9

Structure-Based Kinetics by Time-Resolved X-ray Crystallography

M. Schmidt

9.1 Introduction

Without catalysis there would be no life. All catalysts, also the common catalytic converter in a car, need an active surface where the substrate molecules align specifically to enhance chemical interaction in order to react to product. Catalysis is a multistep process. Binding of substrate, reaction, and dissociation of product are the elementary steps. In a more complex catalyst such as an enzyme, these steps require substantial reorganization of the structure of the catalyst itself on the atomic level. Each step may be accompanied by a multiple of different configurations and conformations. It is of importance for an understanding of catalysis in general that not only the structure of the free catalyst but also the structures of all complexes that form and decay along the reaction are determined on the atomic level. For this purpose, the catalyzed reaction must be recorded with methods such as X-ray structure determination, which are capable of imaging with atomic resolution. Proteins are ideal to study catalysis. Only proteins form crystals with large connected cavities filled by water. The macromolecules are still flexible in the crystals and, hence, are catalytically active. Substrate molecules diffuse freely through the channels in the crystals and provide supply for the catalytic reaction. In almost all cases protein crystals stay stable and do not decompose during the catalytic reaction. The reaction can, therefore, be followed repeatedly from the beginning to the very end. There is no other solid body in the organic or inorganic world from which we can learn so much about catalysis as from proteins.

9.1.1 Structure and Function of Proteins

The living organisms synthesize, from small organic molecules, complex macromolecules with numerous chemical and physical properties. On the one side there are the nucleic acids (DNA and RNA), which store and transmit

the information of life. On the other side are the proteins built from amino acids, which perpetuate the functions necessary for life. In contrast to organic polymers the biopolymers exhibit, on the atomic length scale, a well ordered structure. For polymers, this is a unique property. The biopolymer structure determination contributed substantially to the explosive growth of knowledge in biology and biochemistry. In this context, X-ray crystallography played an outstanding role.

X-Ray Structure Determination

X-ray structure determination was invented in the first decade of the twentieth century in Munich [1] by researchers around Max von Laue who is commonly regarded as the founder of X-ray structure determination. In his famous book "Röntgenstrahlinterferenzen" von Laue commented skeptically: "Es gibt Eiweißstoffe vom Molekulargewicht etwa 35 000, die gut kristallisieren. Die Elektronenverteilung in ihnen zu bestimmen, ist von vornherein aussichtslos." This means that he considered protein structure determination as a hopeless venture in the first place. Already a short period of time later John Kendrew and coworkers solved the first protein structure [2] that of sperm whale myoglobin, a protein with a molecular mass of 17.8 kDa. A few years later, Max Perutz presented the even fourfold larger structure of hemoglobin [3]. Today, roughly 30 000 structures of biomolecules are stored in the protein database [4], 90% of which are determined by X-ray structure determination and 10% by nuclear magnetic resonance (NMR). Seventy-five of the structures are of proteins. The remainder consists of the structures of nucleic acids, macromolecular carbon hydrates, and protein–nucleic acid complexes.

To become familiar with the concepts of protein crystallography the reader is referred to existing textbooks (e.g. ref. [5]). In short, because of the periodicity of the crystalline lattice the scattering vectors, which describe the scattering of radiation, become discrete with respect to length and direction. These vectors form a lattice, the so-called reciprocal lattice, standing perpendicular on the real lattice of the crystal. If a reciprocal lattice point has the correct orientation towards the incident X-ray beam, or in other words, if the wavelength of the X-ray radiation matches, a reflection can be observed. By Bragg's law the scattering angle is connected to the distance between the crystalline lattice planes. The larger the scattering angle at a given wavelength, the smaller this distance and the higher the resolution. The intensity scattered into a reflection is proportional to the square of the structure factor amplitude of the molecules that occupy the unit cells of the crystal. To determine the structure, the intensities of as many reflections as possible have to be collected. For proteins the number of observed reflections in a data set can easily exceed 10^6. The structure factor is a complex number with amplitude and phase. However, only intensities are measured and the phase is lost this way. This is referred to as the phase problem of the X-ray structure analysis. A number of methods exist to retrieve the phase [5]. From the structure

factor amplitudes and the phases an electron density map can be calculated by Fourier synthesis. This map is the experimental result. Atomic models can be used to interpret these maps if the resolution is better than about the diameter of the atoms, which is in the order of 3 Å.

Biocatalysts

Enzymes are proteins with specific catalytic activity. The activity is determined by the 3D structure of the enzyme. Small ligands, which may bind to various sites of the enzyme and which may have a drastic influence on its 3D-structure, decisively determine the catalytic properties of the enzyme, such as the catalytic rate and substrate affinity. With X-ray crystallography it is possible to determine the atomic structure of the enzyme–ligand complexes and elucidate the molecular mechanism of the impact of various ligands on the enzyme activity.

In enzymes the enzyme–ligand complexes exhibit different structures and evolve into each other during the catalytic cycle. These complexes are also refered to as reaction intermediates. A particularly descriptive example is the catalytic cycle of cytochrome $p450_{cam}$. p450 enzymes catalyze the insertion of oxygen into atomic ring systems (here the camphor) and prepare these substances in this way for metabolism and degradation. Therefore, understanding this class of enzymes is important to grasp the mechanism of poison remediation and drug metabolism. As many of five (of probably seven) reaction intermediates on the catalytic pathway of $p450_{cam}$ were structurally characterized [6]. Over several steps, the generation of a mono-oxygen radical is described in atomic detail. This radical attacks the camphor, which is eventually oxidized.

At this point, the question arises how one can determine the structures of transiently occupied intermediates like those within the $p450_{cam}$ enzymatic cycle. Such a cycle takes place within a few millisecond. It has been always a dream to accelerate the structure analysis so much that fast and ultrafast phenomena could be investigated. However, traditional X-ray structure determination is a tedious process and even nowadays synchrotron collection of a conventional monochromatic data set is in the order of minutes. One obvious loophole is to slow down the reaction so much that there is enough time to collect a data set. The methods to perform this are called trapping methods. Although this chapter mainly deals with the converse, which is very fast X-ray data collection, the trapping methods are valuable tools due to their relative simplicity and deserve to be presented, however shortly, here.

9.1.2 Structure Determination of Intermediate States by Stabilization (Trapping) of their Occupation

At room temperature protein intermediates are populated on time scales, which may span as much as 15 orders of magnitude. Some can be observed

by using fast time-resolved spectroscopy already on the picosecond and faster time-scales [7,8], but also life-times in the region of minutes [9,10] are found. These shortly occupied states are therefore also denoted as metastable or transient intermediates. Note: this denotation is meant in the sense that the concentration of molecules in a particular intermediate state but not the intermediate itself is short-lived. The intermediate states can be considered as energy minima in the multidimensional conformational space of the protein and are, therefore, time-independent. The collection of suitable X-ray data may last from minutes at intense synchrotrons to hours in the home laboratory. This is, in nearly all cases, much longer than an entire catalytic cycle. Consequently, also the metastable intermediates cannot be observed directly. However, if the occupation of the intermediates can be hold constant or trapped over a longer period of times, a structure analysis becomes possible also with conventional methods.

Different trapping methods are listed in the literature [11–14] and are resumed in Table 9.1. Besides performing the trapping experiments with catalytically nearly inactive mutants [20], the preferred way to slow down a reaction is to lower the temperature. Typically, a reaction is initiated and the intermediates are stabilized at the temperature of liquid nitrogen (−196°C) or lower using liquid helium. These approaches are usually called freeze-trap experiments. One may also consider starting a reaction at elevated temperatures and plunging the crystal into liquid nitrogen [15,23] after the reaction has evolved for a period of time. This is referred to as trap-freeze in the literature. Then, in contrast to the other trapping methods, the time-scale, on which the reaction proceeds, stays intact. The time-resolution depends on how fast the crystal can be entirely frozen [24]. Because of this, the trap-freeze methods can be used only for processes that take place on the millisecond time-scale. If freeze-trap is used the structural changes are usually restricted due to the low temperatures [25]. This disadvantage can be overcome by slowly increasing the temperature to values above or around 180 K. At elevated temperatures structural relaxation can take place. The structure of the newly occupied state is determined after the crystal is again cooled down. This protocol can be repeated iteratively to subsequently populate more states as successfully demonstrated for CO-Myoglobin [26] and for $p450_{cam}$ [6]. With the trapping methods the structures of a series of reaction intermediates could be observed in the catalytic cycle of numerous proteins and enzymes (see also Table 9.1 for an assortment). Experiments at cryogenic temperatures have the advantage that the crystals are very stable during the X-ray measurements [27].

By using monochromatic X-ray radiation the scattering background can be substantially reduced, and the reflection intensities are collected to highest resolution. The experimentally determined electron density maps are of exquisite quality. Anyhow, the interpretation of the electron density can be difficult as reported, for example, for the proton pump bacteriorhodopsin [28]. Once a reaction is started the molecules accumulate in multiple states whose electron densities mix into each other. Freezing stabilizes the mixture, which

Table 9.1. Advantages and disadvantages of the different trapping methods

	Trap-freeze	Freeze-trap without temperature cycles	Freeze-trap with temperature cycles	Chemical stabilization mutations	Analytical trapping
Examples	β-Lactamase [15] Farnesyltransferase [16]	Myoglobin [17] PYP [18]	^{L29W}Myoglobin [19] $p450_{cam}$ [6]	Isocitrate Deydrogenase [20]	PYP [21, 22]
Advantages					
	– Highest resolution – Time scale preserved – Separation of intermediates possible – Kinetic mechanism can be determined	– Highest resolution	– Highest resolution – Early and late intermediates – Separation of intermediates possible	– Highest resolution	– Authentic intermediates – No cryo-artifacts – No artifacts due to mutation – Time scale preserved – Early and late intermedates – Separation of intermediates possible – Determination of kinetic mechanism
Disadvantages					
	– Only slow processes – Cryo-artifacts possible	– Time scale is lost – Mixture of states – Only early intermediates – Cryo-artifacts possible	– Time-scale is lost – Cryo-artifacts possible	– Time scale is lost – Mixture of states – Structural changes due to mutation	– Resolution and signal-to-noise ratio lower – Complex techniques

Trap-freeze, freeze a period of time after reaction initiation; Freeze-trap, freeze before reaction initiation; Chemical trap, slowing of a reaction to rest due to a site-specific amino-acid exchange; Analytical trapping, Time-resolved crystallography in combination with a kinetic analysis. Examples from the literature are given.

complicates the structure determination. Moreover, it is not clear whether the states trapped at low temperatures match those populated at ambient temperatures.

If the crystallographic experiments are performed at ambient temperatures, the reaction can evolve in an undisturbed way and authentic intermediates may be observed. Then, rapid data collection is an essential requirement and the experiments become time-resolved. In this case, new methods to extract the intermediates from the time-resolved data had to be developed. Since the molecular states relax into each other, admixtures of several intermediates will be observable at any time during the reaction. These admixtures must be separated into the pure contributions.

If the time is to be used as an additional variable, the amount of data, which have to be analyzed simultaneously, grows linearly with the number of time points. Although the well established but slow methods in crystallography had to be given up in favor of rapid data collection, the benefit is substantial: protein crystallography can finally be linked to chemical kinetics. Roughly 50 years after the first protein structure was solved, this is now possible.

9.2 Crystallography Meets Chemical Kinetics

Although crystallography is the main subject of this chapter, chemical kinetics is introduced first. Chemical kinetics leads to the understanding how advantageous it is to use time-resolved crystallographic methods.

9.2.1 Chemical Kinetics

During the course of a reaction catalyzed by an enzyme, the protein molecules relax through their intermediate states [29]. In Fig. 9.1 a reaction is outlined, in which four intermediate states are connected by uni-molecular steps. Different reaction pathways as well as reversible and irreversible steps are possible. The velocity or, in other words, the rate of each uni-molecular step is determined by both the rate coefficients k and the concentrations of the particular intermediate in front of the arrow, which determines the direction of the reaction. The magnitude of a particular rate coefficient is determined among other factors by the height of the activation energy barrier between the corresponding two states.

