Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams

Karol Nass, Lutz Foucar, Thomas R. M. Barends, Elisabeth Hartmann, Sabine Botha, Robert L. Shoeman, R. Bruce Doak, Roberto Alonso-Mori, Andrew Aquila, Saša Bajt, Anton Barty, Richard Bean, Kenneth R. Beyerlein, Maike Bublitz, Nikolaj Drachmann, Jonas Gregersen, H. Olof Jönsson, Wolfgang Kabsch, Stephan Kassemeyer, Jason E. Koglin, Michael Krumrey, Daniel Mattle, Marc Messerschmidt, Poul Nissen, Linda Reinhard, Oleg Sitsel, Dimosthenis Sokaras, Garth J. Williams, Stefan Hau-Riege, Nicusor Timneanu, Carl Caleman, Henry N. Chapman, Sébastien Boutet and Ilme Schlichting

Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams

Karol Nass,[a] Lutz Foucar,[a] Thomas R. M. Barends,[a] Elisabeth Hartmann,[a] Sabine Botha,[a] Robert L. Shoeman,[a] R. Bruce Doak,[a] Roberto Alonso-Mori,[b] Andrew Aquila,[c] Saša Bajt,[d] Anton Barty,[e] Richard Bean,[e] Kenneth R. Beyerlein,[e] Maike Bublitz,[f] Nikolaj Drachmann,[f] Jonas Gregersen,[f] H. Olof Jönsson,[g] Wolfgang Kabsch,[a] Stephan Kassemeyer,[a] Jason E. Koglin,[b] Michael Krumrey,[h] Daniel Mattle,[f] Marc Messerschmidt,[b] Poul Nissen,[f] Linda Reinhard,[f] Oleg Sitsel,[f] Dimosthenis Sokaras,[b] Garth J. Williams,[b] Stefan Hau-Riege,[i] Nicusor Timneanu,[g] Carl Caleman,[e,g] Henry N. Chapman,[e,j] Sébastien Boutet[b] and Ilme Schlichting[a]*

[a]Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Jahnstrasse 29, D-69120 Heidelberg, Germany, [b]SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA, [c]European XFEL GmbH, Albert-Einstein-Ring 19, 22761 Hamburg, Germany, [d]Photon Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany, [e]Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany, [f]Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, Aarhus 8000, Denmark, [g]Department of Physics and Astronomy, Uppsala University, Box 516, Uppsala 75120, Sweden, [h]Physikalisch-Technische Bundesanstalt (PTB), Abbestrasse 2-12, 10587 Berlin, Germany, [i]Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA, and [j]Department of Physics, University of Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany. *E-mail: ilme.schlichting@mpimf-heidelberg.mpg.de

Proteins that contain metal cofactors are expected to be highly radiation sensitive since the degree of X-ray absorption correlates with the presence of high-atomic-number elements and X-ray energy. To explore the effects of local damage in serial femtosecond crystallography (SFX), *Clostridium ferredoxin* was used as a model system. The protein contains two [4Fe–4S] clusters that serve as sensitive probes for radiation-induced electronic and structural changes. High-dose room-temperature SFX datasets were collected at the Linac Coherent Light Source of ferredoxin microcrystals. Difference electron density maps calculated from high-dose SFX and synchrotron data show peaks at the iron positions of the clusters, indicative of decrease of atomic scattering factors due to ionization. The electron density of the two [4Fe–4S] clusters differs in the FEL data, but not in the synchrotron data. Since the clusters differ in their detailed architecture, this observation is suggestive of an influence of the molecular bonding and geometry on the atomic displacement dynamics following initial photoionization. The experiments are complemented by plasma code calculations.

1. Introduction

Radiation damage is one of the primary limiting factors in obtaining structural information from small and/or radiation-sensitive crystals. Ultimately, it limits the maximal resolution of the data (Holton & Frankel, 2010) and can complicate the interpretation of the electron density of redox sensitive cofactors (Kühnel et al., 2007; Yano et al., 2005). Recent developments at third-generation synchrotrons, such as better focusing optics and the use of pink beams resulting in higher flux densities, as well as faster detectors have enabled rapid data collection from micrometre-sized crystals, but have not solved the intrinsic problem of radiation damage (Owen et al., 2014). Although cryocooling techniques reduce the effects of radiation damage for a given dose, the dose limit of 30 MGy (Owen et al., 2006) can still be insufficient, particularly for small crystals. At room temperature, the dose limit is even lower (0.2 MGy) due to the faster diffusion of radicals, and

doi:10.1107/S1600577515002349 **225**

damage becomes dose-rate-dependent (Owen *et al.*, 2012). Importantly, photoreduction of metal centres already occurs at significantly lower doses, complicating if not preventing the collection of chemically meaningful diffraction data using synchrotron sources (Yano *et al.*, 2005).

