen Kristall zu gewinnen.

Der Einsatz von Elektronen- oder Radikalfängern zum Schutz der Biomoleküle gegen reaktive Spezies ist bereits in der Elektronenspektroskopie dokumentiert und mittlerweile gibt es auch in der Röntgenkristallographie vielversprechende Experimente.

Abgesehen von den gut dokumentierten globalen und strukturellen Auswirkungen, kommt es bei Metalloenzymen zusätzlich zur Reduktion des Metalls. Während sich der Oxidationszustand nicht eindeutig aus der Elektronendichte bestimmen lässte, kann er doch entscheidende Auswirkungen auf die Koordination des Metalls und damit seine strukturelle Umgebung haben.

In der vorliegenden Arbeit haben wir verschiedene Radikalfänger auf ihren möglichen Schutzeffekt für Metalloproteine hin untersucht. Kristalle wurden mit Lösungen ausgewählter Substanzen durchtränkt und anschließend Diffraktionsexperimente am ESRF (European Synchrotron Radiation Facility, Grenoble), bei der Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung (BESSY) und am institutseigenen Rotationsanoden-Instrument durchgeführt. Die gesammelten Datensets wurden in fünf aufeinanderfolgende Subsets mit steigender Strahlendosis aufgeteilt, anhand derer die kumulativen lokalen und globalen Strahlenschäden evaluiert werden konnten. Ein besonderer Schwerpunkt lag dabei auf dem Oxidationszustand des Metalls, der mit Hilfe eines UV-vis Mikrospektrophotometers genau festgestellt werden konnte.

Unsere Ergebnisse weisen darauf hin, dass Radikalfänger wahrscheinlich nicht im Stande sind, die Reduktion eines Metalls wesentlich zu verringern. DNQ, HEPES und das schon als wirksamer Scavenger beschriebene Askorbat kann möglicherweise eine schützende Wirkung zugeschrieben werden und sie bieten sich für weiterführende Tests an.

Abstract

The elucidation of high resolution macromolecular structures depends largely on the application of brilliant X-ray radiation to biological matter in crystalline form, which is the basis of macromolecular crystallography. When X-rays interact with organic matter a plethora of electrons is ejected from atomic shells leading to ionizations, formation of radicals and breakage of bonds. In summary, this leads to accumulating damage of protein molecules and as a consequence to progressing deterioration of the crystal lattice.

The rate of damage is reduced dramatically when experiments are done at 100K in a nitrogen stream, which has become a standard technique and kept the problem under control for many years. Today, however, in strong undulator beamlines at third generation synchrotrons radiation damage is re-emerging as a major limiting factor to the amount of information that can be obtained from one crystal.

The use of electron or radical scavengers to reduce the load of reactive species and thereby protect biological macromolecules has been described for electron spectroscopy and in analogy promising experiments have been conducted in macromolecular X-ray crystallography.

Apart from the well-described global and structural damage, enzymes containing metal co-factors run an additional risk. It has been shown that reduction of the metal is one of the first events happening. While the redox state can not be determined from the electron density measured experimentally, it can have critical impact on the metal's structural environment.

In the present work the influence of a range of possible electron scavengers on the damage suffered by crystals of model metalloproteins was investigated. For that purpose, proteins were crystallized and soaked in scavenger solutions. Data was collected from azurin, insulin and myoglobin at three different stations: the ESRF (European Synchrotron Radiation Facility, Grenoble), at BESSY (Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung) and at the rotating anode home source present at the institute. After processing the data in five consecutive subsets of increasing dose, local and global effects of radiation damage were evaluated. A special focus was given to the metal oxidation state, which was determined by UV-vis microspectrophotometry experiments.

Our results suggest that scavengers are not able to prevent the reduction of the metal. Nonetheless, DNQ, hepes and the previously reported scavenger ascorbate, might have a small protective effect and would be candidates for further tests.

Curriculum Vitae

Name:	Maria Pechlaner
Adresse:	Uchatiusgasse 5/4, 1030 Wien
Geburtsdatum:	28.7.1984
Geburtsort:	Innsbruck

1990-1994	Volksschule Hötting West, Innsbruck
1994-2002	Bundesrealgymnasium Adolf-Pichler-Platz, Innsbruck
2002-2008	Studium der Molekularen Biologie an der Universität Wien