Available online at www.sciencedirect.com
SCIENCE DIRECT®
Radiation Physics and Chemistry 71 (2004) 905–916
Radiation Physics and Chemistry
www.elsevier.com/locate/radphyschem

Potential impact of an X-ray free electron laser on structural biology

Richard Neutze[a,*], Gösta Huldt[b], Janos Hajdu[b], David van der Spoel[b]

[a] Department of Chemistry and Bioscience, Chalmers University of Technology, P.O. Box 462, SE-405 30 Göteborg, Sweden
[b] Department of Cell and Molecular Biology, Uppsala University, BMC, Box 596, 75124 Uppsala, Sweden

Abstract

Recent developments in X-ray source technology provide new opportunities for both the rapid imaging of macromolecules in three dimensions, and the observation of short-lived structural intermediates in light-sensitive macromolecules. Pioneering time-resolved X-ray diffraction studies on macromolecules have laid the foundations for similar studies on increasingly complex macromolecular systems at developing X-ray sources. In addition, the increased peak brilliance expected from an X-ray free electron laser source opens the possibility of new imaging technologies for determining X-ray structures from systems which do not readily yield well-diffracting crystals. Here we review the potential impact of extreme brilliance free electron laser X-ray sources on structural biology. We both sketch the biological case for sub-picosecond X-ray studies of light-sensitive proteins and describe the new imaging possibilities which will emerge through ultrafast X-ray imaging. A central issue is that of X-ray-induced radiation damage on macromolecular samples, and the advantages in this context of working in the femtosecond regime, are discussed.

Keywords: Structural biology; Time resolved X-ray diffraction; X-ray free electron laser

1. Introduction

Synchrotron radiation has profoundly influenced the field of structural biology. Along with the development of cryo-techniques for freezing crystals (Garman, 2003) (Fig. 1) and the shift from X-ray film to image plate and image intensified CCD detectors, the rate of determination of X-ray protein structures has accelerated continuously over the last two decades, since synchrotron radiation use has become mainstream. There are 25,000 entries now approaching in the protein data bank (http://www.rcsb.org/pdb/), approximately 85% of which are X-ray structures, and the rate at which macromolecular structures are being determined is currently in excess of 10 new protein data bank entries per day. While this number is somewhat inflated by the presence of closely related mutant structures, complex structures, structures of homologous proteins from different sources, and catalytic intermediate structures, etc., an impressive acceleration of progress cannot be denied.

There are essentially two major trends for the future development of the field of structural biology. One is to increase throughput through the implementation of automated technology, and the expression "*structural genomics*" has become a catch phrase within the community. Structural genomics programmes (Gershon, 2000) aim to determine all unique protein structures from within a specific genome, and thereby build up a database of all protein folds (and structures) in nature. Integral to these programmes is the projected increase in access to synchrotron radiation that the structural biology community is predicted to enjoy over the next

*Corresponding author. Tel.: +46-31-773-3974; fax: +46-31-773-3910.
E-mail address: neutze@molbiotech.chalmers.se (R. Neutze).

0969-806X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.radphyschem.2004.04.121

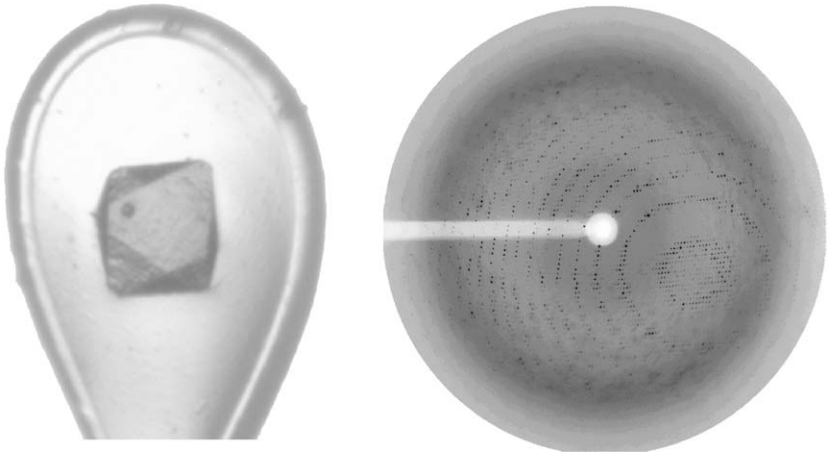

Fig. 1. A protein crystal mounted and frozen within a cryo-loop and the resulting X-ray diffraction pattern recorded at a synchrotron source. Images courtesy S. Törnroth.

decade. The use of automation technology, at all steps of the process from protein production to structural determination, is certainly simplifying the methods and accelerating the rate of progress. A consequence is that, more and more, protein structures are determined primarily to address biological and medical questions put forward by another community of scientists rather than to solve a structure in its own right.

The second major direction for structural biology is to focus on challenging scientific problems which are not easily solved with the existing technology, for which there are a plethora of examples. For example, the structural determination of membrane proteins represents a major challenge within the structural biology community (Lancaster, 2003; Cogdell et al., 2003; Fromme et al., 2003; Landau et al., 2003; MacKinnon, 2003; Bass et al., 2003; Abramson et al., 2003a,b; Chang, 2003; Toyoshima et al., 2003). This is because membrane proteins are located within cell membranes, and therefore contain both hydrophobic and hydrophilic surfaces, which are mutually repulsive and do not naturally form crystal contacts. Another example is the structural determination of large biological complexes such as the ribosome, which consists of a large number of individual components of both proteins and RNA. The remarkable milestone of X-ray structures of the ribosome subunits (Ban et al., 2000; Wimberly et al., 2000; Hanus et al., 2001) is one of the jewels in the crown of structural biology, and, as with most membrane protein structures, is indebted to the high X-ray brilliance of synchrotron radiation. Most membrane proteins and large biological complexes are available in only small quantities, and are difficult to assemble, purify, stabilise and crystallise. An important contribution in this regard has been the availability of micro-focus beamlines at synchrotron sources (e.g. http://www.esrf.fr/UsersAndScience/Experiments/SCMatter/ID13/) through which progress towards well-diffracting crystals can be monitored and optimised using crystals which historically were too small to be viable for structural determination. Nevertheless, many of these problems will, in all likelihood, prove intractable with traditional protein crystallisation techniques and may therefore demand novel approaches. It is here that the next generation of X-ray sources can be expected to make a profound impact (Patel, 2002).

Another growing challenge within structural biology is to go beyond the structural determination of the resting state of an enzyme, and structurally to characterise its functional mechanism throughout its catalytic cycle. Within this sub-discipline of ''time-resolved'' X-ray crystallography the aim is to visualise the details of how a protein functions in four-dimensions (three spatial plus time) and thereby gain insight into how nature has evolved to perform remarkably subtle chemical tasks. Although the numbers of examples where this approach has been successful are relatively few (see, e.g. Hajdu et al., 2000; Moffat, 1998; Wilmouth et al., 2001; Neutze et al., 2002), it is certain that an X-ray free electron laser will offer genuine advantages for such applications, potentially widening the field dramatically.