The femtosecond-duration X-ray pulses produced by X-ray free-electron lasers (FELs) present a novel approach to structure determination by allowing the 'diffraction-before-destruction' approach in which a diffraction pattern is obtained in the short time-frame before the inevitable dis-integration of the sample has progressed too far. This provides a way to circumvent the problem of radiation damage, by outrunning the slower processes (Neutze *et al.*, 2000; Chapman *et al.*, 2011, 2014). The approach seems to work for 'moderate' doses of 30–150 MGy per crystal (Boutet *et al.*, 2012; Kern *et al.*, 2013; Liu *et al.*, 2013); however, clear signs of global radiation damage are observed in the case of higher doses and longer pulses (Lomb *et al.*, 2011; Barty *et al.*, 2012). In fact, our previous experiments at the Linac Coherent Light Source (LCLS) at SLAC National Accelerator Laboratory using soft X-rays showed both a dose- and a dose-rate-dependent resolution deterioration of the diffraction data of lysozyme nanocrystals for doses around 3–6 GGy (Lomb *et al.*, 2011). Both the lysozyme study and experiments using photosystem I (Barty *et al.*, 2012) showed a decrease of the high-resolution diffraction signal with increasing pulse lengths. At these high dose rates, the sample is rapidly ionized. The Coulomb repulsion of the ions and the rapid rise in electron tempera-ture lead to displacements of atoms and ions and this thermal disorder results in loss of crystalline order during the pulse, which ultimately terminates diffraction (Lomb *et al.*, 2011; Barty *et al.*, 2012).

In a first approximation to describing the processes under-lying ultrafast damage, Barty *et al.* (2012) assumed a homo-geneous distribution of atoms/elements in the unit cell, and therefore a homogeneous degree of damage. The study concluded that the FEL-induced disorder gates the diffrac-tion, with undamaged high-resolution intensities collected at the beginning of the pulse which are superimposed with increasingly damaged lower-resolution intensities collected towards the end of the pulse. This effect was dubbed 'Bragg termination' and proposed to be correctable by a scaling algorithm (Barty *et al.*, 2012).

However, homogeneous stochastic atomic disordering of the crystal lattice does not address the possibility of localized electronic damage during the femtosecond duration X-ray pulse. Simulations predict that rapid ionization during the course of a femtosecond-duration X-ray pulse has the capacity to produce local changes in electron density by ionizing different elements at different rates, modifying the atomic scattering factors (Hau-Riege, 2007; Son *et al.*, 2011*a*) and leading to forces and atomic displacements that depend on the local environment (Jurek & Faigel, 2009). This is particularly important since macromolecules and their crystals have distinctly anisotropic distributions of elements, such as oxygen-rich solvent channels, SS-bridges, metal cofactors *etc.* which would lead to different degrees of damage at different

locations. Indeed, the scaling behaviour of integrated powder patterns from lysozyme nanocrystals collected with different pulse lengths (Lomb *et al.*, 2011) was consistent with the existence of local damage hot spots. Such inhomogeneous damage would change the molecular transform during the FEL pulse, compromising the data in a manner that cannot be rectified.

Obviously, a detailed understanding of the nature and extent of radiation damage is of tremendous importance, not only for the emerging method of serial femtosecond crystal-lography (SFX) but also for the development of single-particle imaging techniques using FELs. This is particularly relevant for the study of very small and/or weakly diffracting particles or crystals that require the use of very high flux densities and doses in the high GGy range (Chapman *et al.*, 2014).