The time-dependent concentrations of molecules in the intermediate states can be calculated by integrating the coupled differential equations that describe the mechanism. Equation 9.1 shows the equations of mechanism III in Fig. 9.1 as an example. The concentration in state D is determined by the conservation of mass. The integration can be performed by using the approach that the increase and decay of the concentrations of the intermediates can be

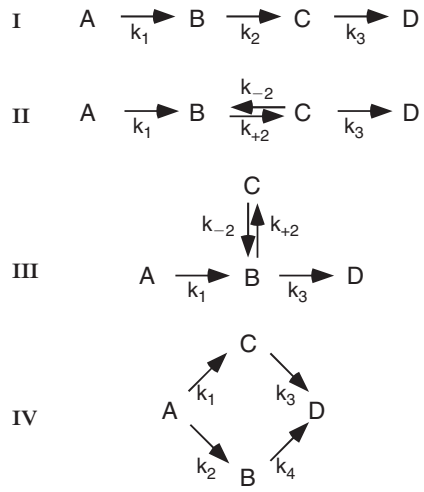

Fig. 9.1. Four different kinetic mechanisms for a reaction with four states, A-D. $k_1 \ldots k_4$ are the rate coefficients that describe the mechanisms

described by exponential functions [30]. For more complicated cases, which are not subject of this chapter, other approaches must be found [31–34].

$$\begin{aligned}
\frac{\mathrm{d}\,[A]}{\mathrm{d}t} &= -k_1\,[A] \\
\frac{\mathrm{d}\,[B]}{\mathrm{d}t} &= k_1\,[A] - k_{+2}\,[B] - k_4[B] + k_{-2}\,[C] \\
\frac{\mathrm{d}\,[C]}{\mathrm{d}t} &= k_{+2}\,[\mathrm{B}] - k_{-2}\,[C] \\
[A] + [B] + [C] + [D] &= \text{const.}
\end{aligned} \tag{9.1}$$

The exponential approach is valid if the thermal equilibration about all degrees of freedom in a protein is much faster than the time needed to depopulate an intermediate. At room temperature and on a time-scale longer than a few nanosecond this is usually the case. Integration proceeds by determining the eigenvalues λ_i and eigenvectors of the so-called coefficient matrix of the reaction (see ref. [35] for illustrative examples). The eigenvalues are the relaxation rates, the (absolute) inverse of which are called the relaxation times. The eigenvectors are also denoted eigenconcentrations [36] and determine (up to a constant of integration) the preexponentials $[P_{ij}]$ of the exponential functions. Equation 9.2 shows the general solution: Anyone of the three exponential terms is called a transient of the intermediate.

$$\begin{aligned}
[A] &= [P_{11}]\ \mathrm{e}^{\lambda_1 t} + [P_{12}]\ \mathrm{e}^{\lambda_2 t} + [P_{13}]\ \mathrm{e}^{\lambda_3 t} \\
[B] &= [P_{21}]\ \mathrm{e}^{\lambda_1 t} + [P_{22}]\ \mathrm{e}^{\lambda_2 t} + [P_{23}]\ \mathrm{e}^{\lambda_3 t} \\
[C] &= [P_{31}]\ \mathrm{e}^{\lambda_1 t} + [P_{32}]\ \mathrm{e}^{\lambda_2 t} + [P_{33}]\ \mathrm{e}^{\lambda_3 t}.
\end{aligned} \tag{9.2}$$

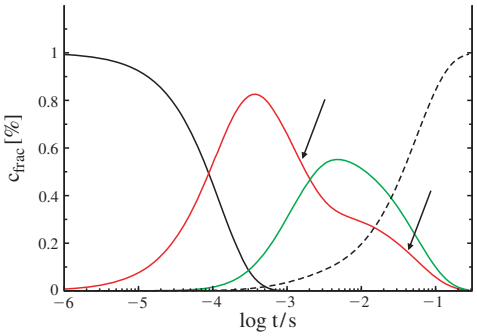

Fig. 9.2. Time-dependent normalized concentrations of the four states in mechanism III shown in Fig. 9.1. *Black solid line*, intermediate A; *red line*, intermediate B; *green line*, intermediate C; *back dashed line*, final state D. Data are on a logarithmic time scale. Arrows show the two observable transients when intermediate B depopulates. Rates: $k_{+1}, 800\,\mathrm{s}^{-1}$; $k_{+2}, 500\,\mathrm{s}^{-1}$; $k_{-2}, 300\,\mathrm{s}^{-1}$; $k_{+3}, 50\,\mathrm{s}^{-1}$

In Fig. 9.2 the coupled differential equations are integrated and the time dependent concentrations are shown for mechanisms III in Fig. 9.1. Not only one but several intermediate states are occupied at almost each time point. This is the reason that admixtures are observed. However, if the time-dependent concentrations are known or can be determined accurately enough from the data, one has, in general, a chance to separate the admixtures. This can be done only if the time-scale is preserved in the experiment.

A convenient way to determine the mechanism is to perform time-resolved spectroscopic experiments. However, the assignment of spectroscopically observable intermediates to atomic structures may be difficult [37, 38]. In many cases it is unknown how far structural changes are connected to changes observed, for example, in the optical absorption spectra. Vice-versa, it is also possible that a structural change remains silent in the wavelength range accessible to the particular experimenter. It is, therefore, desirable to determine the kinetic mechanism based on structural evidence using time-resolved X-ray data and use the results from spectroscopy as complementary information.

9.2.2 Time-Resolved X-Ray Structure Analysis

Fast time-resolved X-ray structure analysis has been developed starting in the mid eighties of the previous century (see ref. [39] for a short historical summary) with the first Laue images on protein crystals. Moffat [40] was well aware of the possibilities inherent to this method: both, the chemical, kinetic mechanism and the structures of the intermediates can be determined.

Time-Resolved Pump-Probe Experiments

To detect the shortly occupied intermediates, the X-ray data must be collected as fast as possible. Intense and very short, in the order of 100 ps, polychromatic X-ray pulses are generated at third generation synchrotrons such as the ESRF in Grenoble/Fr or the APS in Argonne/USA. In these synchrotrons electrons travel in bunches close to light velocity and typically circle around in about 3.5 µs. In the standard operation mode of the synchrotron several hundreds of these bunches may be equally spaced in the storage ring. In special operation modes one bunch is circling either alone – this mode is referred to as the single bunch mode – or the bunch is traveling opposite to and well separated from a super bunch – this mode is sometimes referred to as the hybrid mode (Fig. 9.3).

In pump-probe experiments a reaction is first started, for example, by an intense laser flash and the structure is probed by an X-ray pulse at a time Δt after the laser flash (Fig. 9.3). The minimum Δt that can be resolved is the time-resolution. It depends on the duration of the X-ray flash and that of the laser, whichever is shorter. X-ray flashes are generated using so-called insertion devices such as wigglers or undulators [41]. The radiation can be optimized to the needs of the crystallographic experiment by the design of the insertion device [42–44]. X-ray flashes from a single bunch at 3rd generation synchrotrons last about 100 ps and contain approximately 10^{10} photons at the position of the protein crystal [45, 46]. This is sufficient to collect analyzable scattering patterns. Single pulse experiments become feasible [47, 48] and the maximum available time-resolution can be exploited. Figures 9.3 and 9.4 show

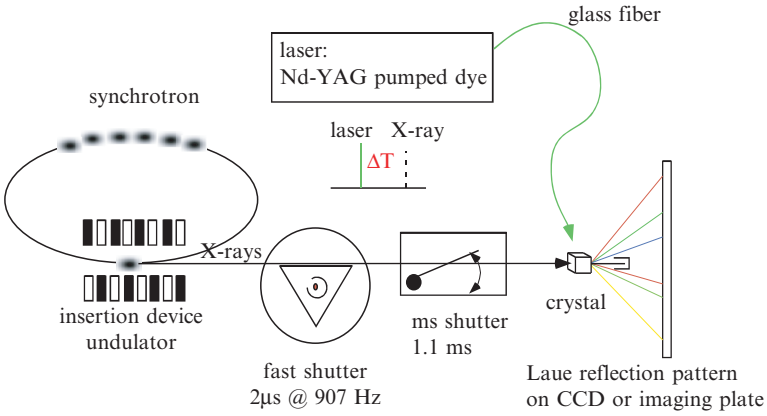

Fig. 9.3. Experimental setup for a time-resolved crystallographic experiment (schematically). Here the "hybride mode" at the Advanced Photon Source is sketched. A single X-ray flash is extracted by a shutter train consisting of a slow and a fast shutter. The crystal is irradiated by this short X-ray flash a time-interval Δt after the reaction has been initiated by an intense laser flash

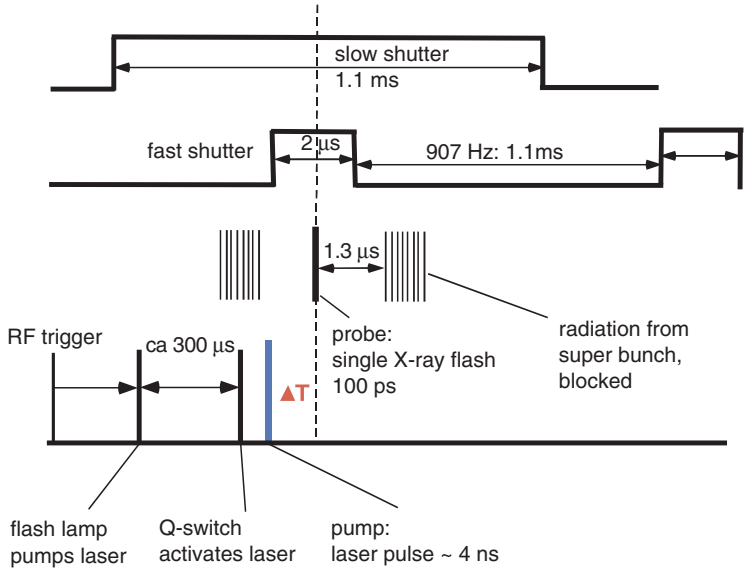

Fig. 9.4. Synchronization of the single X-ray pulse, the fast and slow shutters, and the laser set up for a nanosecond time resolved experiment. Several time delays after the radio frequency (RF) trigger are needed to synchronize the different components. For automation, these time-delays are provided by programmable delay generators

how the various components of the experiment are synchronized. A shutter train consisting of a fast shutter that opens every 1.1 ms for roughly 2 µs and a slow shutter with an opening time of 1.1 ms extracts the 100 ps X-ray flash. By reversing the phase of the fast shutter, the stronger radiation from the super-bunch (c.a. 800 ns) can also be used on the cost of time-resolution.

Slower processes, however, can also be recorded in the normal operation mode. In this case a range of possible time resolutions is available. By slowing down the fast shutter time-resolutions from 2 µs to about 10 µs are usually possible. If the processes are even slower, the fast shutter can be left open and the time resolution can be selected by the opening time of the millisecond shutter.

The integral intensity of the reflections must be collected during the time of the X-ray flash. However, the mosaicity of the protein crystals determine the reflection range with typical angles in the order of 0.1° [49]. An attempt to collect the entire integral intensity of a reflection within a single flash of monochromatic radiation would require that the whole mosaicity is rotated through Ewald's sphere within 100 ps. Hence, the crystal would have to be rotated by about 10^6 rounds per second, which is not possible. Consequently, still exposures must be used.

Still exposures are used with great success in the field of crystallography with quasi monochromatic neutrons pioneered by Niimura and coworkers

[50,51]. In these experiments the $\Delta\lambda/\lambda$ of the incident neutrons is in the order of only a few percent and excellent data to highest resolution can be collected and analyzed as if they were monochromatic. An analogous approach would be conceivable also in time-resolved crystallography. Nevertheless, the volume of the reciprocal lattice sampled within one exposure is quite small. Since the reaction in the crystal has to be reinitiated for each exposure, it is advantageous to collect as many reflections as possible simultaneously. This can be done using polychromatic radiation. Then, the fairly simple and well established monochromatic methods must be replaced by the Laue method [52].

The Laue Method

With the Laue method not only a large fraction of reciprocal space is sampled at once (Fig. 9.5) but also the entire integral intensity of all the reflections in that volume is collected instantaneously without rotating the crystal. Therefore, the Laue method has been established as the method of choice in time-resolved crystallography. By using polychromatic radiation, the determination of accurate structure amplitudes is not quite trivial due to four principal problems:

1. Since a substantial number of reflections is excited at once there is a large probability that the reflections overlap substantially on the detector [53–55]. This spatial overlap can be resolved by using analytical or numerical reflection profiles, which are determined at numerous positions on the detector with the help of nonoverlapping reflections [56, 57].

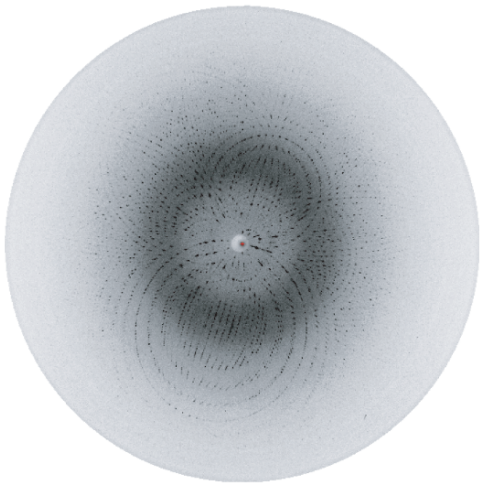

Fig. 9.5. Time-resolved Laue photograph on a crystal of a myoglobin mutant, 35 ns after reaction-initiation. The photograph has been recorded on a Mar345 imaging plate at Lauebeamline 14ID-B, BioCARS, APS, Argonne/USA by using radiation from an undulator

2. If reflections whose indices share a multiple of a basic index triple are exited simultaneously, they scatter in exactly the same direction [58]. This is the case, for example, for the 6 8 10 and the 9 12 15 reflections, which share the basic index 3 4 5. This exact overlap is called the harmonic overlap and earlier, reflections affected by this overlap were dismissed [55]. However, the number of harmonically overlapping reflections (harmonics) is small compared to the number of singlets, which comprise about 80% of the Laue reflections [54, 58]. Since the harmonic overlap especially applies to reflections at low resolution, the completeness at low resolution is particularly low if harmonics are disregarded. This effect is also referred to as the "low resolution hole." However, the harmonic overlap can be resolved and the low resolution hole filled if a particular harmonic or its symmetry mates are measured multiple times at different crystal settings [59]. If such multiple measurements are not available, data statistics such as the resolution dependent distribution of reflection intensities and other a-priori knowledge can be employed to estimate the contributing intensities [60, 61].
3. The intensity of the primary X-ray beam is varying within the bandwidth as a function of the X-ray energy. Therefore, dependent on the orientation of the crystal, the reflections are exited by different frequencies (energies) and intensities of the primary beam. The reflection intensities, therefore, must be brought to a common scale. This is called wavelength normalization accounted by the so-called λ-curve. The λ-curve can be determined directly from the data set of the unscaled reflection intensities and corrects in addition for all wavelength dependent parameters such as the detector sensitivity and the scattering power of the crystal [40]. Highly flexible Chebychev polynomials are successfully used [47, 56, 59, 62] to align the λ-curve. However, if the spectrum of the incident X-rays varies sharply, the λ-curve is difficult to determine. This is the case especially at the critical wavelength of an undulator. Therefore, normalization of data from undulator radiation has been reported only recently [43, 44].
4. A large fraction of the polychromatic X-rays is not reflected at all and generates a scattering background due to, amongst other things, crystal disorder, capillary scattering, scattering by air, and Compton scattering. This background prevents the determination of very weak reflection intensities, which are common at high resolution. The background is substantially reduced using radiation with a smaller bandwidth, which is typically generated by specially designed "narrow bandpath" undulators [43, 44]. With radiation from such a device the data can be collected to higher resolution, and the accuracy of the reflection intensities is greatly enhanced due to the high peak intensity of the radiation. Therefore, the "narrow bandpath" undulator is the preferred device. It balances the advantages of a radiation, which is as much as possible monochromatic, with the needs for a bandwidth, which allows the instantaneous collection of the integral reflection intensities of a larger part of reciprocal space.