Three main features of the next generation of light sources will have to be exploited in order to create genuinely new possibilities for structural biology. The first is simply an evolutionary, albeit dramatic, increase in the number of photons available per X-ray pulse, which is projected to increase to around 10^{12} photons/pulse within a $\Delta E/E \sim 10^{-2}$ bandwidth (Winnik, 1995; Wiik, 1997), (c.w. $\sim 10^{10}$ photons/pulse in $\sim 10\%$ bandwidth at ID09 of the ESRF (Wulff et al., 1997)). This number is comparable with the X-ray exposure required to recover a diffraction image from very small

crystals ($<1000\,\mu m^3$ in volume), hence it will be possible to record interpretable X-ray diffraction data from small crystals in a single shot. As such it is a simple matter of extrapolation from current synchrotron-based experiments to foresee what type of experiments, particularly within the area of time-resolved crystallography, may be possible with such a source. The second feature is the exceptionally high brilliance of this radiation, enabling the X-ray beam to be focused close to the diffraction limit. As such it appears that the X-ray spot focal diameters of 100 nm or less will be easily attained using a free electron laser beam without significantly sacrificing the total X-ray flux. This creates revolutionary new possibilities for "nano-crystallography" which are certainly not feasible at the existing X-ray sources. The final feature is the temporal structure of the planned free electron laser, which will provide X-ray pulses of approximately 100 fs in duration, an improvement (reduction) of three orders of magnitude over existing synchrotron sources. This time structure creates the possibility for both evolutionary experiments in terms of improved temporal resolution in time-resolved X-ray diffraction, and for extraordinary experiments that are not even conceivable at existing sources today.

In this article, we will review the potential impact and benefits of planned X-ray free electron laser facilities on time-resolved X-ray crystallography (Section 2) as well as outline proposals for novel methods of structural determination using ultra-intense X-ray pulses (Section 3). Both applications exploit the increased peak-brilliance of the next generation of light-sources over existing synchrotrons. Finally, in Section 4 we conclude by sketching a personal view of how the projected development of X-ray source technology may influence structural biology.

2. Potential impact on time-resolved X-ray crystallography

Protein functions are remarkably diverse. In order to achieve the full spectrum of biological catalysis within the cell, enzymes fold to form carefully designed scaffolds upon which an active site where chemical catalysis performed is constructed. Enzyme function, however, is seldom passive and the enzyme must coordinate the arrival and binding of substrates, the release of products, and the input of energy in both space and time. Time-resolved crystallography, as the name suggests, seeks to follow enzymatic reactions fully along the reaction coordinate in order to characterise structurally all details of an enzyme's function.

There are two distinct branches within time-resolved protein crystallography. The first branch is the sub-discipline of pump–probe crystallography, whereby a reaction is rapidly triggered at room temperature using a "pump" (most commonly laser light) and the structural changes within the protein are followed in time using an X-ray "probe" (Moffat, 1998). The most spectacular successes within this branch include 100 ps studies on the photo-dissociation of a carbon-monoxide molecule bound to the haeme group of myoglobin following nano-second (Srajer et al., 1996, 2001) and femto-second (Schotte et al., 2003) laser excitation. Related studies concern the negative phototaxis receptor photoactive yellow protein (Perman et al., 1998; Ren et al., 2001). Due to the need to collect X-ray diffraction data rapidly, Laue-diffraction (i.e. using a white beam X-ray probe) has been the only feasible method for such studies (Bourgeois et al., 1996). Exceptionally well-diffracting protein crystals have been required in these studies because of problems with high background and spatial overlaps using the Laue-diffraction technique. The first proof-of-principle experiments used a broad wavelength X-ray probe within $\Delta E/E \sim 100\%$, although it has since been demonstrated that the quality of X-ray diffraction data can be improved significantly using narrower bandwidth radiation (Bourgeois et al., 2000). The conception of multi-layered X-ray optics is soon likely to see the development of narrow-beam "pink-Laue" diffraction (i.e. $\Delta E/E \sim 1\%$), which may soon allow these specialist techniques to be applied to a broader set of biologically systems.

The second major branch of time-resolved crystallography is the use of physical techniques to control the rates of reactions within crystals (Hajdu et al., 2000). If a specific reaction can, for all practical purposes, be stopped entirely, then monochromatic X-ray diffraction data may be collected at low temperature, thereby avoiding the technical, and frequently overwhelming, challenges associated with white-beam data collection. The most widespread technique in this respect is coined "kinetic crystallography", since the enzyme's kinetics are controlled using physical parameters (e.g. lowering the crystal's temperature, or varying the crystal's pH), frequently followed by plunging the crystal into liquid nitrogen and thereby "freezing" the protein into a desired conformation. A correspondence between enzymatic reactions at room temperature and at low temperature can be drawn using single-crystal micro-spectrophotometry (Hadfield and Hajdu, 1993; Bourgeois et al., 2002), through which spectral properties of specific intermediates can be compared within crystals. For example, with light-driven proteins one can illuminate crystals at low temperature, allowing the reaction to evolve up to the point where the thermal energy needed to cross a specific energy barrier is unavailable. As such a high population of a specific reaction intermediate may be achieved and this method has been successful for studies on carbon-monoxide: myoglobin complexes (Schlichting et al.,

1994), photoactive yellow protein (Genick et al., 1998), bacteriorhodopsin (Neutze et al., 2002; Edman et al., 1999; Royant et al., 2000) (Figs. 2 and 3), and sensory rhodopsin II (Edman et al., 2002). Alternatively, one can start a reaction at room temperature and subsequently quench the reaction by freezing the crystal in liquid nitrogen. Examples of this technique include studies of serine protease catalysis in which the reaction was activated using a pH jump (Wilmouth et al., 2001), or studies of quinoprotein amine oxidase (Wilmot et al., 1999) and dethiobiotin synthetase (Kack et al., 1998), for which the reactions were activated using substrate soaking.

Fourth generation X-ray sources will have a significant impact on both of these sub-disciplines. In the first instance, the significantly superior X-ray flux within a narrow X-ray bandwidth will immediately make available a tremendous improvement in signal-to-noise ratios for pump–probe studies on those systems which have already proven accessible using synchrotron radiation (Srajer et al., 1996; Perman et al., 1998). This consequent improvement in resolution will enable the structural details of light-driven processes to be visualised in more detail with considerably more confidence. Since the problem lies in the (subtle structural) details, these gains are perfectly consistent with the aims of such experiment and represent evolutionary improvements with the next generation of light sources. An interesting technical point is it is likely that, for very well-diffracting macromolecular systems, it will prove advantageous to use the "background" undulator radiation rather than the "laser-line" in such studies, since the former will provide a broader accessible bandwidth. With appoximately 10^{12} photons per pulse predicted within both the laser line and the undulator pedestal, there shall be tremendous flexibility for the scientist to optimise experimental conditions.