Radiation damage is initiated by inner-shell photoabsorp-tion. Absorption cross sections are larger than scattering cross sections for most atomic species, meaning that, for every scattered photon that contributes to the diffraction pattern, many more photons are absorbed and deposit their energy in the sample. At the wavelengths used in X-ray diffraction experiments, K-shell absorption dominates for low-Z elements like C and S. The same holds for Fe above its K-shell absorption-edge at 7.1 keV. Below this absorption edge, the photoionization cross sections of Fe and S are similar (5272 barns and 7703 barns, respectively, at 6.86 keV, and about 40 and 55 times higher than for C at 137 barns, respectively). The photoionization cross section of Fe increases dramatically above the K-shell absorption edge (37640 barns), where Fe absorbs 5.4 times more strongly than S (and 306 times more than C) and therefore could be significantly more highly ionized through photoionization. Multiple ionization of iron in the high-intensity FEL regime affects X-ray absorption and anomalous scattering para-meters, and was proposed as a new route to phasing (Son *et al.*, 2011*b*). As a first approximation, the assumption was made that only the heavy Fe atom undergoes ionization dynamics during the X-ray pulse; neither the ionization dynamics of the low-Z atoms, atomic recombination associated with the large free-electron density nor structural rearrangements (Erk *et al.*, 2013) including those causing Bragg termination (Barty *et al.*, 2012) were considered, which may affect the phasing algo-rithm and reconstruction (Galli *et al.*, 2015). These are important issues to explore. To minimize effects of radiation damage on measured diffraction data, a detailed under-standing of the influence of FEL pulse parameters such as pulse fluence and duration will enable the optimization of experiments and their analysis, and the FEL sources them-selves.

Here, we investigate the effects of X-ray damage using the iron-containing protein ferredoxin as a model system. *Clostridium acidurici* ferredoxin is a small protein (55 amino acids), containing two iron–sulfur [4Fe–4S] clusters separated by $\sim$12 Å, which each have a cubane-like structure with a well defined spatial arrangement of iron and sulfur atoms (see Fig. 1). The structure of ferredoxin has been determined previously to 1.8 Å resolution from crystals kept at room
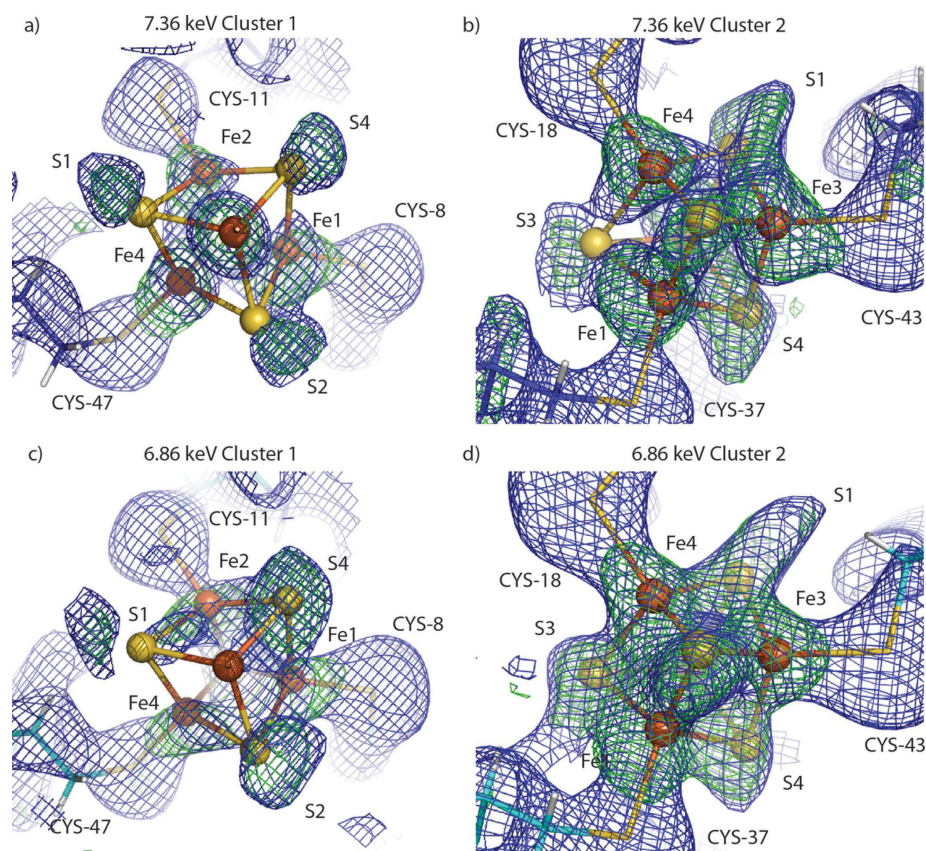
Figure 1
The [4Fe–4S] clusters in the high-dose SFX data at (a, b) 7.36 keV and (c, d) 6.86 keV. The $2mF_{obs} - DF_{calc}$ (blue, 1.0σ) and $F_{obs} - DF_{calc}$ (green, 2.5σ) maps were calculated after ten cycles of restrained refinement in *Refmac5*. The coordinates of the [4Fe–4S] clusters were removed from the model prior to refinement. The [4Fe–4S] clusters are superposed as a reference for the original positions of Fe and S atoms in the clusters.