9.3 From the Reaction Initiation to Difference Electron Density Maps

9.3.1 Reaction Initiation

To determine the structures of the intermediates, the intermediate states must be transiently occupied at a sufficiently high level. The X-ray structure analysis averages over a large number (typically 10^{14}) of molecules in the crystal. Therefore, it is indispensable, as many molecules as possible are in the same state at the beginning of the reaction. This state must be prepared by the reaction initiation [63]. Protein crystals suitable for experiments with the time-resolved X-ray structure analysis have edge lengths of approx. 150 µm. For slow reactions it is sufficient to diffuse the substrate into the crystal [64]. The crystals can be saturated with substrate within seconds to minutes [13]. Indeed, the catalytic rate must be much slower to ensure a significant population of intermediates. The time-resolved X-ray structure analysis is particularly qualified to analyze fast processes. Consequently, the reaction initiation must be also performed very fast. Processes activated by light are well suited for this purpose, since they can be triggered by short and ultrashort Laser pulses (Figs. 9.3 and 9.4).

For enzymes, "caged" substrates that are inactive in the dark can be used. After the crystal is saturated by theses substrates, the reaction is initiated by an intense light flash and the substrate is enzymatically processed to the product. The rate of activation of the "caged" substrates is usually in the order of microsecond and intermediates that become populated on similar time scales can be observed [65]. The fastest processes, however, can be observed if the protein itself contains a photoactive functional group. After irradiation with light, the earliest intermediates may be populated on very fast time scales (ps and faster). Then, the observability of these intermediates depends on the time-resolution of the method itself, which is determined here either by the duration of the X-ray pulse or by the duration of the laser flash that initiates the reaction, whichever is slower.

If the protein molecule contains a chromophore, the crystals are unusually optically dense at or close to the absorption maximum of the chromophore. Small crystals with edge lengths around 100 µm should be used and the wavelength of the exciting light should be selected substantially off the peak absorption to ensure homogeneous illumination throughout the volume of the crystal probed by the X-ray beam. Pulsed lasers are convenient to initiate reactions in protein crystals with cross-sections of about 0.01 mm^2. Shortest pulse durations and the precise adjustment of the wavelength are possible [66–69]. The laser beam can be split to illuminate both the bottom and the top of the crystal to maximize the volume which is homogenously illuminated.

In crystals of proteins with a moderate molecular mass of about 20 kDa, the concentration of a chromophore is very often around 50 mmol l^{-1}. In a crystal volume of 1 nl 50 pmol (3×10^{13}) of molecules are present. In ideal

cases, each chromophore absorbs one photon. 3×10^{13} blue photons have a total energy of about 15 μJ. Common laser setups deliver in the order of 5 mJ per pulse into a cross-section of about 1 mm^2 after the light pulse is coupled into and transported through a glass fiber. Even if 99% of the photons miss the crystal enough photons are available.

The time point when the laser flash impinges the crystal can be controlled exactly (Fig. 9.4). The experimentally determined jitter of light pulses from a Q-switched Nd:YAG laser is in the sub-nanosecond regime (V. Srajer, BioCARS, personal communication). However, the pulse duration of such a laser setup is typically around 4 ns, which is more than one order of magnitude larger than that of the X-ray flash. Recently, experiments with light pulses from a femtosecond laser were performed on myoglobin mutants at the European Synchrotron Radiation Facility [45, 46]. To avoid damage by the intense laser intensity and to facilitate the penetration of the laser light in the crystal, the femtosecond pulses were stretched to a few picosecond. With these pulses the maximum possible time-resolution of 100 ps, the duration of the X-ray flash, became possible and early processes were observed.

Usually, the number of photons in one incident X-ray pulse is not large enough to produce a sufficiently strong diffraction pattern. For each crystal setting multiple X-ray exposures (of the order of 10–100) are necessary even with the strongest X-ray sources [45]. Accordingly, the reaction must be reinitiated each time. To collect an entire data set multiple crystal settings with an angular spacing in the order of 3° (depending on the bandwidth of the radiation from the insertion device [39]) are necessary.

9.3.2 Detectors

In protein crystallography, area detectors are used to collect the reflection patterns. These should posses an aperture as large as possible to collect reflections at highest scattering angles, they should be highly sensitive to determine the intensity of even the weakest reflections accurately and they should have a very good pixel resolution to separate even closely spaced reflections. About 10^7 X-ray photons are scattered on the detector during the 100 ps duration of the X-ray pulse. This means that the detector must be able to handle 10^{17} events per second. Common gas detectors or multiwire detectors with typical counting rates of 10^5–10^6 Hz cannot be considered. The X-ray photographic film that has been used frequently to collect Laue-data is not sensitive enough. However, area detectors like the image plate and the CCD detector fulfill the above criteria and are therefore used throughout. Since these detectors must be read out, which lasts some seconds, all experiments to date are therefore of the "pump-probe" type. For each X-ray exposure the reaction in the crystal must be restarted. New detectors will consist of a mosaic of individually addressable pixels that can be read out extremely fast [70]. If these pixel detectors can be "time gated," a stroboscopic mode of data collection will

become possible, where an entire time-course of diffraction patterns can be recorded after the reaction is initiated only once.

9.3.3 Data Reduction

The results of a successful scattering experiment are digitized images of reflection patterns (see Fig. 9.5). From these accurate structure factor amplitudes must be obtained, which is performed by the data reduction software. This software has to cope with the special requirements associated with the Laue method. Only a limited number of software packages should be mentioned here (see ref. [39] for a larger assortment): (1) The Daresbury Laue Software Suite Lauegen [71] is used for the initial indexing and geometric refinement. The separation of spatial overlaps and integration is done by Prow [57]. Finally the λ-curve is determined and the harmonics are deconvoluted by Lscale [62]. (2) The Chicago LaueView suite [56, 59] integrates all necessary steps into one program. Recently, a new, commercial program package, Precognition and Epinorm, has been released (www.renzresearch.com) in which the experiences gathered primarily from data reduction with LaueView have been condensed. With automatic pattern recognition of nodals and ellipses observable in the Laue reflection pattern, the data reduction becomes semi-automatic, robust, and particularly easy to use also for novices.

9.3.4 Difference Maps

To follow the reaction in the crystal, data sets of Laue structure amplitudes $|F_t|$ at different time-points after the reaction initiation are collected. The time points are ideally arranged equidistant on a logarithmic time-scale. Because of this, fast and slow relaxations are equally considered. In addition, the structure amplitudes of the dark state $|F_D|$ are collected as a reference. By subtraction of the structure amplitudes of the dark state from the time-dependent Laue amplitudes time-dependent difference structure amplitudes $\Delta F_t = |F_t| - |F_D|$ are obtained. Using the phases ϕ_{D} from the ground state a set of time-dependent difference electron density maps $\Delta\rho_t$ can be calculated (9.3):

$$\Delta\rho_t = \frac{1}{V_{\mathrm{e}}} \sum_{hkl} w \Delta F_t \, \mathrm{e}^{\mathrm{i}\phi_{\mathrm{D}}} \mathrm{e}^{-2\pi\mathrm{i}(hX+kY+lZ)}, \tag{9.3}$$

where hkl are the reflection indices, X, Y, Z are fractional coordinates, V_{e} is the volume of the unit cell, and w is a weighting factor. Difference maps should be preferentially weighted [72–75] to reduce the influence of outliers and inaccurately determined reflection intensities. The maps are typically contoured on a positive or negative multiple of their root mean square deviation (σ-value, see Fig. 9.6). Negative difference density refers to locations from which atoms migrated to other positions and positive values refer to newly occupied positions. The time series of difference electron density maps contain both the

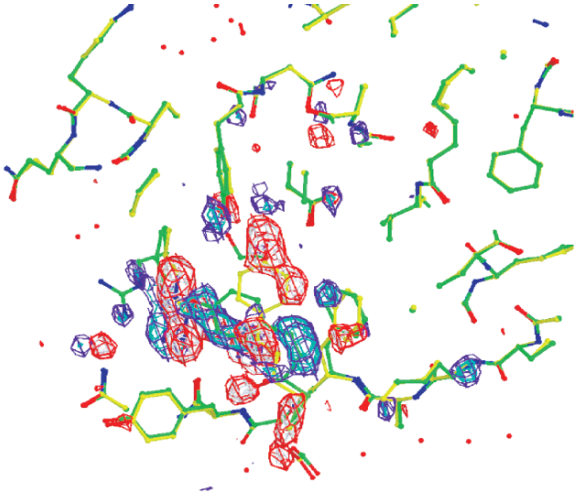

Fig. 9.6. Time-dependent, weighted difference electron density map, 2 ms after reaction initiation Contour levels, negative, red/gray $-3\sigma/-4\sigma$; positive, blue/cyan $3\sigma/4\sigma$. Atomic structure in yellow, ground state model; atomic structure in green, guide to the eye, approximate explanation of the positive difference electron density features

structure information, which eventually leads to the determination of structures of the intermediates, as well as the time information, which leads to the determination of a chemical, kinetic mechanism.

9.4 Experiments

A substantial number of successful time-resolved experiments using the Laue method have been performed so far on different biological systems [39]. However, in the context of this chapter, only myoglobin and the photoactive yellow protein are considered, since the importance of both molecules was the driving force for the development of the new, fast methods. Results from crystallography are subsumed, however, shortly here.

9.4.1 Myoglobin

Myoglobin is a small heme-protein with a molecular weight of 17.8 kDa and 153 amino acids. Carbon monoxide binds as a sixth ligand to the iron(II) of the heme. This coordinative bond can be ruptured by a photon absorbed by the heme. Figure 9.7 subsumes most of the findings from photoflash experiments at low temperatures and at room temperature in the wild-type Mb and in the mutants L29F and L29W [17, 19, 25, 26, 45, 46, 73, 76–81].

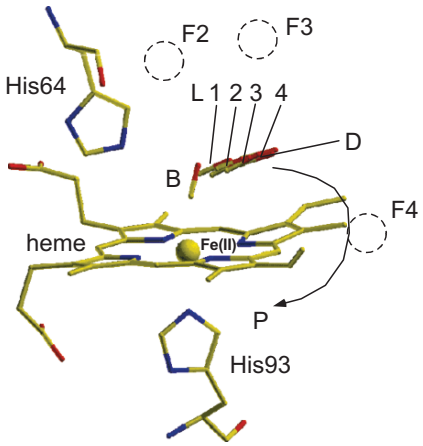

Fig. 9.7. Close view of the active site in Myoglobin. Site B, bound CO; sites L1–L4, flash experiments at low temperatures; site D, distal docking site at room temperature, equivalent to site L3; site P, proximal docking sites; F2, F3, and F4, additional sites observed in the 150 ps study on the L29F mutant

The first successful, fast time-resolved experiments with single bunch X-ray exposures were performed by Srajer et al. [78] on wild-type CO-myoglobin at room temperature. A series of six time-dependent difference maps between 4 ns and 1.9 ms showed the position of the CO after the flash. The CO is in site D, also called the distal docking site. Site D is equivalent to site L3 found at low temperatures in combination with long illumination times. After photolysis, subsequent rebinding of the CO was followed. However, a meaningful kinetic analysis became possible as soon as difference electron density maps were available at significantly more time points [73].

In addition to the distal docking site another binding place for the CO was identified on the proximal side of the heme (Fig. 9.7, position P), which could also be observed in the wild-type and in both the L29W and the L29F mutants [26, 45, 46, 79–81].

For a quantitative analysis, the total electron count in the various sites was identified by integrating the difference electron density. In the wild type, about 40% of the CO molecules have been flashed away. The distal docking site is fully occupied already 1 ns after the flash and is depopulated on a time scale of 100 ns. Simultaneously, site P is occupied. The CO escapes in less than 1 µs to other places and most likely also to the water space. However, it seems that an equilibrium is established between the molecules in site P with those outside the protein. At around 100 µs after the flash the CO rebinds to the iron. It should be stressed here that the exact trajectory of CO molecules between the binding sites cannot be determined due to principal reasons. A sufficient occupation of states is only possible at those sites, which can be thought of local but pronounced energy minima on the reaction coordinate.