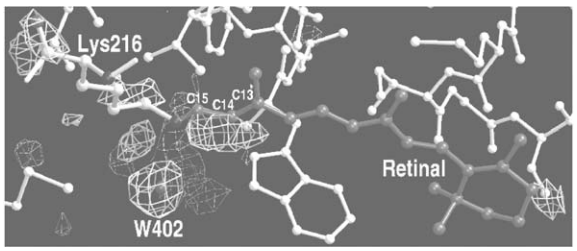

Fig. 2. Difference electron density resulting from structural changes associated with an early intermediate in the photocycle of bacteriorhodopsin. Negative electron density changes are lightly shaded, and positive electron density changes are darker grey. This figure is reproduced with permission from Ref. Edman et al. (1999).

The second spectacular gain of an X-ray free electron laser (XFEL) will be the reduction of the available X-ray pulse duration by three orders of magnitude over synchrotron radiation. Since the typical period of atomic vibrations within macromolecular systems is from 100 to 1000 fs, the availability of an X-ray probe with 100 fs duration will create the revolutionary possibility for visualising, in four dimensions, many basic photochemical phenomena. For example, an important class of biological systems are retinal proteins. This family of integral membrane proteins includes the visual rhodopsins, which are the primary photoreceptor for vision in animals (Palczewski et al., 2000), the sensory rhodopsins,

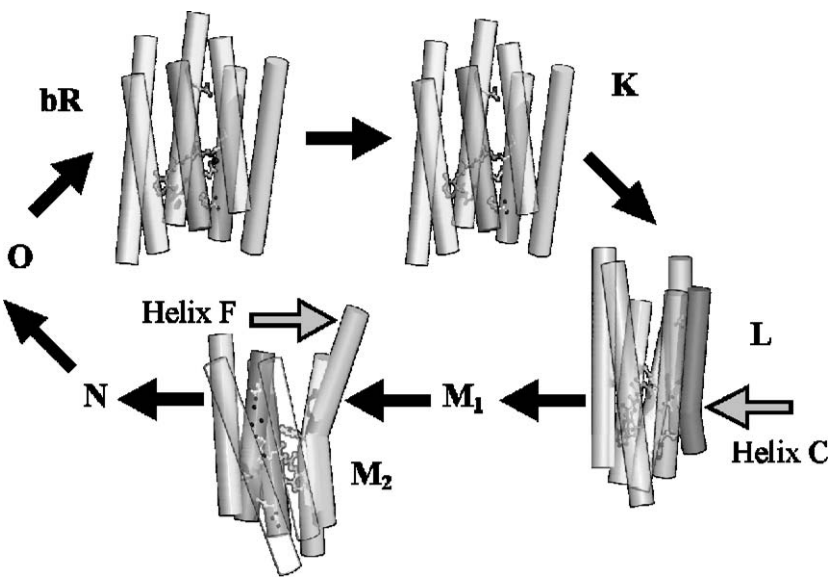

Fig. 3. Overview of the structural changes in the photocycle of bacteriorhodopsin. This figure is reproduced with permission from Ref. Neutze et al. (2002).

which govern negative or positive phototaxis in their host bacteria (Landau et al., 2003), and the transport rhodopsins (Landau et al., 2003), which harvest light so as to pump ions across cell membranes. Low-temperature studies on bacteriorhodopsin have shown that a water molecule hydrogen-bound to the retinal becomes disordered by light early in the photocycle (Edman et al., 1999) (Fig. 2). Due to the high X-ray flux and short pulse duration of the predicted XFEL beam, it should be feasible to follow retinal isomerisation from an all-*trans* conformation to its 13-*cis* conformation and clock the time-scale with which this water molecule becomes disordered.

Another exciting example falls within the field of photosynthesis, whereby photons are first captured by an integral membrane reaction centre complex, whereupon their energy is utilised by other processes within the cell. Ultrafast stimulated Raman scattering studies on bacterial reaction centres have implicated persistent large-scale coherent motions, with a period of $\sim$1 ps, following rapid photoexcitation (Vos et al., 1993). If the interpretation of these spectral results is correct, then an XFEL should make it possible to characterise these motions fully. Certainly, time-resolved X-ray diffraction studies with 100 fs temporal resolution will be performed on many macromolecular systems. This will provide an exciting and creative direction for future XFEL users, and structural results will shed new light on the fastest biological processes in nature.

Related to these studies are the emerging possibilities for time-resolved structural probes of light-dependent macromolecules without using X-ray difffraction. In particular two proof-of-principle demonstrations have recently been published for solution phase X-ray scattering (Neutze et al., 2001) and X-ray absorption spectroscopy (Saes et al., 2003) on simple photochemical systems in solution. A common theme to both experiments was that the structural information recovered lay in the local-vicinity of (relatively) heavy-atoms dissolved within a lighter solvent matrix. The increased X-ray flux available from an XFEL should make it possible to recover with confidence smaller signals over a higher background, and hence improve the general applicability of these techniques. Significant improvements in photon statistics will create the possibility for these techniques to also be applied to macromolecular systems. Since enzymes contain heavy atoms near the active site, or since they may be introduced using chemical means, then information on structural changes in the immediate vicinity of these atoms could prove invaluable in piecing together the structural details of their functional mechanism.

A third important, yet surprisingly under-sung, application of an XFEL within the discipline of time-resolved crystallography is to structurally characterise slower enzymatic processes through time-resolved X-ray diffraction. In particular, the proposed pulse structure at the TESLA facility (Wiik, 1997) foresees thousands of 100 fs X-ray pulses within microseconds long pulse trains separated by nano-second intervals, each pulse with sufficient X-ray photons to generate an interpretable X-ray diffraction image from very small crystals. It is not difficult to imagine triggering a reaction within a crystal by a variety of means, and designing the pulse-structure to match your chosen time window. Furthermore, the X-ray beam profiles may be tuned from bunch-to-bunch, and this provides an opportunity to design the experiment so as to optimise the X-ray diffraction data completeness. For instance, one might arrange a series of narrow bandwidth X-ray pulses which do not overlap spectrally, such that a quasi-Laue diffraction experiment may be built up of composite parts, thereby incorporating the primary benefit of Laue diffraction (i.e. the ability to sample many diffraction spots in a single experiment (Bourgeois et al., 1996)) while avoiding many of the pit-falls of the technique (e.g. spatial and spectral overlaps, and prohibitive restrictions on the degree of mosaic spread of the crystals). It may also prove it possible to overcome some of the technical challenges on the slower (milliseconds to seconds) time-scales using a rapidly spinning crystal in combination with rapid-readout 2D X-ray detectors, such as the developing pixel detectors (Sellin et al., 2001; Ayzenshtat et al., 2002, 2003).