described previously (Hong & Rabinowitz, 1970). For the LCLS experiments microcrystals were grown by batch crystallization mixing 100 µl protein (12 mg ml^{-1} in 15 mM Tris HCl pH 7.5, 100 mM NaCl) and 500 µl 4 M ammonium sulfate in 25 mM Tris HCl pH 7.8, 12 mM Na MOPS pH 5.8. This resulted in a rapid growth of relatively large crystals of inhomogeneous size distribution [(1.6 ± 0.5) µm $\times$ (1.6 ± 0.5) µm $\times$ (17 ± 7.5) µm] as shown in Fig. S1 of the supporting information[1] and used for the high-dose SFX datasets. While significantly smaller crystals of very homogeneous size distribution were obtained using a protein concentration around 3 mg ml^{-1}, these crystals did not diffract to as high resolution. They were used for the lower dose SFX dataset. The same protein batch used to grow the crystals for FEL data collection was used to grow macroscopic crystals for synchrotron measurements. The crystals were obtained in Linbro plates using the vapour diffusion method in the hanging drop geometry. Equal volumes of protein (6 mg ml^{-1} in 15 mM Tris HCl pH 7.5, 100 mM NaCl) and precipitant solution (3 M ammonium sulfate in 0.05 M MES pH 6.5) were mixed and equilibrated against the precipitant solution. Rod-shaped crystals appeared overnight and grew for a few days reaching a size of $\sim$30 µm $\times$ 30 µm $\times$ 100 µm.

temperature during data collection with a rotating anode (Duée *et al.*, 1994) and to 0.94 Å resolution using cryocooled crystals and synchrotron radiation, accurately defining the geometry of the [4Fe–4S] clusters (Dauter *et al.*, 1997). In the current study, the clusters serve as extremely sensitive markers of ionization and thermalization effects caused by FEL-induced radiation damage which we probed by SFX both below and above the iron K-edge. We deliberately chose to use highly damage-inducing conditions for the measurements, and to this end we used the highest available pulse energy ($\sim$1.5 mJ) and relatively long pulse duration (80 fs). For comparison, we also collected data at the LCLS using a lower flux density as well as datasets at the Swiss Light Source at room temperature.

2. Material and methods

2.1. Protein purification and crystallization

C. acidurici was obtained from the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH) using the order number DSM 604 (ATCC 7906). The cells were cultured under microaerobic conditions at 303 K using DSMZ medium #76. Ferredoxin was purified essentially as

2.2. Sample injection, post-sample attenuator characterization and data collection

FEL measurements were performed in June 2013 (experiment cxi76413) in the nanofocus chamber of the Coherent X-ray Imaging (CXI) instrument at the LCLS (Boutet & Williams, 2010). A suspension of ferredoxin microcrystals was injected into the FEL interaction region using a liquid microjet from a gas dynamic virtual nozzle (GDVN) (DePonte *et al.*, 2008; Weierstall *et al.*, 2012) at a flow rate of 25–30 µl min^{-1}. The LCLS provides pulses at 120 Hz, which were focused to approximately 200 nm $\times$ 200 nm using a pair of Kirkpatrick–Baez mirrors. The X-ray pulse duration was estimated to be 80 fs from the measured duration of the electron beam (Behrens *et al.*, 2014). The photon energy and pulse energy was 7.36 keV and 1.4 mJ (above the iron K-edge at 7.11 keV), respectively, and 6.86 keV and 1.6 mJ below the iron K-edge. Taking into account a beamline transmission of

[1] Supporting information for this paper is available from the IUCr electronic archives (Reference: XH5045).