An approximate CO migration pathway can be identified by comparing time-resolved X-ray data on the Wild-Type Myoglobin with those on the L29W mutant [81]. The bulky tryptophane that replaces the leucine at position 29 on the distal side of the heme pocket serves as a dynamic plug, which keeps the CO particularly long in the proximal docking site P. It is interesting, however, that the CO leaves site P about one order of magnitude faster (1.5 ms) than it rebinds to the iron (25 ms). Obviously, the Trp29 keeps the CO from migrating to the distal side and it has to escape to the solvent on the proximal side. In the Wild-Type escape from site P as well as rebinding happens roughly simultaneously on a time-scale of some 100 μs, much faster than in the L29W mutant. The reason for this is that in the Wild-Type the migration pathway to the distal side is not blocked, CO can easily migrate to the distal side and escape to the solvent from there [82]. This shows that the proximal pathway is insignificant for CO escape. Since the L29W mutant rebinds CO so slowly, a particularly long-time window is opened to follow the initial relaxations in the protein without interference from CO rebinding. The initial relaxation phase appears to be nonexponential, which is evidence for a more complicated reaction dynamics at fast time-scales.

9.4.2 The Photoactive Yellow Protein

The Photoactive Yellow Protein (PYP) is a 14 kDa protein with bright yellow color, which was initially isolated from the bacterium *Ectothiorhodsopira halophila* [83]. PYP is proposed to be the primary acceptor for a light induced signaling cascade in the cell. *E. halophila* is negatively phototactic [84]; the absorption spectrum of PYP is similar to the action spectrum of the bacterium. The mechanism how the signal is transduced from the PYP-photoreceptor into the bacterium, however, is not fully understood.

In Fig. 9.8 the X-ray structure of PYP is shown [85]. The central chromphore, para-coumaric acid (pCA, also called 4-hydroxicinnamic acid) is embedded in the typical PAS domain fold [86], a five-stranded antiparallel β–sheet flanked by 3 helical segments, which accounts for roughly 80% of the PYP.

The pCA chromophore can be excited by a blue photon, upon which the configuration of the pCA changes from *trans* to *cis*. This drives the PYP molecule into a photocycle with four in the UV/vis range identifiable intermediates I_0, $I_0{}^{\#}$, pR, and pB [7,87] (see also Fig. 9.9). It is assumed that the pB state is the signaling state and, therefore, is responsible for the negative phototaxis. Further spectroscopic investigations suggest that both states, pR and pB, are structurally heterogeneous [37,38,89–91].

Up to now the structures of two or probably three of the intermediates associated with the photo-cycle were determined. An early intermediate has been trapped and its structure was determined at cryogenic temperatures [18]. The structure most likely corresponds to that of an early intermediate I_0 or probably $I_0{}^{\#}$.

Fig. 9.8. Crystal structure of the photoactive yellow protein. *red*, helices; *green*, sheets; *yellow*, connecting loops. The position and the structure of the pCA-chromophore is displayed in addition. The pCA chromophore is bound to Cys69

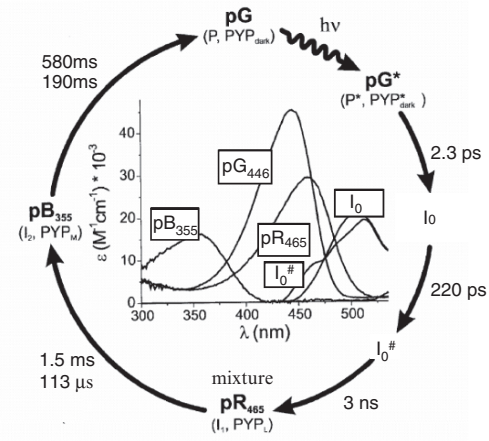

Fig. 9.9. The PYP photocycle and the optical absorption spectra of the PYP intermediates in solution [90]. Ground state pG (absorption maximum at 465 nm), exited state pG*, early occupied intermediates I_0 and $I_0{}^{\#}$ (absorption in the green), later intermediates pR (red shifted absorption maximum), pB (blue shifted aborption maximum). An additional intermediate I_1' between pR and pB is postulated [37] and evidence for an inhomogeneous pR state is found [89], time scales from [7, 37]

The structures of the other two intermediates were determined from only one electron density map each using the time-resolved X-ray structure analysis. The structure of one of the two, probably also $I_0{}^{\#}$, was determined from a pump-probe experiment at 1 ns [74, 92]. That of the other, probably pB [93] was determined from the photostationary state of PYP, which can be generated by exposing the PYP crystal to a 200 ms pulse of blue laser light.

From PYP well scattering, rod-shaped hexagonal crystals of space group $P6_3$ can be grown, which are very well suited for the time-resolved X-ray structure analysis. A number of studies that involve a series of time-dependent difference maps on PYP and its mutants are presented so far [21,22,74,94–96]. For the first time a fit of exponential functions to time-dependent difference electron density features found in the time-dependent difference maps became possible [74]. Using these exponential function a movie with frames closely spaced in time could be prepared by interpolation. The reader should not fall to the intuitive idea that this movie shows the trajectory of the atoms moving through the unit cell. In fact, this movie showed the appearance and disappearance of positive and negative difference electron density features at particular locations in space. These features indicate the population and depopulation of intermediate states, whose electron densities and structures must be extracted from the movie. For this purpose, Schmidt et al. [21] applied for the first time a new method, which is based on the singular value decomposition, to a time-series of 15 time-dependent difference maps collected on a time-scale from 5 µs to 100 ms. From this, the structures of the intermediates as well as plausible kinetic mechanisms can be derived. This new method is explained in the following.

9.5 A New Method for the Analysis of Time-Resolved X-ray Data

Difference maps consist of numerous positive and negative density features. It is not sufficient to determine the time-course of the integral density of only a few features. The temporal variation of all features should rather be considered and described simultaneously and in a global way. New methods for the analysis of these maps are required.

9.5.1 The Singular Value Decomposition

As explained in Sect. 9.2.1 the velocity, or rate, of uni-molecular reactions is proportional to the occupation of the intermediate states in the crystal and to the rate coefficients k. If the concentration of molecules in one intermediate state is declining, one or more subsequent intermediate states are occupied, until finally, the ground or dark state is reached. As a result at each time-point multiple intermediate contribute simultaneously to the time-dependent difference electron density maps (see also Fig. 9.2).

These maps are extremely difficult to interpret, unless the mixture can be separated into the pure contributions. For this purpose the time information can be exploited by a component analysis such as the singular value decomposition (SVD). The SVD is widely employed in spectroscopy [97–103], and it is also applied to analyze genome wide expressions of RNA [104] or to trajectories from molecular dynamics simulations [105, 106]. However, for the analysis of time-resolved X-ray data, new tools had to be developed [75]. The underlying principles of these new tools are described here.

The SVD is only successful if the data depend linearly on the components, which are, here, the time-dependent occupations, or fractional concentrations, of the intermediates. Reflection intensities, for example, are not linearly dependent on the concentrations. Therefore, it is not clear what results if the SVD is applied to them. In contrast, electron density is equivalent to concentration. Because of the difference approximation [107] this holds also for difference electron density values.

For an SVD analysis, the time-series of $t = 1 \ldots T$ difference maps must be related one by one and in temporal order to corresponding column vectors of a data matrix, called matrix $\boldsymbol{A}$ here. This can be done with relative ease since the maps are represented in the computer on a grid with $m = 1 \ldots M$ grid points. For a difference map at time point t, the difference electron density found in the mth grid point can be written to the nth element of the tth column vector of matrix $\boldsymbol{A}$. The order, how the mth grid point is matched to the nth element of the vector in $\boldsymbol{A}$ does not play a role, even if it is done in a completely random fashion. Of course, a once constituted order, which has been applied to the first map/vector pair, must stay the same throughout the $t = 1 \ldots T$ maps and vectors.

There is no need to relate an entire difference map to a vector of matrix $\boldsymbol{A}$, only the asymmetric unit needs to be considered. Smaller volumes such as the volume occupied by protein atoms can be used and the remaining water space left out. Moreover, also those grid points that do not contain significant difference electron densities throughout the time-course can be disregarded.

Once a protein mask is used to constrain the volume, typically $M = 100\,000$ grid points have to be considered per map for a protein of 20 kDa molecular mass. This number reduces to less than 20 000 if only those grid points that contain difference electron density larger than 2σ or smaller than -2σ are considered. A matrix, which consists of 20 difference maps with 100 000 grid points each, can be decomposed (9.4) in less than half a minute on a personal computer into a $T \times M$ dimensional Matrix $\boldsymbol{U}$, which contains the left singular vectors (lSV), the $T \times T$ square diagonal matrix $\boldsymbol{S}$, diagonal elements of which are called the singular vaules (SV), and the transpose of the $T \times T$ square matrix $\boldsymbol{V}$. The rows of $\boldsymbol{V}^{\mathrm{T}}$ are called the right singular vectors (rSV).

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{T} \tag{9.4}$$

The lSV are the spatial main components and are difference density maps. The rSV contain their temporal variation, whereas the SV are weighting factors. Later, the decomposition matrix $\boldsymbol{A}$ can be approximated or reconstructed in a least squares sense using only the significant singular vectors and values.

9.5.2 The Noise Filter

The number of significant singular values and vectors is related to the (minimal) number of intermediates in the reaction. Consider a (hypothetical) reaction, where molecules in the ground state (D) are excited to only one intermediate state, IS1. The molecules simply relax back to the ground state: IS1 $\rightarrow$ D. In this hypothetical experiment, $T = 10$ time-dependent difference maps follow the relaxation. Because of the nature of this reaction, there is only one kind of difference map present, the IS1-D (IS1 minus D) difference map, which vanishes with time. Obviously, the mean or average difference map is equivalent to the IS1-D difference map. Consequently, after the SVD there is only one significant lSV and the corresponding rSV describes its temporal variation. The lSV occupies the first column, the rSV the first row, respectively, in the matrices $\boldsymbol{U}$ and $\boldsymbol{V}^T$, since the SVD algorithm orders the singular vectors and values according to there significance. Additional nine lSV and rSV are present, in which the density features vary fast in time. They do not contribute to the true signal and contain only noise. One should realize that all 10 difference maps in data matrix $\boldsymbol{A}$ can be approximated by another matrix $\boldsymbol{A}'$ simply by reversing (9.4) and taking into consideration only the first vector of matrix $\boldsymbol{U}$ and $\boldsymbol{V}$ and the first singular value from matrix $\boldsymbol{S}$.

Now consider a reaction, where the molecules relax through two intermediates. IS1 $\rightarrow$ IS2 $\rightarrow$ D. Then, the first significant lSV_1 will be again the average electron density of the relaxation processes and the rSV_1 describes its temporal variation. However, this time, the lSV_1 will consist of a mixture of the IS1-D and IS2-D difference maps. There will be a second significant lSV_2, which contains the average deviation from the average difference map, and of course, there is also a second rSV_2 describing its temporal variation. However, the remaining eight insignificant vectors at position 3 to 10 in the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ contain only noise. In the case of two intermediates, data matrix $\boldsymbol{A}'$ can be constructed using the first two significant singular vectors and values only.

These examples can be expanded easily to more intermediates and demonstrate that it is important to judge the number of significant singular values and vectors. This can be done conveniently if the lSV, which are difference maps, are displayed on a graphics screen. Figure 9.10 shows one significant and one insignificant lSV from mock (simulated) difference maps at 21 time points. It is of pivotal advantage that the structure of the ground state can be displayed in addition, because it guides the eye to positions where signal should be present. This structural constraint exquisitely facilitates the discrimination of significant from insignificant lSV. It should be mentioned that the sign of

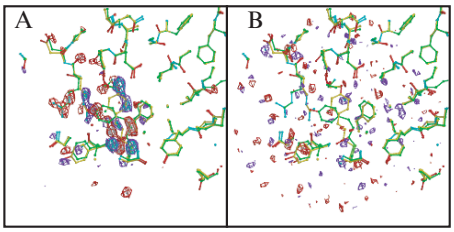

Fig. 9.10. Significant and insignificant left singular vectors from a simulation employing three intermediates plus the dark state. (**a**) First most significant lSV, contour level: red/gray, $-3.5\sigma/-4.5\sigma$; blue/cyan, $3.5\sigma/4.5\sigma$. (**b**) 5th (insignificant) lSV, contour level: red/gray, $-2.0\sigma/-3.0\sigma$; blue/cyan, $-2.0\sigma/-3.0\sigma$. Yellow atomic model: structure of the ground state. Green atomic model: guide to the eye

the difference electron density is irrelevant in the lSV. In the reconstruction of data matrix $\boldsymbol{A}'$, the correct sign is restored by the rSV. Therefore, signal can be identified if spatially contiguous density of whatever sign is found on top of the atoms of the ground state. Very often, signal with opposite sign is found in the vicinity of the atoms. In contrast, noise is randomly sprinkled around, and is not contiguous. The lSVs in Fig. 9.10 demonstrate this behavior. However, in the presence of noise, some of the signal may smear into insignificant lSV. Indeed, there is some small signal found in the fourth lSV (not shown), whereas the time course was generated using only three intermediates. The additional, structural information introduced by the dark state model facilitates the detection of this small signal. Even in the presence of substantial noise all the signal can be collected safely into the reconstructed matrix $\boldsymbol{A}'$ and discriminated from noise this way [95]. This noise filter property is used to enhance the signal to noise level in the time-dependent difference maps. For a more thorough discussion of the multiple origins of noise in difference maps see ref. [75].

At this point, the goal is still to determine the structures of the intermediates from the main components of the SVD. However, the noise filter property is a very important feature of the decomposition, which should be described first. We will come back to the structure determination later in this chapter. Since the SVD exhibits a noise filter property it can be used to determine better phases for the difference structure factor amplitudes.

After matrix $\boldsymbol{A}'$ has been successfully reconstructed using only the significant lSV, SV, and rSV, the vectors in this matrix contain noise-reduced difference maps. The maps are Fourier-transformed from which difference structure factors with amplitudes and, more important, with phases are obtained. These difference structure factors can be used in a phase recombination scheme outlined in Fig. 9.11 to generate improved difference structure factors $\Delta\mathbf{F}^{\mathrm{svd}}$ [21, 74, 75].