These possibilities are particularly exciting since the vast majority of enzymes do not use light as their reaction trigger, and enzyme catalysed reactions invariably involve structural changes on the microseconds to milliseconds regime, with the femtosecond to picosecond regime being something of a special case. The methods available for reaction triggering (Hajdu et al., 2000) could involve substrate soaks, caged compounds (i.e. compounds containing a reactive species, such as ATP, which are blocked until photoactivated), pH jumps, temperature jumps, etc. Since it is possible to record all diffraction data from a single multi-bunch pass, it will also become possible to study non-reversible processes or systems involving such large-scale structural changes within their catalytic cycle. Large-scale movements, in particular, are frequently incompatible with the crystal lattice and the desired movements themselves tend to destroy the crystal contacts. A well-known example of this problem occurs in the later half of the photocycle of bacteriorhodopsin, for which the only successful experiment to structurally characterise the late M-intermediate (Fig. 3) relied upon crystallising a triple mutant analogue for this intermediate (Subramaniam et al., 1993), rather than directly observing the light-driven conformational change within a crystal. From an aesthetic viewpoint, one particularly appealing aspect of an application of an XFEL to time-resolve structural changes on the microsecond to millisecond time-scale, is

that the field would come full circle, since for optimal results both kinetic and time-resolved crystallography techniques would need to be incorporated into the design of each experiment targeting biologically relevant structural questions.

3. Potential impact on structural problems not accessible through crystallisation

Despite the impressive and accelerating rates at which new protein structures are being determined, there is no cause for complacency within structural biology. Many structural problems take years of painstaking work to reach to conclusions, as the proteins or protein:DNA or RNA complexes must be over-expressed or purified from natural sources, characterised biophysically, purified, stabilised, and thereupon crystallisation experiments may begin. Whereas crystals are frequently recovered even for difficult problems, in the overwhelming majority of crystallisation experiments on challenging systems, it takes thousands of crystallisation experiments to optimise crystals to the point where they are suitable for structural determination. A few notorious cases include bacteriorhodopsin, where the first crystals were published in 1980 (Michel and Oesterhelt, 1980) yet it was another 16 years before the most successful crystal form emerged (Landau and Rosenbusch, 1996); LacY, where cystine-mutant cross-linking studies were performed on every residue within the protein, and a "cross-linking" model was proposed, well before diffracting crystals were recorded (Abramson et al., 2003a, b); and the ribosome, where the first crystals were published in 1981 (Appelt et al., 1981; Yonath et al., 1984) yet an X-ray structure did not emerge until almost two decades later (Ban et al., 2000; Wimberly et al., 2000; Harms et al., 2001).

As the field of structural biology moves into the post Human-genome (or "post-genomics") era, emphasis is moving away from the study of individual protein structures towards understanding the role of protein (or DNA and RNA) components within larger complexes, and their interactions with other components within the cell at the structural level. Protein:protein interactions, signalling pathways, and regulation pathways are just a few of the challenges for which there currently is only a limited understanding at a structural level. Since it is frequently difficult, or even impossible, to recover large quantities of correctly assembled stable biological complexes, a prerequisite even to begin crystallisation experiments, new approaches will be needed to address many of these challenges at a structural level.

An obvious evolutionary development from the current success enjoyed by micro-focus beamlines at third generation synchrotron sources, is to construct nano-focus beamlines at an XFEL. In practice, along the path towards recovering micro-crystals (volumes of $1000\,\mu m^3$ or more) or larger crystals, one frequently follows leads with thousands of nano-crystals (i.e. volumes of $1\,\mu m^3$ or less) or "crystal-like" material, which is still visible under a microscope. However, as one reduces the focal spot of an XFEL to the sub-micron level, one has to become increasingly concerned with issues of radiation damage, even at cryogenic-temperatures. Furthermore, if one wishes to record useful diffraction data from crystals of $1\,\mu m^3$ or smaller, rather than $1000\,\mu m^3$ or larger, then the data-collection strategy must necessarily change. Primarily, it is unlikely that such a small crystal will survive more than a single exposure from the full XFEL beam. Consequently, rather than build up a complete X-ray diffraction data set by rotating a single crystal within the X-ray beam and collecting a sequence of diffraction images, one will be forced to scale together individual diffraction images from hundreds of different nano-crystals (more or less randomly aligned) in order to build up complete diffraction data. It is yet to be proven that one can effectively scale together such data, but it is reasonable to expect that this computational problem will be solved as nano-focus XFEL sources become reality.

The concept of nano-focus single-shot X-ray beams raises the issue of radiation damage to biological samples. A new possibility which is specific to an XFEL is to try to understand the extent to which the very short-pulse lengths can be exploited to push the radiation damage barrier beyond very real limits at synchrotron sources (Garman, 2003), i.e. can one rapidly image an exploding sample before the effects of X-ray damage destroy all useful structural information. To investigate this problem we performed a series of molecular dynamics simulations on a single particle of lysozyme exposed to a rapid, extremely intense X-ray beam (Neutze et al., 2000). We modified the molecular dynamics package GROMACS (Lindahl et al., 2001) to include: heating by the X-ray beam whereby the recoil momentum was conserved; bond breaking, through the use of Morse rather than harmonic potentials; sample ionization due to inelastic scattering, photoelectric ionization, and the emission of Auger electrons from the atoms of the sample. These simulations kept an inventory of the number of electrons removed from each atom through these processes, since their ionization states strongly affected their dynamics, as well as their X-ray scattering cross-sections.

Fig. 4 shows how the number of X-ray-induced ionization events increased in time for a sample exposed to a 12 keV X-ray pulse containing 3×10^{12} photons focused through a 100 nm diameter focal spot. In the first simulation the FWHM of the X-ray pulse is assumed to be 10 fs, whereas the second is set at 50 fs. While the number of primary ionization events (photoelectric and inelastic) is the same for both simulations,

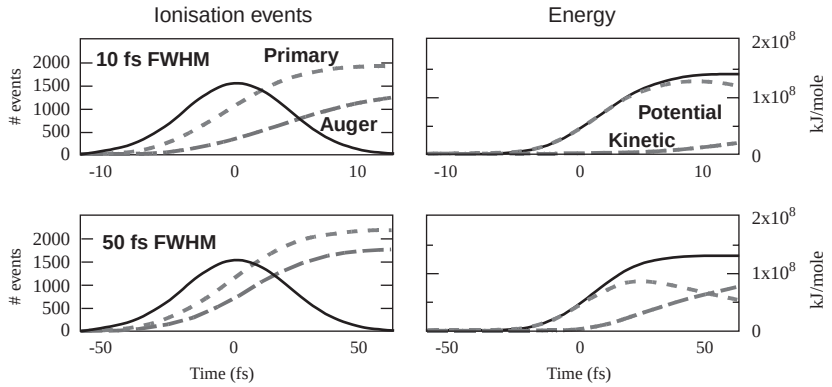

Fig. 4. Ionisation events and energy changes during the exposure of a single molecule of lysozyme to an intense X-ray pulse containing 3×10^{12} X-ray photons focused through a 0.1 μm diameter focal spot. This figure is reproduced from Ref. Neutze et al. (2000).