~35–40%, the power density in the focus was estimated to be of the order of 1.8×10^{19} W cm^{-2}.

Diffraction patterns were recorded on the CSPAD detector, version 1.2 (Philipp *et al.*, 2010). The nominal detector distance was 90 mm. To avoid saturation of and damage to the CSPAD detector, a filter was placed 22 mm downstream of the interaction region, *i.e.* between the liquid jet delivering the microcrystals and the CSPAD detector. Any material of uniform thickness and density will attenuate high-resolution diffraction by a much larger factor than the on-axis beam, due to the increased path length of high-angle diffracted rays through a slab of finite thickness. To counteract this unfortunate situation and prevent the weakest high-angle diffraction from being measured with worse statistics, filters were fabricated with a thickness that decreased from the centre to the edge. These were made such that they give uniform transmission when placed a distance 1 cm from the sample, which results in uniformly increasing transmission with scattering angle when used at distances greater than this. The particular filter used here consisted of two foils coated with a thickness-graded alloy of $W_{0.84}Ta_{0.16}$ with a density of 16 g cm^{-3} and a combined thickness of 7.7 µm on-axis as designed and verified in DESY's multilayer coating laboratory. At the 22 mm sample-to-filter distance used here, the transmission of this filter varied approximately linearly from 6% (7.9%) at the lowest scattering angle to 15% (18%) at the highest angle for 6.86 keV (and 7.36 keV) (see Fig. S2 of the supporting information). The graded filter transmission was measured at more than 7000 positions using the four-crystal monochromator beamline of PTB at BESSY II where highly monochromatic radiation is available (Krumrey & Ulm, 2001).

In preparation for the SFX damage experiment, a test SFX dataset from smaller ferredoxin microcrystals had been collected in June 2012 (experiment cxi59712) in the microfocus chamber of the CXI instrument using the full X-ray beam. The photon and pulse energy was 7.23 keV and 1.3 mJ, respectively. This dataset serves as a lower dose SFX comparison dataset, collected above the Fe *K*-absorption edge.

In addition, synchrotron ferredoxin datasets were collected at the PXIII beamline of the Swiss Light Source (SLS) using a Pilatus 2M detector. The macroscopic crystals were mounted on a Mitegen mesh surrounded by a mylar sleeve and kept at room temperature during data collection. The X-ray beam was focused to 80 µm × 45 µm. The flux was ~1×10^{11} photons s^{-1} and 3×10^{10} photons s^{-1} for the datasets collected using the same photon energy as for the FEL measurements above (7.36 keV) and below (6.86 keV) the iron *K*-edge, respectively, using one crystal (above) and two crystals (below) for each dataset.

2.3. Data processing, refinement and electron density map calculation

LCLS data in XTC format were processed using *CASS* (Foucar *et al.*, 2012). Snapshots containing more than ten Bragg reflections were classified as hits and saved to individual HDF5 files, together with the data describing each individual FEL pulse (pulse energy, wavelength, current detector distance, *etc.*) corresponding to each exposure. Bragg peaks were identified with the *PostProcessor 208* in the offset-corrected and bad-pixel-masked frames using a set box size of 5 × 5 pixels, which results in a box size of 11 × 11 pixels, a threshold of 400 ADU and a signal-to-noise ratio of 3. *PostProcessor 208* is implemented in *CASS* (Foucar, 2015; to be published). HDF5 files were analysed using *CrystFEL* 0.5.3a (White *et al.*, 2012). Diffraction patterns were indexed and individual intensities were then merged using Monte Carlo integration (Kirian *et al.*, 2011) using *CrystFEL* while excluding reflections that contained pixels with more than 3500 ADU (corresponding to ~100 8 keV X-ray photons). The position as well as the roll, pitch and yaw parameters of each of the detector tiles on the multi-tile CSPAD detector were refined by minimizing the distance between indexed observed and theoretically predicted Bragg peak positions using a custom program, which will be described elsewhere. This procedure increased the indexing rate of the diffraction patterns and improved the data quality as judged by R_{split} (White *et al.*, 2012). The merged intensities from *CrystFEL* were converted to structure factors and saved to CCP4 (mtz) format using *XDSCONV* (Kabsch, 2010). The synchrotron data were processed with *XDS* (Kabsch, 1988).