From these, new difference maps are calculated and the procedure iterated. The phase improvements are of the order of 10°–15° in the average [75]

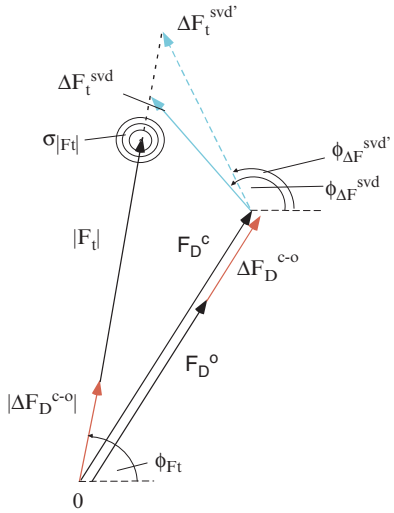

Fig. 9.11. Phase recombination in SVD-flattening: $\Delta\mathbf{F}_t{}^{\mathrm{svd}'}$ is derived from a Fourier-inversion of the difference map, which has been reconstructed from the most significant singular vectors/values. $\Delta\mathbf{F}_t{}^{\mathrm{svd}'}$ is added to the calculated structure factor $F_{\mathrm{D}}^{\mathrm{c}}$ derived from a very precise dark state model. This determines the phase ϕ_{Ft}. The Laue amplitude $|F_t|$ is added to $|\Delta F_{\mathrm{D}}{}^{\mathrm{c-o}}|$. $|\Delta F_{\mathrm{D}}{}^{\mathrm{c-o}}|$ is calculated by subtracting the observed dark state Laue amplitude, $F_{\mathrm{D}}{}^{\mathrm{o}}$, from $F_{\mathrm{D}}{}^{\mathrm{c}}$. This procedure corrects for crystal-to-crystal differences, which influence both the dark state Laue amplitudes and the time-dependent Laue amplitudes. The sum is aligned with the phase ϕ_{Ft}. However, the triangle formed by the origin 0 and the tips of $|\Delta F_{\mathrm{D}}{}^{\mathrm{c-o}}| + |F_t|$, $\mathbf{F}_{\mathrm{D}}{}^{\mathrm{c}}$, and $\mathbf{F}_{\mathrm{D}}{}^{\mathrm{c}} + \mathbf{F}_{\mathrm{D}}{}^{\mathrm{svd}'}$ does not close. It depends on the uncertainty, $\sigma_{|\mathrm{Ft}|}$, of $|F_t|$ relative to the mean uncertainty how $\Delta\mathbf{F}_t{}^{\mathrm{svd}'}$ with phase $\phi_{\Delta_F}{}^{\mathrm{svd}'}$ is corrected to $\Delta\mathbf{F}_t{}^{\mathrm{svd}}$ with phase $\phi_{\Delta_F}{}^{\mathrm{svd}}$. If $\sigma_{|\mathrm{Ft}|}$ is small $\Delta\mathbf{F}_t{}^{\mathrm{svd}}$ will end at the tip of $|\Delta F_{\mathrm{D}}{}^{\mathrm{oc}}| + |F_t|$, if $\sigma_{|\mathrm{Ft}|}$ is large $\Delta\mathbf{F}_t{}^{\mathrm{svd}}$ equals $\Delta\mathbf{F}_t{}^{\mathrm{svd}'}$. Here, a situation in between is sketched. $|\Delta\mathbf{F}_t{}^{\mathrm{svd}}|$ and $\phi_{\Delta_F}{}^{\mathrm{svd}}$ can be used to calculate phased difference maps

depending on the noise level. This procedure is called SVD-flattening in accordance to solvent flattening [108]. In contrast to solvent flattening, which utilizes a-priori chemical knowledge, SVD-flattening uses both chemical information and time information in an objective way for the phase improvement. The resulting maps can be contoured at higher sigma values, and the difference electron density features become better analyzable.

9.5.3 Transient Kinetics and Kinetic Mechanisms from the SVD

The right singular vectors (rSV) of the SVD describe the temporal variation of the corresponding left singular vectors (lSV), hence the kinetics, if any, is observed in the rSV. The time courses in the rSV are linear combinations of the true time courses of the difference electron density values in the

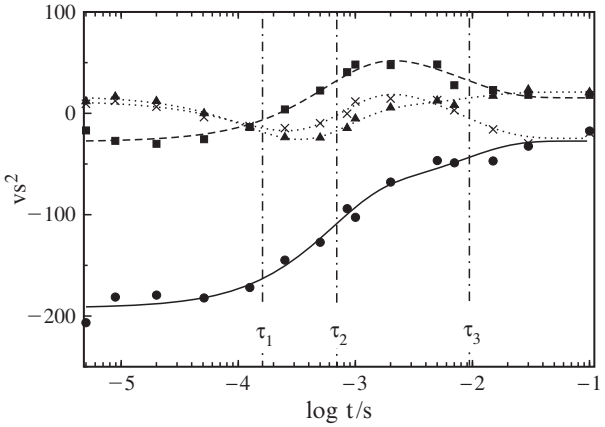

Fig. 9.12. The right singular vectors as determined from an SVD of 15 time-dependent difference electron density maps from the photo cycle of PYP on a logarithmic time-scale. The ordinate shows the magnitude of the elements of the rSV weighted with the square of the appropriate singular value. *Filed circle*, *filled square*, *filled trianlge*, *cross*, first through fourth significant right singular vector. *Solid lines*: global fit with three exponential funcitions. *Vertical dashed dotted lines* denote the relaxation times $\tau_1 \ldots \tau_3$ at 170 µs, 620 µs, and 8.5 ms

time-dependent difference maps. Likewise, the lSVs consist of linear combinations of the true intermediate difference electron densities. Hence both the mechanism and the intermediates cannot be observed directly.

However, the biggest advantage of the SVD is that one is able to observe a set of relaxation steps or phases in the rSV (Fig. 9.12). These steps are commonly referred to as relaxation times of the transient kinetics (see also Fig. 9.2). If only a few difference electron density features are taken into account, the relaxation times are usually difficult to find due to signal-to-noise problems. However, if all difference electron density features are analyzed simultaneously or globally by SVD the relaxation times can be extracted accurately.

The numerical values of the relaxation times are determined by globally fitting exponential functions to the significant rSVs (Fig. 9.12). The number of functions that have to be used to accomplish a satisfactory fit is equivalent to the number of relaxation times. After the fit the relaxation times are found at those times where the exponential functions, each, have decreased to $1/e$ of their initial value. They provide the key access to the determination of mechanism since they establish the (minimal) number of intermediates present. For example, in Fig. 9.12, three relaxation times are observable. Therefore, $N = 3$ intermediates plus the dark state are present and the general mechanism has therefore $N + 1$ states. If a fit with exponential functions, one per relaxation step each, can be performed successfully, a simple mechanism [109] holds.

The analysis proceeds by fitting a mechanism selected from the general mechanism to the rSV. One difficulty is that the rSV correspond to a linear combinations of the true time courses. Since this linear transformation is not known, it is not trivial to restore a common scale. Consequently, usually, the amplitudes of the rSV, which are the magnitudes of the vector elements in each rSV, are not used in the fitting procedure and the mechanism is fitted using only the relaxation times. It has been shown above that the relaxation rates are the (absolute) eigenvalues of the coefficient matrix. Usually, the number of rate coefficients is larger than the number of observed relaxation rates and the system of linear equations is underdetermined and all candidate mechanisms fit equally well. Structural and stoichiometric constraints can be used to further exclude or retain certain mechanism (see "posterior analysis" below).

Once the concentrations are successfully fitted to a particular rSV the relative contribution of the corresponding lSV to the time-independent difference electron density of an intermediate is determined [21, 75], see also [97]. The time-independent difference map of the intermediate state Ij ($j = 1 \ldots N$) is composed of the contribution of all significant lSV_s ($s = 1 \ldots S$) averaged over all measured time points t_i. In other words, the significant lSV_s are projected on the time independent difference electron density with the help of the fitted concentrations. This projection, of course, is done for all the N intermediates and results in a set of N time-independent difference maps, $\Delta\rho_{Ij}$, one for each intermediate state IS_j, from the time-course of time dependent difference maps.

To be acceptable the $\Delta\rho_{Ij}$ must be interpretable by a valid atomic model. If this cannot be done, the mechanism is not correct. An example is shown by Schmidt et al. [75]: consider the case that among the intermediates of the (unknown) true mechanism one intermediate employs two transients (see Fig. 9.2 for an example). If the chosen candidate mechanism is such that each intermediate has only one transient, mixtures, which cannot be explained by a unique atomic structure, are likely to occur in the time-independent difference maps and this particular mechanism must be discarded.

In addition, the number of suitable candidate mechanisms can be reduced by re-evaluating the mechanisms on the absolute scale present in the crystallographic data. Since this becomes possible only after the structures of intermediates are determined, this method has been called posterior analysis [21, 75] and is explained later in this chapter.

9.5.4 Determination of the Structures of the Intermediates

From the time-independent difference electron densities, $\Delta\rho_{Ij}$, the structures of the intermediates must be determined. It is not trivial to model a structure into difference maps due to the following reason: if atom A1 is replaced by another atom A2, atom A1 would leave a negative difference electron density. However, the positive difference electron density of atom A2 replenishes this

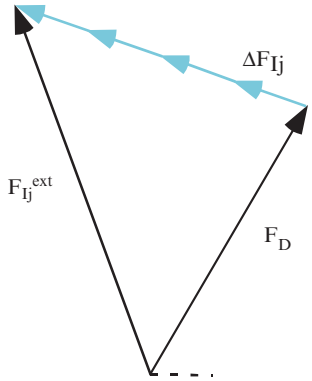

Fig. 9.13. Argand-diagram for the construction of extrapolated structure factors from phased difference structure factors. $\Delta\mathbf{F}_{Ij}$, difference structure factor obtained by Fourier-inversion of the time-independent difference map of intermediate Ij. $\mathbf{F}_{Ij}^{\text{ext}}$, extrapolated structure factor of intermediate Ij

negative "hole." The result is a flat difference electron density at this position and the structure cannot be modeled.

Extrapolated, conventional electron density maps, ρ^{ext}, however, show the correct electron densities and are, therefore, suitable to model structures. The extrapolated maps are constructed according to Fig. 9.13. The difference map, $\Delta\rho_{Ij}$, is Fourier-inverted and a multiple f of the resulting difference structure factor $\Delta\mathbf{F}_{Ij}$ is added to the structure factor of the dark state $\mathbf{F}_{\text{D}}$. Note, the $\Delta\mathbf{F}_{Ij}$ are phased due to the averaging properties of the SVD. The resulting extrapolated structure factor $\mathbf{F}_{Ij}^{\text{ext}}$ is used to calculate the extrapolated map ρ_{Ij}^{ext}. The factor f is determined in such a way that the extrapolated map is free of electron density at positions where prominent negative difference electron density features can be observed in the time-independent difference map $\Delta\rho_{Ij}$. Atomic models are aligned conveniently into the extrapolated map and are refined conventionally against the extrapolated amplitudes $\left|F_{Ij}^{\text{ext}}\right|$ or with the difference refinement [110,111] against $\mathbf{F}_{\text{D}}^{\text{calc}} + \Delta\mathbf{F}_{Ij}$. Both methods can be performed using, for example, the frequently used refinement program CNS [112].

9.5.5 Posterior Analysis

Posterior analysis partially resolves the degeneracy of the chemical, kinetic mechanisms by introducing at the same time both stoichiometric and structural constraints. It can be applied only after the structures of the intermediates are determined. The basic idea is that the observed time-dependent difference electron densities, which may or may not be SVD-flattened, are compared on a common and preferentially absolute scale with calculated difference electron densities.

A common scale, on which the time-independent difference electron density is represented, is lost after the SVD. Similar problems are described in the field of time-resolved spectroscopy and they are circumvented by introducing stoichometric constrains, which ultimately lead to the development of SVD with self modeling (SVD-SM) [101,102,114]. In time-resolved crystallography, the intermediate structures are used to restore a common, absolute scale. Then, both the structural and stoichometric constraints are automatically included.