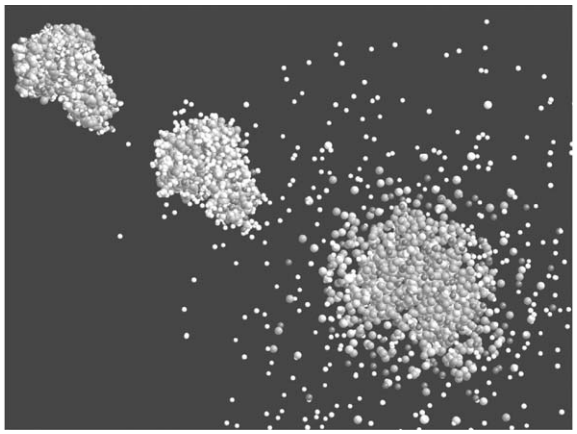

Fig. 5. Visual representation of the explosion of a single molecule of lysozyme when exposed to an extremely intense femtosecond duration X-ray pulse. This figure is reproduced with permission from Ref. Neutze et al. (2000).

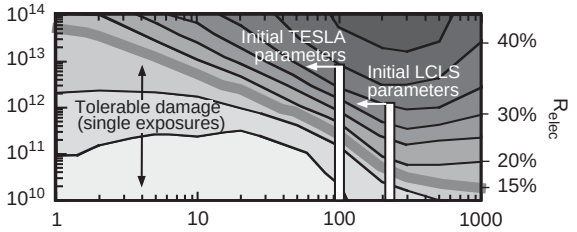

Fig. 6. Landscape of damage tolerance. This figure plots the effective R-factor values for a range of pulse durations and X-ray pulse intensity (given in units of number of photons focused through a 0.1 μm focal spot). Values of R below 15% are suggested to be acceptable. This figure is reproduced with permission from Ref. Neutze et al. (2000).

the natural time scale for Auger emission (approximately 10 fs for the light atoms) leads to there being more secondary ionization events in the second simulations. The creation of a large number of charges in close proximity provides an enormous driving force for a Coulomb explosion, and this is delayed slightly by inertial effects. Most dramatically, for the shorter simulation there has been insufficient time for the efficient conversion of electrostatic potential energy into kinetic energy, and the sample has only begun to explode at the completion of the X-ray pulse. In contrast, by the end of the 50 fs X-ray pulse, half of the potential energy is converted to kinetic energy, and the X-ray-induced explosion of the sample is well underway. A visual representation of this fact is given in Fig. 5, where a sequence of pictures show the structure of the lysozyme molecule at the beginning, near the centre, and towards the end of the 50 fs X-ray pulse.

To quantify the results from these simulations, X-ray scattering distributions from the single lysozyme molecule were calculated throughout the simulation and were averaged over the simulation ($I(t)$), and were compared against a hypothetical "ideal case" where the effects of radiation damage are turned-off "in silico" (I_0). An agreement factor defined as

$$R = |\sqrt{I(t)} - \sqrt{I_0}|/\sqrt{I_0}$$

which is analogous to the R-factor used as a representation of the quality of the crystallographic data and model. Fig. 6 provides a graphical representation of this analysis, plotting R as a two-dimensional contour versus the pulse-duration and the total X-ray flux. Somewhat arbitrarily an R-factor of 15% is chosen as a boundary marking the borderline of tolerable damage, and it is clearly seen that in order to sustain 3×10^{12} photons through a 100 nm focal spot, the X-ray pulse should be of the order of 10 fs or less. For longer X-ray pulse it is necessary to use a lower X-ray flux through the sample. This result quantifies what one would intuitively expect from the results of the simulations (Fig. 4) and their structural representation (Fig. 5).

Working from the assumption that $R<15\%$ is a reasonable criterion for judging which X-ray scattering data could be useful, it is possible to calculate the X-ray scattering power for a number of samples, and from that determine what resolution interpretable data may be recovered (Fig. 7). It is evident that even very small well-ordered crystals, for example $10\times10\times10$ of lysozyme ($\sim0.1\,\mu m^3$) which is too small to be easily seen under an optical microscope, can still produce X-ray diffraction to reasonable resolution when using pulses with a duration consistent with the planned XFEL facilities. This could prove to be a significant advantage over existing experiments at synchrotron sources. To recover useful X-ray diffraction data to good resolution from even smaller non-ordered clusters appears to be challenging unless the X-ray pulse duration can be shortened. Nevertheless, there should be sufficient statistics to recover some useful structural information from well ordered yet non-crystalline systems, such as an icosahedral virus, even with X-ray pulses longer than 100 fs. This situation is rather similar to the current state of the art in electron microscopy, where low-resolution electron microscopy structures are now rapidly recovered from a large-number of images of single virus particles (Bottcher et al., 1997). Likewise, it is possible to recover low-resolution structural information from non-symmetric samples (Mueller et al., 2000) using electron microscopy as long as the sample is sufficiently large, and it appears that ultra-fast X-ray scattering approaches may provide complementary information. In all likelihood, the best structural results will emerge through electron microscopy studies to recover a low-resolution structural model, followed by X-ray scattering studies aiming to extend the resolution of the model. To a structural biologist, the difference between data to 6 Å resolution and to 3 Å resolution is effectively entering a new regime, since it is only at higher resolution that the details of structural mechanism become tangible.

Since these simulations were published, there have been a number of suggestions as to how to improve the physical model for the Coulomb explosion of the protein. Most attention has been paid to the influence of secondary electrons which lead to additional ionization events within the sample (Ziaja et al., 2001, 2002). Recent calculations, which explicitly incorporated electron trajectories and their interactions with other atoms within the simulated sample (Jurek et al., 2004), have suggested that the rate of onset of radiation damage effects may be somewhat enhanced over the earlier model (Neutze et al., 2000). The availability of the first experimental results from the Tesla Test Facility Vis-UV FEL (Wabnitz et al., 2002) is very promising in this regard. It should soon be possible to measure ionization and protein fragmentation effects due to interaction with a soft X-ray FEL beam. While a different regime, these experiments will help guide improvements in computational models.

Progress has also recently been made on the treatment of data resulting from an X-ray scattering experiment without crystals. In electron microscopy studies, the cross-correlation between images is widely used to sort individual images into specific image classes, such that numerical averaging techniques can be applied. Due to the enhanced signal-to-noise ratio recovered by sorting and averaging, a common-line projection theorem can

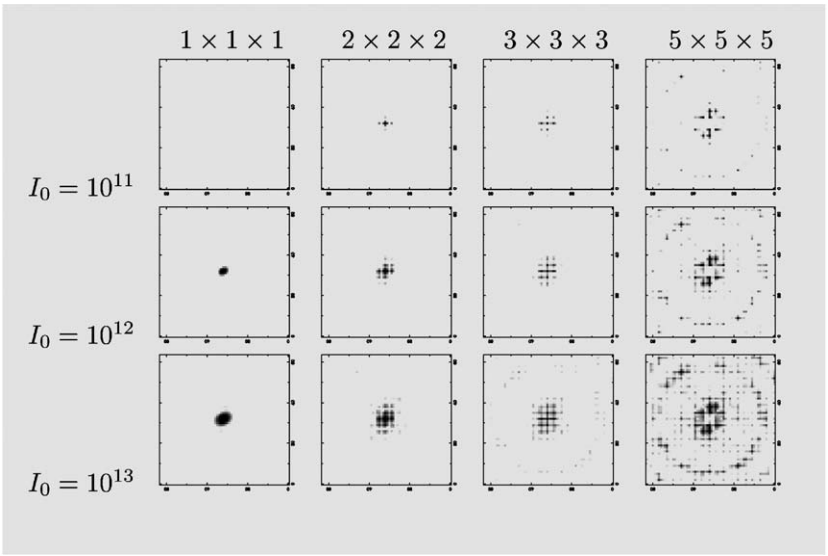

Fig. 7. X-ray scattering from small clusters of lysozyme nano-crystals as a function of the total number of 12 keV photon (I_0) passing through a 0.1 μm focal spot.