Both the LCLS and SLS data were phased by molecular replacement with *Phaser-MR* (McCoy *et al.*, 2007) using the high-resolution ferredoxin coordinates from the Protein Data Bank (2fdn) (Dauter *et al.*, 1997) as a search model. After removing the [4Fe–4S] clusters from the model obtained with molecular replacement, ten cycles of restrained refinement were performed using *Refmac5* (Murshudov *et al.*, 2011). $2mF_{obs} - DF_{calc}$ and $mF_{obs} - DF_{calc}$ (Read, 1986) electron density maps were calculated, yielding unbiased omit maps of the clusters. As an additional check, the normalized $2mF_{obs} - DF_{calc}$ maps from *Refmac5* were subtracted after superposition using *Phenix* (Adams *et al.*, 2010) to exclude the possibility that the features observed in the $F_{obs} - F_{obs}$ maps are due to non-isomorphism. In all cases, except for the difference between the low-dose LCLS data and the SLS data, the maps obtained by subtraction of $2mF_{obs} - DF_{calc}$ maps were consistent with the corresponding $F_{obs} - F_{obs}$ maps (results not shown).

Anomalous difference Fourier maps were calculated with the *FFT* program (Collaborative Computational Project, Number 4, 1994) using phases obtained from molecular replacement. After scaling the datasets (see Tables S1–S5 of the supporting information) with *XSCALE* (Kabsch, 1988), difference electron density maps between SLS and LCLS datasets were calculated with the *FFT* program, using molecular replacement phases. For the calculation of the simulated annealing omit maps, the Tyr30 residue was deleted in its entirety followed by simulated annealing in torsion angle and Cartesian space. Electron density maps were visualized with *Pymol* (DeLano, 2002).

2.4. Plasma code calculations

Photon–ferredoxin interactions were modelled to study the ultrafast damage dynamics in the experiment using the non-local thermodynamic equilibrium (non-LTE) plasma approach as implemented in the software package *CRETIN* (Scott, 2001). *CRETIN* is a well established plasma code, developed to simulate astrophysical plasmas and plasmas generated in the laboratory (Rubiano *et al.*, 2007). It calculates the electronic level populations and transition rates, from which opacities, heating rates and other material properties are generated with time. The non-LTE code treats the sample as a continuum and it is not limited by the size or the number of molecules in the system. This allows the simulation of large systems such as protein crystals at affordable computational costs. The trade-off is that it gives no specific information about the changes in atomic structure. The code provides statistical quantities of the system, such as ion and electron temperatures and average ion displacements, but it can give no information about local damage at specific locations in the sample. In the type of samples studied here, *i.e.* micrometre-sized protein crystals, the plasma model is expected to give a good average description of the physics, and it has proven to do so in several earlier studies in the soft X-ray regime at the FLASH free-electron laser in Hamburg (Andreasson *et al.*, 2011; Bergh *et al.*, 2008; Hau-Riege *et al.*, 2007) as well as in the hard X-ray regime at LCLS (Chapman *et al.*, 2014).

The simulation model does not contain detailed structural information, such as atomic positions; however, it follows the atomic populations in time and includes ionization rates, electron collisions and radiation transfer. The rates from these processes are calculated, and a self-consistent rate-equation is solved, describing the evolution of the ionization and collision rates, that ultimately lead to structural degradation. This model was used by Barty *et al.* (2012) to describe the global atomic disorder. Further details can be found in the simulations described by Jönsson and co-workers in this issue (Jönsson *et al.*, 2015). For simplicity, here a flat-top X-ray pulse model is used. This keeps track of the average temperatures (a two-temperature approach, for electrons and ions is used) and average ionization but excludes any local structural information, assuming that the atomic population is homogeneous throughout the sample during the exposure. The sample is simulated in nanometre-sized zones, with an average atomic composition corresponding to crystalline ferredoxin ($C_{228}H_{359}N_{61}O_{83}Fe_8S_{16}$) with water as the buffer. The density was modelled to be 1.35 g cm^{-3}. The simulations were performed with X-ray beam parameters matching the experimental conditions, *i.e.* photon energies of 6.86 keV and 7.36 keV, pulse length 80 fs, pulse power density 1.4×10^{19} W cm^{-2}.