On the absolute scale, the integrated electron density of an atom is directly related to its concentration or occupation. With the structures of the intermediates and the initial (dark) state, the concentrations determined from the particular candidate mechanism can be used to calculate difference maps, $\Delta\rho(k,t)^{\mathrm{calc}}$, which then become dependent on the time, on the mechanism, and the magnitude of the rate coefficients. For this purpose, structure factors $\mathbf{F}_{Ij}$ ($\mathbf{F}_{I0}, \mathbf{F}_{I1}..\mathbf{F}_{IN}$) are calculated from the structures of the dark state and the $j = 1 \ldots N$ intermediates, respectively. Difference structure factors $\Delta\mathbf{F}_{Ij}$ are determined by subtraction of the dark state structure factors $\mathbf{F}_{I0}$ from those of the intermediates. With the $\Delta\mathbf{F}_{Ij}$ time independent difference maps, $\Delta\rho_{Ij}^{\mathrm{calc}}$, one for each intermediate, are calculated by Fourier-synthesis. Candidate mechanism with rate coefficients k are employed as previously fitted to the rSV to determine the time-dependent concentrations, $c_{\mathrm{j}}(k,t)$, of the intermediates. Difference maps, which now depend on time and the rate coefficients, are derived on the absolute scale from these concentrations according to (9.5):

$$\Delta\rho(k,t)^{\mathrm{calc}} = \sum_{j=1}^{N} c_j(k,t)\Delta\rho_{Ij}^{\mathrm{calc}}. \tag{9.5}$$

These maps are fitted to the observed difference maps by varying the rate coefficients using (9.6) as the kernel of a minimization problem:

$$\sum_{t=1}^{T}\sum_{m=1}^{M} \frac{1}{\left\langle |\Delta\rho(t)|^{\mathrm{SVD}} \right\rangle} \left(\Delta\rho_m(t)^{\mathrm{SVD}} - C_{\mathrm{PA}}\Delta\rho_m(k,t)^{\mathrm{calc}}\right)^2, \tag{9.6}$$

where $\Delta\rho_m(k,t)^{\mathrm{calc}}$ is the calculated time- and mechanism-dependent difference electron density values and $\Delta\rho_m(t)^{\mathrm{SVD}}$ is the SVD-flattened difference electron density values at grid point m of the difference maps [21]. If all grid points were considered, the noise features would largely influence the fit, since the signal is usually concentrated only at a small number of grid points [73]. Therefore, those M grid points can be taken into consideration where the signal exceeds 2σ or falls below -2σ. The calculation is done at all T time points. The fit is weighted by the average of the absolute difference electron density values in the observed difference maps $\langle|\Delta\rho(t)^{\mathrm{SVD}}|\rangle$. This ensures that maps at later time points, where the signal may have been already greatly reduced, are equally considered.

C_{PA} is a linear fit parameter and corresponds to the concentration of activated molecules at the beginning of the analysis. More formally, C_{PA} is the constant of integration in the solutions to the coupled differential integrations, which describe the mechanism. C_{PA} can be determined from the first one or two pairs of observed and calculated maps.

9.5.6 Verification of the Functionality of the SVD-Driven Analysis by Mock Data

Since it was unknown how an SVD driven analysis performs with time-resolved X-ray data, mock data were used to determine the limits of such an analysis [75]. Numerous realistic influences on the mock data were considered:

1. The signal to noise levels in the difference amplitudes, which were used to calculate the difference electron densities, were varied from 0.5 times to about 10 times the signal-to-noise values observed in experimental structure amplitudes.
2. Simple and more complex chemical, kinetic mechanisms were taken into account. Typically 20% reaction initiation was assumed, which is a realistic value for time-resolved experiments. As a result, for some mechanisms, the peak concentrations of some intermediates were not higher than about 5%.
3. The number of time points was varied between five time points to only one time point per logarithmic decade.
4. The extent of reaction initiation was also varied randomly from time point to time point. In addition to the random noise in the structure amplitudes, substantial systematic noise across the time axis was modeled. In some mock time-courses the reaction initiation was allowed to vary even between 5 and 17%. Hence, the variation is much larger and the extent of reaction initiation is much smaller than expected from the experiment.

The results can be subsumed as follows. First of all, the SVD is quite insensitive to random noise in the difference amplitudes. Only in the case of unrealistically large noise levels the SVD-driven analysis fails, whereas on the experimentally observed noise levels the SVD analysis is stable. Intermediates with only 5% peak occupation can be extracted. This result attenuates the common belief that such low occupancy values are not suitable for a structural characterization.

Three to five time points per logarithmic decade are usually sufficient for a successful data analysis. The variation of reaction initiation from time point to time point has a large influence. If it is (unrealistically) large, the analysis might fail even if the noise in the structure amplitudes is moderate. However, if an outlier is present, which means that the reaction initiation at single time points is either too high or too low compared to the average, it can be identified and corrected. Once the results with mock data showed how the

SVD performs with X-ray data [75], the analysis was applied to experimental time-resolved data [21].

9.6 The SVD Analysis of Experimental Time-Resolved Data

In the next paragraphs the first application of the SVD based analysis [21] to 15 difference maps collected on the late photocycle of PYP (from 5 µs to 100 ms) is described. All new methods listed in this chapter (from the SVD-flattening to the posterior analysis) were used to find the structures of the intermediates and identify a set of plausible mechanisms. Newest results on almost the entire photocycles of the PYP Wild-type [22] and the E46Q mutant [96] are also summarized shortly.

9.6.1 SVD-Flattening

Figure 9.14 shows on the left side selected observed (weighted) difference maps near the chromophore of the PYP. Numerous features are present. Noticeable, there are also some negative features (red difference electron density) distant from the dark state model (in yellow). Since the negative features are expected to be exclusively on top of the dark state atoms, it is evident that these features must result from the noise in the maps. However, if the time-series of maps is decomposed by SVD and the first four significant singular values and vectors are used for a resynthesis, the maps on the right side of Fig. 9.14 are obtained. Here, the number of negative features on incorrect positions is reduced. In addition, the magnitude of the remaining features is greatly enhanced. As a consequence the signal-to-noise ratio is substantially larger and the maps can be contoured on a larger level and can be much better inspected and analyzed further by a mechanistic analysis.

9.6.2 The Mechanistic Analysis of the PYP data

In Fig. 9.12 the right singular vectors (rSV) as derived from an SVD-analysis of the 15 difference maps spanning the temporal range from 5 µs to 100 ms are shown. The fit with exponential functions reveals three relaxation rates at 170 µs, 620 µs, and 8.5 ms. Hence the general mechanism consists of three intermediate states plus the ground state. As mentioned earlier, a common or absolute scale is lost in the rSV. So, only the relaxation rates but not the amplitudes of the rSV can initially be used to fit the mechanism. Since the relaxation rates are the eigenvalues of the coefficient matrix, the exponential approach is unambiguous only if the number of reaction coefficients is equal to the number of relaxation times. This is true only if an irreversible, sequential mechanism holds (Fig. 9.1, mechanism I). For the present PYP

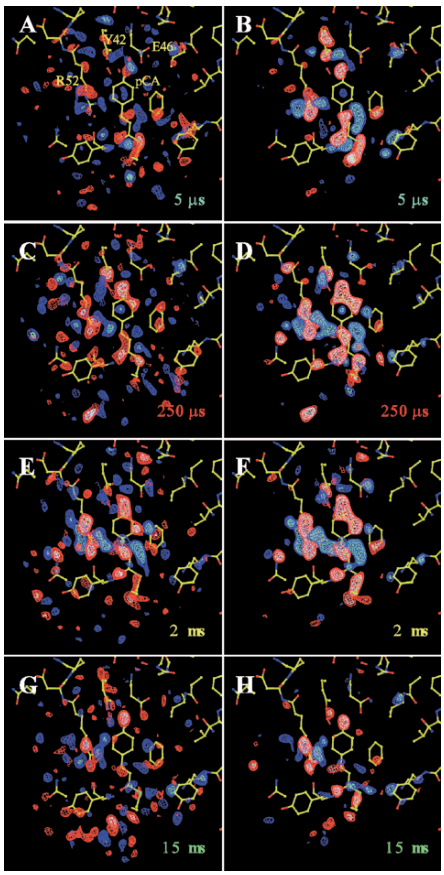

Fig. 9.14. Effect of the SVD-flattening. (**a, c, e, g**) Experimental, weighted difference electron density at 5 µs, 250 µs, 2 ms, and 15 ms; (**b, d, f, h**) same as left panels after SVD-flattening. Contour levels left: red/white, $-2\sigma/-3\sigma$; blue/cyan, $2\sigma/3\sigma$; right: red/white, $-3\sigma/-4\sigma$; blue/cyan, $3\sigma/4\sigma$. Structure of the dark state in yellow. Amino acid residues Tyr42, Glu46, Arg52 as well as the chromophor pCA are marked in (**a**)

data the sequential mechanism employs three rate coefficients which have to be related to three relaxation rates. It is shown later in this chapter that this simple mechanism does not hold. More complicated mechanisms with more rate coefficients must be employed.

It is useless to fit all rate coefficients that may occur in a general mechanism. Nevertheless, it makes sense to fit candidate mechanisms with four rate coefficients (Fig. 9.1, mechanisms II to IV). Some of these mechanisms should be able to reproduce the true but unknown concentrations so well that they become indistinguishable from the true concentrations, given the noise in the data. Since the fit is underdetermined, some user input is required. For example, it should be possible to fix the ratio of some rate coefficients, or one

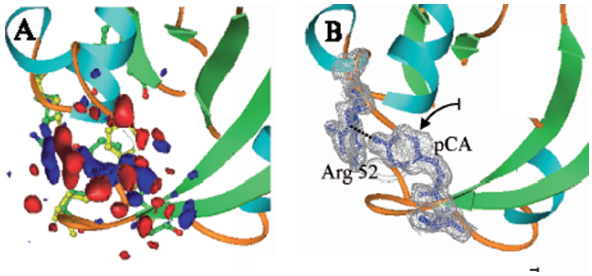

Fig. 9.15. (**a**) Pure, authentic difference electron density extracted with the SVD-driven analysis. Contour levels: red, -4σ; blue, $+4\sigma$. Structure of the dark state in yellow. (**b**) Extrapolated, conventional electron density of one of the identified intermediates, structure of the intermediate in red. Contour level $1.5\,\sigma$, chromophore, and Arg52 depicted in blue

should think about reasonable start values for the rate coefficients and do not allow them to vary too much. In any case, the extracted difference maps must be homogeneous.

In the late photocycle of PYP, three time-independent difference maps were found, from which map 2 and 3 where homogeneous and the first, earlier map was not. Since the analysis was started in the middle of the photocycle, an early intermediate most likely mix into the first map. This admixture can then be separated only if data at earlier time points become available. In Fig. 9.15 the second intermediate identified in this study is shown.

9.6.3 The Structures of the Intermediates in the Late Photocycle Between 5 µs and 100 ms

The structure of the PYP dark state is sketched in yellow as a reference in Fig. 9.15a. The chromophore is in the *trans* configuration. The pCA head is fixed by a hydrogen bond network between Tyr42 and Glu46. After 5 µs the head is disrupted from this network in part of the molecules. Residual features suggest that still another intermediate with intact head hydrogen bonds is present (not shown). The next intermediate populates between 125 µs and 2 ms. Here, the head has been flipped to the solvent (Fig. 9.15b). Arg52 is pushed into the solvent space already on a 100 µs time scale. In the third intermediate which is populated between 1 and 100 ms Arg52 is still swung out and the chromophore head points to the solvent. Although *cis*, the foot of the chromophore in IS3 is closer to the dark state position than in IS2.

9.6.4 Plausible Kinetic Mechanisms

Posterior analysis (see Sect. 5.5) was applied to the PYP data to reduce the number of possible kinetic mechanisms. Time-dependent difference maps were

calculated (9.5) from IS1, the two homogenous structures IS2 and IS3, the dark state structure, and the kinetic mechanisms listed in Fig. 9.1 as explained earlier. Figure 9.16 shows the result of the fit of the calculated maps to the SVD-flattened maps. The extent of reaction initiation was about 25%.

The irreversible, sequential mechanism could be discarded based on gross residual electron density. However, the others were still indistinguishable. Figure 9.17 compares the concentrations of the intermediates derived from the incompatible mechanism I (Fig. 9.17a) with those calculated from one of the compatible mechanisms (Fig. 9.17b). Obviously, the concentrations of

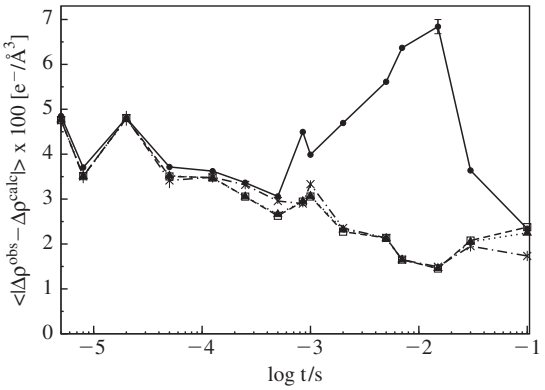

Fig. 9.16. Mean squared residual density in the residual maps $\Delta\rho^{\mathrm{svd}} - \Delta\rho^{\mathrm{calc}}$ after the fit of the different candidate mechanisms from Fig. 9.1. *Filled circle*, candidate mechanism I; *open square*, candidate mechanism II; *filled square*, candidate mechanism III; *cross*, candidate mechanism IV. The significance of the fit can be estimated from the error bar, which was calculated from the variation (σ^{svd}) of the SVD-flattened difference maps around their mean values

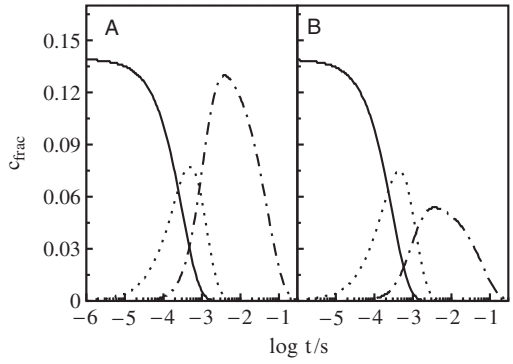

Fig. 9.17. Fractional concentrations (C_{frac}) of the intermediates after posterior analysis. *Solid*, *dotted*, *dashed dotted lines*: concentration of intermediates IS1 (a), IS2 (b), IS3 (c), respectively. (**a**) Mechanism I; (**b**) mechanism III in Fig. 9.1

intermediate C in mechanism I deviate grossly in magnitude from the other compatible mechanism. With mechanism I, large residual density features were observed in the residual maps $\Delta\Delta\rho_t = {\Delta\rho_t}^{\mathrm{svd}} - {\Delta\rho_t}^{\mathrm{calc}}$, whereas the residual maps obtained from the other candidate mechanisms were almost empty. Based on this, the irreversible sequential mechanism I in Fig. 9.3 could be disregarded in favor of the other, however, still degenerate, candidate mechanisms.