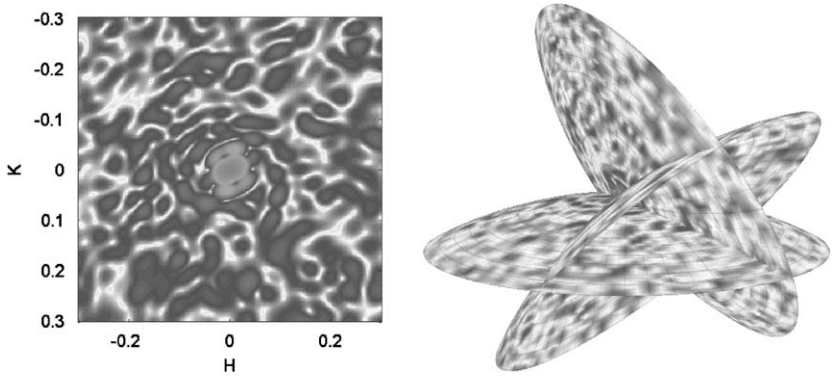

Fig. 8. Molecular transform of a single macromolecule of lysozyme, and its representation in reciprocal space. Any two Ewald-spheres, upon which the X-ray scattering patterns are projected, share a common line of intersection in reciprocal space. This can be used to classify different images and recover the orientation of a randomly oriented sample. This figure is reproduced with permission from Ref. Huldt et al. (2003).

then be used to find the mutual orientation of the averaged images. This theorem is based upon the idea that any two projections through a three-dimensional object have at least one line in common (Van Heel, 1987). Huldt et al. have extended this idea into the X-ray domain (Huldt et al., 2003) by noting that any two X-ray scattering images will share at least one curve in common where the Ewald spheres intersect in reciprocal space (Fig. 8). From this observation, after sorting-and-averaging based upon the cross-correlations between individual X-ray scattering images, it will be possible to orientate each of the averaged image classes (deriving from randomly orientated biological samples) relative to each other such that the X-ray scattering intensities can be accurately mapped throughout the Ewald sphere. This averaging approach will significantly improve the resulting photon statistics, and will thereby overcome many of the difficulties associated with single-shot imaging. As such the resolution which one can achieve in the final model will be limited by how accurately individual images can be sorted into image classes and oriented relative to each other, and the proposed alignment protocol (Huldt et al., 2003) appears to be very powerful in this respect.

Theoretical work has also been performed on generating electron density from images recorded from non-crystalline samples. Miao et al. (2001) have shown that, when noise is added to X-ray scattering intensities (sampled along a regular three-dimensional grid) from a single Ribisco molecule, it was possible to invert these intensities and recover electron density. In particular, due to the continuous sampling of reciprocal space, rather that the discreet sampling of reciprocal space when recording diffraction from a crystalline lattice, there was sufficient information available to determine the phases directly. The algorithm used information regarding the boundary of the molecule, with a key concept being that over-sampling allowed the electron density outside of the molecule to be calculated, which should be zero, providing a powerful constraint on the experimental phases. A noteworthy theoretical issue remaining to be addressed is how best to map experimental data onto a regular grid in reciprocal space when starting from randomly oriented raw images. Sorting, averaging, orientation and interpolation methods will produce data with quite unusual properties, which may influence the convergence and noise sensitivity of any given iterative phasing algorithm. Nevertheless, the numerical experiment of Miao et al. (2001) has since been experimentally demonstrated in a different arena using a frozen *Escherichia coli* bacteria as a sample and coherent synchrotron generated X-rays as the source (Miao et al., 2003). While several technical challenges still remain in moving from theoretically generated data to genuine experimental data, it appears that the key ingredients concerning how to process experimental data and recover electron density have already been well formulated.

4. Conclusion

Structural biology has enjoyed a period of extended success and growth during the last decade. As many of the traditional challenges from crystal growth to model building are becoming automated, effort is increasingly turning to more challenging problems. Two of the future challenges for structural biology are to understand the structural details of reaction mechanisms, and to study the structural biology of large complexes, which are difficult to produce and purify in quantities suitable for traditional X-ray crystallography approaches.

Time-resolved X-ray crystallography is a powerful tool when probing beyond the resting state structure of

an enzyme. While there have been several milestones in this field using synchrotron generated X-rays (Hajdu et al., 2000; Moffat, 1998), the projected properties of an XFEL should make it possible to study both light-driven systems with improved temporal and spatial resolution, and to study the catalytic reactions of non-reversible systems. The latter point is particularly poignant since most biological reactions do not follow fully reversible photo cycles, and this new technology may open up a large number of biological problems to study which were not accessible using synchrotron radiation. This is particularly true for systems undergoing large-scale movements, which can disrupt the crystal lattice.

X-ray scattering from non-crystalline samples promises a potentially revolutionary avenue for many of the most challenging problems in structural biology (Patel, 2002; Neutze et al., 2000). It is clear that there are benefits to be had from using ultra-fast X-ray exposures, but the technique will push both the technical limits of the XFEL source, and the physical limits of the biological sample. As such many significant challenges remain and this new regime is only beginning to be understood, yet the structural biology community should look forward to a period of excitement as novel X-ray scattering approaches are realized experimentally, and unanticipated opportunities arise and are exploited. Ultimately, as these new X-ray facilities emerge (Patel, 2002) they will open new windows providing unique structural insight into the most challenging biological problems of the day.

Acknowledgements

We wish to thank Jan Davidsson, Jörgen Larsson, Remco Wouts and Michael Wulff for many years of collaboration. We are guilty of borrowing from the phrase "evolutionary, revolutionary, extraordinary" from several excellent lectures by Keith Hodgson. We acknowledge financial support from the Swedish Research Council (VR), the Swedish Strategic Research Foundation (SSF), SWEGENE, and the European Commission Improving Human Potential programme.

References

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Iwata, S., Kaback, H.R., 2003a. The lactose permease of *Escherichia coli*: overall structure, the sugar-binding site and the alternating access model for transport. FEBS Lett. 555, 96–101.

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., Iwata, S., 2003b. Structure and mechanism of the lactose permease of *Escherichia coli*. Science 301, 610–615.

Appelt, K., Dijk, J., Reinhardt, R., Sanhuesa, S., White, S.W., Wilson, K.S., Yonath, A., 1981. The crystallization of ribosomal proteins from the 50 S subunit of the *Escherichia coli* and *Bacillus stearothermophilus* ribosome. J. Biol. Chem. 256, 11787–11790.

Ayzenshtat, G.I., Bimatov, M.V., Tolbanov, O.P., Vorobiev, A.P., 2002. Charge collection in X-ray pixel detectors based on semi-insulating GaAs doped with Cr. Nucl. Instrum. Methods Phys. Res. Sect. A—Accelerators Spectrometers Detectors Assoc. Equip. 494, 210–213.

Ayzenshtat, G.I., Bimatov, M.V., Tolbanov, O.P., Vorobiev, A.P., 2003. Charge collection in X-ray pixel detectors based on SI-GaAs doped with Cr. Nucl. Instrum. Methods Phys. Res. Sect. A—Accelerators Spectrometers Detectors Assoc. Equip. 509, 52–55.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 289, 905–920.

Bass, R.B., Locher, K.P., Borths, E., Poon, Y., Strop, P., Lee, A., Rees, D.C., 2003. The structures of BtuCD and MscS and their implications for transporter and channel function. FEBS Lett. 555, 111–115.

Bottcher, B., Wynne, S.A., Crowther, R.A., 1997. Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. Nature 386, 88–91.

Bourgeois, D., Ursby, T., Wulff, M., Pradervand, C., Legrand, A., Schildkamp, W., Laboure, S., Srajer, V., Teng, T.Y., Roth, M., Moffat, K., 1996. Feasibility and realization of single-pulse Laue diffraction on macromolecular crystals at ESRF. J. Synchrotron Radiat. 3, 65–74.

Bourgeois, D., Wagner, U., Wulff, M., 2000. Towards automated Laue data processing: application to the choice of optimal X-ray spectrum. Acta Crystallogr. D Bio. Crystallogr. 56, 973–985.

Bourgeois, D., Vemede, X., Adam, V., Fioravanti, E., Ursby, T., 2002. A micro spectrophotometer for UV-visible absorption and fluorescence studies of protein crystals. J. Appl. Crystallogr. 35, 319–326.

Chang, G., 2003. Multidrug resistance ABC transporters. FEBS Lett. 555, 102–105.

Cogdell, R.J., Isaacs, N.W., Freer, A.A., Howard, T.D., Gardiner, A.T., Prince, S.M., Papiz, M.Z., 2003. The structural basis of light-harvesting in purple bacteria. FEBS Lett. 555, 35–39.

Edman, K., Nollert, P., Royant, A., Belrhali, H., Pebay-Peyroula, E., Hajdu, J., Neutze, R., Landau, E.M., 1999. High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. Nature 401, 822–826.

Edman, K., Royant, A., Nollert, P., Maxwell, C.A., Pebay-Peyroula, E., Navarro, J., Neutze, R., Landau, E.M., 2002. Early structural rearrangements in the photocycle of an integral membrane sensory receptor. Structure (Camb) 10, 473–482.

Fromme, P., Melkozernov, A., Jordan, P., Krauss, N., 2003. Structure and function of photo system I: interaction with its soluble electron carriers and external antenna systems. FEBS Lett. 555, 40–44.

Garman, E., 2003. 'Cool' crystals: macromolecular cryocrystallography and radiation damage. Curr. Opin. Struct. Biol. 13, 545–551.

Genick, U.K., Soltis, S.M., Kuhn, P., Canestrelli, I.L., Getzoff, E.D., 1998. Structure at 0.85: a resolution of an early protein photocycle intermediate. Nature 392, 206–209.

Gershon, D., 2000. Structural genomics—from cottage industry to industrial revolution. Nature 408, 273–274.

Hadfield, A., Hajdu, J., 1993. A fast and portable microspectrophotometer for protein crystallography. J. Appl. Crystallogr. 26, 839–842.

Hajdu, J., Neutze, R., Sjogren, T., Edman, K., Szoke, A., Wilmouth, R., Wilmot, C.M., 2000. Analyzing protein functions in four dimensions. Nat. Struct. Biol. 7, 1006–1012.

Hanus, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., Yonath, A., 2001. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell I 07, 679–688.

Huldt, G., Szoke, A., Hajdu, J., 2003. Diffraction imaging of single particles and biomolecules. J. Struct. Biol. 144, 219–227.

Jurek, Z., Faigel, G., Tegze, M., 2004. Dynamics in a cluster under the influence of intense femtosecond hard X-ray pulses. Eur. Phys. J. D 29, 217–229.

Kack, H., Gibson, K.J., Lindqvist, Y., Schneider, G., 1998. Snapshot of a phosphorylated substrate intermediate by kinetic crystallography. Proc. Natl. Acad. Sci. USA 95, 5495–5500.

Lancaster, C.R.D., 2003. Wolinella succinogenes quinol: fumarate reductase and its comparison to *E. coli* succinate: quinone reductase. FEBS Lett. 555, 21–28.

Landau, E.M., Rosenbusch, J.P., 1996. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. Proc. Natl. Acad. Sci. USA 93, 14532–14535.

Landau, E.M., Pebay-Peyroula, E., Neutze, R., 2003. Structural and mechanistic insight from high resolution structures of archaeal rhodopsins. FEBS Lett. 555, 51–56.

Lindahl, E., Hess, B., van der Spoel, D., 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J. Mol. Mod. 7, 306–317.

MacKinnon, R., 2003. Potassium channels. FEBSLett. 555, 62–65.

Miao, J., Hodgson, K.O., Sayre, D., 2001. An approach to three-dimensional structures of biomolecules by using single-molecule diffraction images. Proc. Natl. Acad. Sci. USA 98, 6641–6645.

Miao, J., Hodgson, K.O., Ishikawa, T., Larabell, C.A., LeGros, M.A., Nishino, Y., 2003. Imaging whole *Escherichia coli* bacteria by using single-particle X-ray diffraction. Proc. Natl. Acad. Sci. USA 100, 110–112.

Michel, H., Oesterhelt, D., 1980. Three-dimensional crystals of membrane proteins: bacteriorhodopsin. Proc. Natl. Acad. Sci. USA 77, 1283–1285.

Moffat, K., 1998. Ultrafast time-resolved crystallography. Nat. Struct. Biol. 5 (Suppl), 641–643.

Mueller, F., Sommer, I., Baranov, P., Matadeen, R., Stoldt, M., Wohnert, J., Gorlach, M., van Heel, M., Brimacombe, R., 2000. The 3D arrangement of the 23 S and 5 S rRNA in the *Escherichia coli* 50 S ribosomal subunit based on a cryo-electron microscopic reconstruction at 7.5 A resolution. J. Mol. Biol. 298, 35–59.