9.6.5 The Entire Photocycles of the Wild-Type PYP and its E46Q-Mutant

After the successful analysis of the slower part, the entire reaction cycles of the Wild-Type PYP and of its E46Q mutant was investigated on time scales from 1 ns to 2 s and 10 ns to 100 ms, respectively [22, 95, 96]. In the Wild-Type five intermediates were identified and their structures determined. The structure of the shortest-lived one, I_{cp}, most likely resembles that of ${I_0}^{\#}$ (Fig. 9.18). This intermediate decays on a time scale of 10 ns in favor of an admixture of two distinct intermediates called $\mathrm{pR_{CW}}$ and $\mathrm{pR_{E46Q}}$ (Fig. 9.18). The structure of the former is similar to I_{cp} but has a distorted chromophore geometry, the structure of the latter is also identified in the E46Q mutant. This admixture has also been observed in the earlier study [21] on the microsecond time scale. The admixture decays on the microsecond time scale to typical pB states, structures of which are called $\mathrm{pB_1}$ and $\mathrm{pB_2}$. The dark state pG is reached after 1 s.

In the E46Q-mutant intermediate I_{cp} was not observed most likely because it has already relaxed at the fastest time of 10 ns. An intermediate similar to I_{cw} was also not observed in this mutant. Two states with structures similar to $\mathrm{pR_{E46Q}}$ in Fig. 9.18 were identified on the nanosecond time range. Two pB states are populated up to about 50 ms. By focusing on the chromophore the mutant apparently cycles about one order of magnitude faster than the Wild-Type. In the Wild-Type as well as in the mutant, a step-wise dislocation of the chromophore from its hydrogen bonding network is observed. The final, putatively signaling state in the Wild-Type is $\mathrm{pB_2}$. In this state structural changes at the N-terminus of PYP were identified. However, in the mutant

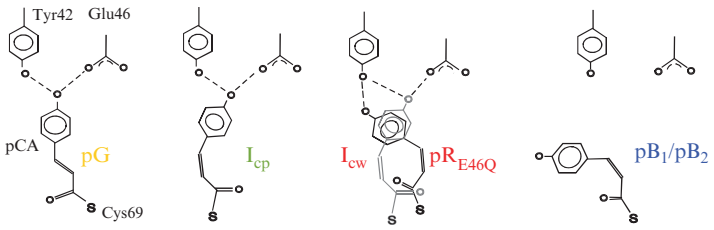

Fig. 9.18. Structures of the chromophore and the nearest environment in the photocycle of the Wild-Type PYP)

an additional, putatively optically silent, species with the chromophore in the dark position but with signal on the protein exists, which relaxes on a similar same time scale as pB_2 in the Wild-Type.

9.7 Picosecond Time Resolution and Beyond

Picoseconds

With existing third generation synchrotrons, X-ray pulses in the range of 100 ps can be exploited [45, 46] to observe short lived intermediates at the beginning of a reaction cycle such as I_0 of PYP (Fig. 9.8). At room temperature and at times faster than some nanoseconds the relaxation kinetics of proteins becomes complicated [81, 114]. This is because the time scale of the time-resolved experiment and the time scale of the diffusive processes within the biomolecules [115] match. The simple picture of chemical kinetics has to be replaced by another type of relaxation kinetics, which explicitly considers these specific motions [116]. In some respect, room temperature experiments at ultrafast times have some similarities to experiments at low temperatures and longer times [31]. Extending time-resolved protein crystallography to the picosecond time-range will give new insights in the physical nature of this very elementary behavior of biomolecules. Traveling of signal from the source, or epicenter, of the excitation through the molecule [74] can be followed in a larger time-range, giving direct evidence for the viscoelastic properties of these molecules.

Femtoseconds

The femtosecond time scale is a challenge for the design of hard X-ray sources. Two concepts exist in the moment, which will make these experiments possible.

1. *The plasma sources*: X-rays are generated by focusing a powerful ultrashort laser pulse on a thin band of metal such as copper. A plasma is formed at the surface, electrons of which are accelerated deeper into the material, hence generating intense, but quite divergent [117] X-ray radiation during the pulse duration of the laser. These X-ray pulses can be used for femtosecond time-resolved experiments.
2. *The free-electron laser (FEL)*: In this machine electrons are accelerated to several GeV either by a linear accelerator or by other means [118, 119] and channeled into an undulator system, which can be very long (in the order of 100–200 m). Here, in a self amplified spontaneous emission (SASE) process [120], hard X-ray beams with enormous brilliance are produced. About 10^{33} photons/(s mm^2 mrad^2 0.1% bandwidth) of wavelength 1 Å are expected for the European FEL at the "Deutschen Elektronen Synchrotron" (DESY) in Hamburg [70]. This brilliance is at least 10 orders of

magnitude higher compared to the strongest existing synchrotron X-ray sources at that wavelength. Depending on the design, the expected pulse duration is in the 10–100 fs range and the number of X-ray photons per pulse is in the order of 10^{13}. This opens the opportunity to collect a diffraction image even from single molecules [121, 122]. For time-resolved experiments pump-probe sequences on tiny crystals are conceivable. This would facilitate, for example, a homogenous reaction initiation. However, the small bandwidth of this device and the complicated spectrum of the XFEL radiation complicate the use of the Laue method and the experiments probably must be conducted in a different way. In addition, issues like radiation damage or the orientation of the sample at the intersection of the laser and the X-ray beams have to be addressed. Then, however, the gap between femtosecond and picosecond can be closed.

Protein crystals can be envisioned as containers that preserve a periodic arrangement of small molecules around which they wrap. Because proteins are highly flexible, the small molecules can react in them without risking the integrity of the crystalline lattice. Hence, numerous chemical reactions most likely become directly observable. By pushing the limits with respect to time and spatial resolution, the direct four-dimensional observation of electronic displacements in these small molecules will become possible not only but also on the femtosecond time scale, which probably can even be expanded to attoseconds in the future [123]. The technical challenges, however, can not be foreseen by now.

9.8 More Applications

The time-resolved X-ray structure analysis has evolved to a true structure determining method. It is generally applicable for the investigation of fast and slow processes in protein crystals. Cyclic, light initiated processes in PYP and bacteriorhodopsin [28, 124, 125] are paradigms for such types of reactions. In bacteriorhodopsin mixed states have been reported to generate problems with the determination and attribution of structure [28]. The application of the time-resolved X-ray structure analysis restores the time-scale to separate this admixture. This will lead to a cleaner and unambiguous determination of the structures of the intermediates. In any case, the proton pump bacteriorhodpsin is a good example to point out future challenges, since the pumped proton usually cannot be observed in X-ray structure determination and must be derived spectroscopically [126] or calculated theoretically [127,128] from the structure of the heavier atoms. Time-resolved neutron crystallography would directly show its position. However, for this purpose extremely strong, preferentially pulsed neutron sources, which do not exist at present, are required.

The new methods become indispensable if noncyclic reactions in (potentially pharmacologically relevant) enzymes are to be depicted. Apart from the

Michaelis complex numerous other intermediates along the reaction pathway are likely to be identified and structurally characterized. Activation of an initially inactive (caged) substrate will usually be necessary to initiate the reaction [9, 14, 129]. However, after activation and consumption of the substrate the product must be washed away and the caged compound reloaded. Multiple X-ray exposures, which to date are still necessary to collect one reflection pattern, must be avoided. Any attempt to enhance the primary beam flux density is therefore highly desirable. New devices such as the mentioned plasma sources, the free electron lasers for hard X-rays, advanced optics for shaping and focusing the incident X-ray beam, and new detector systems [70] will pave the way to a general and fast application of time-resolved crystallography to numerous reactions in proteins.

Acknowledgements. M.S. is most grateful to Keith Moffat whose insights in protein structure and chemical kinetics created a new field in the biological sciences. The author thanks Vukica Srajer and Sudar Rajagopal for valuable discussions. The work was sponsored by the DFG, Sonderforschungsbereich 533 and grant SCHM 1423/2-1.

References

1. M. von Laue, *Physik und Chemie und ihre Anwendungen in Einzeldarstellungen*, Band IV, 2nd edn. (Akademische Verlagsgesellschaft Geest & Portig, Leipzig, 1948)
2. J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips, Nature **181**, 662 (1958)
3. H. Muirhead, M.F. Perutz, Nature **199**, 633 (1963)
4. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucleic Acids Res. **28**, 235 (2000)
5. J. Drenth, *Principles of Protein Crystallography* (Springer, New York, 1994)
6. I. Schlichting, J. Berendzen, K. Chu, A.M. Stock, S.A. Maves, D.E. Benson, R.M. Sweet, D. Ringe, G. Petsko, S. Sligar, Science **287**, 1615 (2000)
7. L. Ujj, S. Devanathan, T.E. Meyer, M.A. Cusanovich, G. Tollin, G.H. Atkinkson, Biophys. J. **75**, 406 (1998)
8. J.L. Martin, A. Migus, C. Poyart, Y. Lecarpentier, R. Astier, A. Antonetti, Proc. Natl Acad. Sci. USA **80**, 173 (1983)
9. I. Schlichting, S.C. Almo, G. Rapp, K. Wilson, K. Petrakos, A. Lentfer, A. Wittinghofer, W. Kabsch, E.F. Pai, G.A. Petsko, R.S. Goody, Nature **345**, 309 (1990)
10. J.R. Helliwell, Y.P. Nieh, J. Habash, P.F. Faulder, J. Raftery, M. Cianci, M. Wulff, A. Hädener, Faraday Discuss. **122**, 131 (2002)
11. K. Moffat, R. Henderson, Curr. Opin. Struct. Biol. **5**, 656–663 (1995)
12. B.L. Stoddard, Methods **24**, 125 (2001)
13. J. Hajdu, R. Neutze, T. Sjögren, K. Edman, A. Szöke, R.C. Wilmouth, C.M. Wilmot, Nat. Struct. Biol. **7**, 1006 (2000)
14. I. Schlichting, K. Chu, Curr. Opin. Struct. Biol. **10**, 744 (2000)

15. C.C.H. Chen, O. Herzberg, J. Mol. Biol. **224**, 1103 (1992)
16. S.B. Long, P.J. Casey, L.S. Beese, Nature **419**, 654 (2002)
17. T.Y. Teng, V. Srajer, K. Moffat, Nat. Struct. Biol. **1**, 701 (1994)
18. U.K. Genick, S.M. Soltis, P. Kuhn, I.L. Canestrelli, D.E. Getzoff, Nature **392**, 206 (1998)
19. K. Nienhaus, A. Ostermann, U. Nienhaus, F. Parak, M. Schmidt, Biochemistry **44**, 5095 (2005)
20. J.M. Bolduc, D.H. Dyer, W.G. Scott, P. Singer, R.M. Sweet, D.E. Koshland, B.L. Stoddard, Science **268**, 1312 (1995)
21. M. Schmidt, R. Pahl, V. Srajer, S. Anderson, H. Ihee, Z. Ren, K. Moffat, Proc. Natl Acad. Sci. USA **101**, 4799 (2004)
22. H. Ihee, S. Rajagopal, V. Srajer, R. Pahl, M. Schmidt, F. Schotte, P.A. Anfinrud, M. Wulff, K. Moffat, Proc. Natl Acad. Sci. USA **102**, 7145 (2005)
23. M.H.B. Stowell, T.M. McPhillips, D.C. Rees, S.M. Soltis, E. Abresch, G. Feher, Science **276**, 812 (1997)
24. T.Y. Teng, K. Moffat, J. Appl. Crystallogr. **31**, 252 (1998)
25. T.Y. Teng, V. Srajer, K. Moffat, Biochemistry **36**, 12087 (1997)
26. A. Ostermann, R. Waschipky, F.G. Parak, G.U. Nienhaus, Nature **404**, 205 (2000)
27. T.Y. Teng, K. Moffat, J. Synchrotron Radiat. **9**, 198 (2002)
28. R. Neutze, E. Pebay-Peyroula, K. Edman, A. Royant, J. Navarro, E.M. Landau, Biochim. Biophys. Acta **1565**, 144 (2002)
29. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics* (Portland Press, London, 1999)
30. G.M. Fleck, *Chemical Reaction Mechanism* (Holt, Rinehart and Winston, New York, 1971)
31. R.H. Austin, K.W. Beeson, L. Eisenstein, H. Frauenfelder, I.C. Gunsalus, Biochemistry **14**, 5355 (1975)
32. N. Agmon, J.J. Hopfield, J. Chem. Phys. **79**, 2042 (1983)
33. V. Srajer, L. Reinisch, P.M. Champion, J. Am. Chem. Soc. **110**, 6656 (1988)
34. R.J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B.R. Cowen, D. Ehrenstein, H. Frauenfelder, J.B. Johnson, D.C. Lamb, S. Luck, J.R. Mourant, G.U. Nienhaus, P. Ormos, R. Philipp, A. Xie, R.D. Young, Biochemistry **30**, 3988 (1991)
35. J.I. Steinfeld, J.S. Francisco, W.L. Hase, *Chemical Kinetics and Dynamics* (Prentice Hall, Englewood Cliffs NJ, 1989)
36. F.A. Matsen, J.L. Franklin, J. Am. Chem. Soc. **72**, 3337 (1950)
37. R. Brudler, R. Rammelsberg, T.T. Woo, E. Getzoff, K. Gerwert, Nat. Struct. Biol. **8**, 265 (2001)
38. A. Xie, L. Kelemen, J. Hendriks, B.J. White, K.J. Hellingwerf, W.D. Hoff, Biochemistry **40**, 1510 (2001)
39. Z. Ren, D. Bourgeois, J.R. Helliwell, K. Moffat, V. Srajer, B.L. Stoddard, J. Synchrotron Radiat. **6**, 891 (1999)
40. K. Moffat, Annu. Rev. Biophys. Biophys. Chem. **18**, 309 (1989)
41. P.F. Lindley, Acta Crystallogr. D. Biol. Crystallogr. **55**, 1654 (1999)
42. M. Wulff, F. Schotte, G. Naylor, D. Bourgeois, K. Moffat, G. Mourou, Nucl. Instrum. Methods Phys. Res. A **398**, 69 (1997)
43. D. Bourgeois, U. Wagner, M. Wulff, Acta Crystallogr. D. Biol. Crystallogr. **56**, 973 (2000)