Neutze, R., Wouts, R., van der Spoel, D., Weckert, E., Hajdu, J., 2000. Potential for biomolecular imaging with femtosecond X-ray pulses. Nature 406, 752–757.

Neutze, R., Wouts, R., Techert, S., Davidsson, J., Kocsis, M., Kirrander, A., Schotte, F., Wulff, N., 2001. Visualizing photochemical dynamics in solution through picosecond X-ray scattering. Phys. Rev. Lett. 8719, art. no.-195508.

Neutze, R., Pebay-Peyroula, E., Edman, K., Royant, A., Navarro, J., Landau, E.M., 2002. Bacteriorhodopsin: a high-resolution structural view of vectorial proton transport. Biochim. Biophys. Acta 1565, 144–167.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M., Miyano, M., 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. Science 289, 739–745.

Patel, N., 2002. Shorter, brighter, better. Nature 415, 110–111.

Perman, B., Srajer, V., Ren, Z., Teng, T., Pradervand, C., Ursby, T., Bourgeois, D., Schotte, F., Wulff, M., Kort, R., Hellingwerf, K., Moffat, K., 1998. Energy transduction on the nanosecond time scale: early structural events in a xanthopsin photocycle. Science 279, 1946–1950.

Ren, Z., Perman, B., Srajer, V., Teng, T.Y., Pradervand, C., Bourgeois, D., Schotte, F., Ursby, T., Kort, R., Wulff, M., Moffat, K., 2001. A molecular movie at 1.8 A resolution displays the photocycle of photo active yellow protein, a eubacterial blue-light receptor, from nanoseconds to seconds. Biochemistry 40, 13788–13801.

Royant, A., Edman, K., Ursby, T., Pebay-Peyroula, E., Landau, E.M., Neutze, R., 2000. Helix deformation is coupled to vectorial proton transport in the photocycle of bacteriorhodopsin. Nature 406, 645–648.

Saes, M., Bressler, C., Abela, R., Grolimund, D., Johnson, S.L., Heimann, P.A., Chergui, M., 2003. Observing photochemical transients by ultrafast X-ray absorption spectroscopy. Phys. Rev. Lett. 90, 047403-1–047403-4.

Schlichting, I., Berendzen, J., Phillips Jr., G.N., Sweet, R.M., 1994. Crystal structure of photo lysed carbonmonoxymyoglobin. Nature 371, 808–812.

Schotte, F., Lim, M., Jackson, T.A., Smimov, A.V., Soman, J., Olson, J.S., Phillips Jr., G.N., Wulff, M., Anfinrud, P.A., 2003. Watching a protein as it functions with 150-ps time-resolved X-ray crystallography. Science 300, 1944–1947.

Sellin, P.J., Rossi, G., Renzi, M.J., Knights, A.P., Eikenberry, E.F., Tate, M.W., Barna, S.L., Wixted, R.L., Gruner, S.M., 2001. Performance of semi-insulating gallium arsenide X-ray pixel detectors with current-integrating readout. Nucl. Instrum. Methods Phys. Res. Sect. A—Accelerators Spectrometers Detectors Assoc. Equip. 460, 207–212.

Srajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M., Moffat, K., 1996. Photolysis of the carbon monoxide complex of myoglobin: nanosecond time-resolved crystallography. Science 274, 1726–1729.

Srajer, V., Ren, Z., Teng, T.Y., Schmidt, M., Ursby, T., Bourgeois, D., Pradervand, C., Schildkamp, W., Wulff, M., Moffat, K., 2001. Protein conformational relaxation and ligand migration in myoglobin: a nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction. Biochemistry 40, 13802–13815.

Subramaniam, S., Gerstein, M., Oesterhelt, D., Henderson, R., 1993. Electron diffraction analysis of structural changes in the photocycle of bacteriorhodopsin. EMBO J. 12, 1–8.

Toyoshima, C., Nomura, H., Sugita, Y., 2003. Structural basis of ion pumping by Ca2+-ATPase of sarcoplasmic reticulum. FEBS Lett. 555, 106–110.

Van Heel, M., 1987. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. Ultramicroscopy 21, 111–123.

Vos, M.H., Rappaport, F., Lambry, J.C., Breton, J., Martin, J.L., 1993. Visualization of coherent nuclear motion in a membrane-protein by femtosecond spectroscopy. Nature 363, 320–325.

Wabnitz, H., Bittner, L., de Castro, A.R., Dohrmann, R., Gurtler, P., Laarmann, T., Laasch, W., Schulz, J., Swiderski, A., von Haeften, K., Moller, T., Faatz, B., Fateev, A., Feldhaus, J., Gerth, C., Hahn, U., Saldin, E., Schneidmiller, E., Sytchev, K., Tiedtke, K., Treusch, R., Yurkov, M., 2002. Multiple ionization of atom clusters by intense soft X-rays from a free-electron laser. Nature 420, 482–485.

Wiik, B.H., 1997. The TESLA project: an accelerator facility for basic science. Nucl. Instrum. Methods 398, 1–17.

Wilmot, C.M., Hajdu, J., McPherson, M.J., Knowles, P.F., Phillips, S.E., 1999. Visualization of dioxygen bound to copper during enzyme catalysis. Science 286, 1724–1728.

Wilmouth, R.C., Edman, K., Neutze, R., Wright, P.A., Clifton, I.J., Schneider, T.R., Schofield, C.J., Hajdu, J., 2001. X-ray snapshots of serine protease catalysis reveal a tetrahedral intermediate. Nat. Struct. Biol. 8, 689–694.

Wimberly, B.T., Brodersen, D.E., Clemons Jr., W.M., Morgan-Warren, R.J., Carter, A.P., Vomhein, C., Hartsch, T., Ramakrishnan, V., 2000. Structure of the 30S ribosomal subunit. Nature 407, 327–339.

Winnik, H., 1995. The linac coherent light source (LCLS): a fourth-generation light source using the SLAC linac. J. Elec. Spec. Rei. Phenom. 75, 1–8.

Wulff, M., Schotte, F., Naylor, G., Bourgeois, D., Moffat, K., Mourou, G., 1997. Time-resolved structures of macromolecules at the ESRF: single-pulse Laue diffraction, stroboscopic data collection and femtosecond flash photolysis. Nucl. Instrum. Methods 398, 69–84.

Yonath, A., Bartunik, H.D., Bartels, K.S., Wittmann, H.G., 1984. Some X-ray diffraction patterns from single crystals of the large ribosomal subunit from *Bacillus stearothermophilus*. J. Mol. Biol. 177, 201–206.

Ziaja, B., van der Spoel, D., Szoke, A., Hajdu, J., 2001. Auger-electron cascades in diamond and amorphous carbon. Phys. Rev. Ii 6421, art. no.-214104.

Ziaja, B., Szoke, A., van der Spoel, D., Hajdu, J., 2002. Space-time evolution of electron cascades in diamond. Phys. Rev. B 66, 024116-1–024116-9.