44. V. Srajer, S. Crosson, M. Schmidt, J. Key, F. Schotte, S. Anderson, B. Perman, Z. Ren, T.Y. Teng, D. Bourgeois, M. Wulff, K. Moffat, J. Synchtrotron Radiat. **7**, 236 (2000)
45. F. Schotte, M. Lim, T.A. Jackson, A.V. Smirnov, J. Soman, J.S. Olson, G.N. Philips Jr., M. Wulff, P.A. Anfinrud, Science **300**, 1944 (2003)
46. F. Schotte, J. Soman, J.S. Olson, M. Wulff, P.A. Anfinrud, J. Struct. Biol. **147**, 235 (2004)
47. D.M.E. Szebenyi, D.H. Bilderback, A. LeGrand, K. Moffat, W. Schildkamp, B. Smith Temple, T.Y. Teng, J. Appl. Crystallogr. **25**, 414 (1992)
48. D. Bourgeois, T. Ursby, M. Wulff, C. Pradervand, A. Legrand, W. Schildkamp, S. Laboure, V. Srajer, T.Y. Teng, M. Roth, K. Moffat, J. Synchrotron Radiat. **3**, 65 (1996)
49. H.D. Bellamy, E.H. Snell, J. Lovelace, M. Pokross, G.E. Borgstahl, Acta Crystallogr. D. Biol. Crystallogr. **56**, 986 (2000)
50. A. Ostermann, I. Tanaka, N. Engler, N. Niimura, F. Parak, Biophys. Chem. **95**, 183 (2002)
51. N. Engler, A. Ostermann, N. Niimura, F. Parak, Proc. Natl Acad. Sci. USA **100**, 10243 (2003)
52. J.L. Amoros, M.J. Buerger, M. Canut de Amoros, *The Laue Method* (Academic Press, New York, 1975)
53. J. Hajdu, P.A. Machin, J.W. Campbell, T.J. Greenhough, I.J. Clifton, S. Zurek, S. Gover, L.N. Johnson, M. Elder, Nature **329**, 178 (1987)
54. D.W.J. Cruickshank, J. Helliwell, K. Moffat, Acta Crystallogr. A **47**, 352 (1991)
55. H.D. Bartunik, H.H. Bartsch, H. Qichen, Acta Crystallogr. A **48**, 180 (1992)
56. Z. Ren, K. Moffat, J. Appl. Crystallogr. **28**, 461 (1995)
57. D. Bourgeois, Acta Crystallogr. D. Biol. Crystallogr. **55**, 1733 (1999)
58. D.W.J. Cruickshank, J. Helliwell, K. Moffat, Acta Crystallogr. A **43**, 656 (1987)
59. Z. Ren, K. Moffat, J. Appl. Crystallogr. **28**, 482 (1995)
60. G.P. Bourenkov, Thesis (2003)
61. G.P. Bourenkov, A.N. Popov, H.D. Bartunik, Acta Crystallogr. A **52**, 797 (1996)
62. S. Arzt, J. Cambell, M.M. Harding, Q. Hao, J. Helliwell, J. Appl. Crystallogr. **32**, 554 (1999)
63. I. Schlichting, Biospektrum **2**, 153 (2003)
64. B.L. Stoddard, G.K. Faber, Structure **3**, 991 (1995)
65. I. Schlichting, R. Goody, Meth. Enzymol. **277**, 467 (1997)
66. M.D. Perry, G. Mourou, Science **264**, 917 (1994)
67. B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics* (Wiley, New York, 1991)
68. G. Steinmeyer, D.H. Sutter, L. Gallmann, N. Matuschek, U. Keller, Science **286**, 1507 (1999)
69. M.H. Dunn, M. Ebrahimzadeh, Science **286**, 1513 (1999)
70. M. Altarelli, R. Brinkmann, M. Chergui, W. Decking, B. Dobson, S. Düsterer, G. Grübel, W. Graeff, H. Graafsma, Hajdu J., et al. DESY 2006 (2006)
71. J.W. Campbell, J. Appl. Crystallogr. **28**, 228 (1995)
72. T. Ursby, D. Bourgeois, Acta Crystallogr. A **53**, 564 (1997)
73. V. Srajer, Z. Ren, T.Y. Teng, M. Schmidt, T. Ursby, D. Bourgeois, C. Pradervand, W. Schildkamp, M. Wulff, K. Moffat, Biochemistry **40**, 13802 (2001)

74. Z. Ren, B. Perman, V. Srajer, T.Y. Teng, C. Pradervand, D. Bourgeois, F. Schotte, T. Ursby, R. Kort, M. Wulff, K. Moffat, Biochemistry **40**, 13788 (2001)
75. M. Schmidt, S. Rajagopal, Z. Ren, K. Moffat, Biophys. J. **84**, 2112 (2003)
76. I. Schlichting, J. Berendzen, G.N. Phillips, R.M. Sweet, Nature **371**, 808 (1994)
77. H. Hartmann, S. Zinser, P. Komninos, R.T. Schneider, G.U. Nienhaus, F. Parak, Proc. Natl Acad. Sci. USA **93**, 7013 (1996)
78. V. Srajer, T.Y. Teng, T. Ursby, C. Pradervand, Z. Ren, S. Adachi, W. Schildkamp, D. Bourgeois, M. Wulff, K. Moffat, Science **274**, 1726 (1996)
79. K. Chu, J. Vojtchovsky, B.H. McMahon, R.M. Sweet, J. Berendzen, I. Schlichting, Nature **403**, 921 (2000)
80. D. Bourgeois, B. Vallone, F. Schotte, A. Arcovito, A.E. Miele, G. Sciara, M. Wulff, P. Anfinrud, M. Brunori, Proc. Natl Acad. Sci. USA **100**, 8704 (2003)
81. M. Schmidt, K. Nienhaus, R. Pahl, A. Krasselt, U. Nienhaus, F. Parak, V. Srajer, Proc. Natl Acad. Sci. USA **13**, 11704 (2005)
82. J.S. Olson, G.N. Phillips Jr., J. Biol. Chem. **271**, 17593 (1996)
83. T.E. Meyer, Biochim. Biophys. Acta **806**, 175 (1985)
84. W.W. Sprenger, W.D. Hoff, J.P. Armitage, K.J. Hellingwerf, J. Bacteriol. **175**, 3096 (1993)
85. G.E.O. Borgstahl, D.R. Williams, D.E. Getzoff, Biochemistry **34**, 6278 (1995)
86. J.L. Pellequer, K.A. Wager-Smith, S.A. Kay, E.D. Getzoff, Proc. Natl Acad. Sci. USA **95**, 5884 (1998)
87. W.D. Hoff, I.H.M. van Stokkum, H.J. van Ramesdonk, M.E. von Brederode, A.M. Brouwer, J.C. Fitch, T.E. Meyer, R. van Grondelle, K.J. Hellingwerf, Biophys. J. **67**, 1691 (1994)
88. K. Ng, E.D. Getzoff, K. Moffat, Biochemistry **34**, 879 (1995)
89. K. Takeshita, Y. Imamoto, M. Kataoka, F. Tokunaga, M. Terazima, Biochemistry **41**, 3037 (2002)
90. K. Hellingwerf, J. Hendriks, T. Gensch, J. Biol. Phys. **28**, 395 (2003)
91. K. Hellingwerf, J. Hendriks, T. Gensch, J. Phys. Chem. A. **107**, 1082 (2003)
92. B. Perman, V. Srajer, Z. Ren, T.Y. Teng, C. Pradervand, T. Usrby, D. Bourgeois, F. Schotte, M. Wulff, R. Kort, K.J. Hellingwerf, K. Moffat, Science **279**, 1946 (1998)
93. U.K. Genick, G.E. Borgstahl, K. Ng, Z. Ren, C. Pradervand, P.M. Burke, V. Srajer, T.Y. Teng, W. Schildkamp, D.E. McRee, K. Moffat, D.E. Getzoff, Science **275**, 1471 (1997)
94. S. Rajagopal, K.S. Kostov, K. Moffat, J. Struct. Biol. **147**, 211 (2004)
95. S. Rajagopal, M. Schmidt, S. Anderson, K. Moffat, Acta Crystallogr. D. Biol. Crystallogr. **60**, 860 (2004)
96. S. Rajagopal, S. Anderson, V. Srajer, M. Schmidt, R. Pahl, K. Moffat, Structure **13**, 55 (2005)
97. E.R. Henry, J. Hofrichter, Meth. Enzymol. **210**, 129 (1992)
98. A. Ansari, C.M. Jones, E.R. Henry, J. Hofrichter, W.A. Eaton, Science **256**, 1796 (1992)
99. A. Ansari, C.M. Jones, E.R. Henry, J. Hofrichter, W.A. Eaton, Biochemistry **33**, 5128 (1994)
100. E.R. Henry, Biophys. J. **72**, 652 (1997)
101. L. Zimanyi, A. Kulcsar, J.K. Lanyi, D.F. Sears Jr., J. Saltiel, Proc. Natl Acad. Sci. USA **96**, 4408 (1999)

102. L. Zimanyi, A. Kulcsar, J.K. Lanyi, D.F. Sears Jr., J. Saltiel, Proc. Natl Acad. Sci. USA **96**, 4414 (1999)
103. L. Zimanyi, J.K. Lanyi, Biophys. J. **64**, 240 (1993)
104. O. Alter, P.O. Brown, D. Botstein, Proc. Natl Acad. Sci. USA **97**, 10101 (2000)
105. P. Doruker, A.R. Atilgan, I. Bahar, Proteins **40**, 512 (2000)
106. T.D. Romo, J.B. Clarage, D.C. Sorensen, G.N. Phillips Jr., Proteins **22**, 311 (1995)
107. R. Henderson, J.K. Moffat, Acta Crystallogr. B **27**, 1414 (1971)
108. B.C. Wang, in *Diffraction Methods for Biological Macromolecules*, ed. by H. Wyckoff, C.H.W. Hirs, S.N. Timasheff. Methods in Enzymology, vol. 115 (Academic Press, New York, 1985)
109. K. Moffat, Chem. Rev. **101**, 1569 (2001)
110. T.C. Terwilliger, J. Berendzen, Acta Crystallogr. D. Biol. Crystallogr. **51**, 609 (1995)
111. T.C. Terwilliger, J. Berendzen, Acta Crystallogr. D. Biol. Crystallogr. **52**, 1004 (1996)
112. A.T. Brunger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, N. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Acta Crystallogr. D. Biol. Crystallogr. **54**, 905 (1998)
113. A. Kulcsar, J. Saltiel, L. Zimanyi, J. Am. Chem. Soc. **123**, 3332 (2001)
114. T.A. Jackson, M. Lim, P.A. Anfinrud, Chem. Phys. **180**, 131 (1994)
115. F. Parak, Curr. Opin. Struct. Biol. **13**, 552 (2003)
116. S.J. Hagen, W.A. Eaton, J. Chem. Phys. **104**, 3395 (1996)
117. A. Bonvalet, A. Darmon, J.-C. Lambry, J.-L. Martin, P. Audebert, Opt. Lett. **31**, 2753 (2006)
118. J. Faure, Y. Glinec, A. Pukhov, S. Kiselev, S. Gordienko, E. Lefebvre, J.P. Rousseau, F. Burgy, V. Malka, Nature **431**, 541 (2004)
119. C.G.R. Geddes, Cs. Toth, J. van Tilborg, E. Esarey, C.B. Schroeder, D. Bruhwiler, C. Nieter, J. Cary, W.P. Leemans, Nature **431**, 538 (2004)
120. A.M. Kondratenko, E.L. Saldin, Part. Accelerators **10**, 207 (1980)
121. R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, J. Hajdu, Nature **406**, 752 (2000)
122. R. Neutze, G. Huldt, J. Hajdu, D. van der Spoel, Rad. Phys. Chem. **71**, 905 (2004)
123. A.A. Zholents, W. Fawley, Phys. Rev. Lett. **92**, 224801-1 (2004)
124. U. Haupts, J. Tittor, D. Oesterhelt, Annu. Rev. Biophys. Biomol. Struct. **283**, 67 (1999)
125. J.K. Lanyi, H. Luecke, Curr. Opin. Struct. Biol. **11**, 415 (2001)
126. R. Rammelsberg, G. Huhn, M. Lübben, K. Gerwert, Biochemistry **37**, 5001 (1998)
127. C. Scharnagl, S.F. Fischer, Chem. Phys. **212**, 231 (1996)
128. K. Murata, Y. Fujii, N. Enomoto, M. Hata, T. Hshino, M. Tsuda, Biophys. J. **79**, 982 (2000)
129. T. Ursby, M. Weik, E. Fioravanti, M. Delarue, M. Goeldner, D. Bourgeois, Acta Crystallogr. D. Biol. Crystallogr. **58**, 607 (2002)