3. Results and discussion

Radiation damage at conventional X-ray sources is typically divided into two categories: global and local damage; the latter is also called specific damage. With the accumulation of the dose during data acquisition, global damage manifests itself as an overall change of the diffracted signal, such as a loss of high-resolution diffraction, an increase of the Wilson *B*-factor and mosaicity, as well as changes of the unit-cell parameters. In contrast, local damage, such as the breakage of disulfide bonds, decarboxylation of aspartate and glutamate residues (Ravelli & McSweeney, 2000; Burmeister, 2000), or changes in oxidation and coordination states of metal cofactors (Kühnel *et al.*, 2007; Schlichting *et al.*, 2000) results in local changes in the molecule, producing changes in individual Bragg peak intensities that cannot be accounted for by an overall scaling factor. While there is a rich history of studying radiation damage at both cryogenic and ambient temperatures using synchrotron radiation (*e.g.* Burmeister, 2000; Owen *et al.*, 2006, 2012; Ravelli *et al.*, 2003; Ravelli & McSweeney, 2000), few experiments have been performed so far at FEL sources. The very bright femtosecond pulses allow most damage processes to be outrun, in particular those involving radical chemistry, which depend on radicals diffusing through the crystal. This has enabled the structure determination of the fully oxidized resting state of bovine cytochrome c oxidase, demonstrating a peroxide ligand (Hirata *et al.*, 2014) and the undamaged oxygen evolving complex of photosystem II at the SACLA FEL (Suga *et al.*, 2015). These studies used exposure conditions that resulted in a dose of 9.9 and 1.4 MGy, respectively. A particular dose level of 400 MGy for FEL radiation was described by Chapman and co-workers (Chapman *et al.*, 2014). This is the dose for which, for an average protein crystal, each atom is ionized once at the end of the pulse. Thus, during the pulse a photon is more likely to be scattered from a neutral atom than an ionized atom. Ionization is dominated by collisional ionizations from the high-temperature electrons in the plasma; direct photoabsorption accounts for less than 1% of all ionizations in this case of a 400 MGy dose. At higher doses it is expected that the pulse-integrated diffraction will be affected by the ionizations. Indeed, a decrease of high-resolution diffraction with increasing FEL pulse length (nominal electron bunch length 70–400 fs) was observed for doses in the GGy range, which was both dose- and dose-rate-dependent (Barty *et al.*, 2012; Lomb *et al.*, 2011). In particular, an increase in atomic disorder during the pulse causes thermal disordering of the crystalline lattice, resulting in a resolution-dependent termination of crystalline diffraction during the course of the X-ray pulse.

Bragg termination is reminiscent of the global damage observed at synchrotron sources, despite fundamentally different regimes due to extremely different time scales and mechanisms of the radiation damage processes at the two X-ray sources. Importantly, radiation damage at synchrotrons typically accumulates over several exposures, allowing correction by zero-dose extrapolation (Diederichs, 2006). In the diffraction-before-destruction regime at FELs, however, the sample may be completely destroyed before the pulse is even over, with the decaying diffraction being integrated over the pulse duration. Measurements and simulations at different pulse irradiances may offer a way to correct for the effect of Bragg termination (Barty *et al.*, 2012).

The question of local damage in the high-intensity FEL regime remains. Based on the scaling behaviour of integrated powder patterns of lysozyme nanocrystals exposed to FEL X-ray pulses with different pulse lengths, we had also predicted the existence of the equivalent of local damage in FEL data (Lomb *et al.*, 2011). However, the photon energy and thus the resolution of the data available at the time did not permit a study of the underlying structural changes.

Here, the nature of radiation damage in the high-intensity regime at FEL sources is explored, using ferredoxin as a model system. Of particular interest is the behaviour of the two electron-rich [4Fe–4S] clusters and their surrounding environments. High-intensity SFX data from ferredoxin microcrystals were collected using 80 fs duration X-ray pulses above and below the iron K-edge at an irradiance in the low 10^{19} W cm^{-2} range. These conditions ensure significant damage and result in an average dose of $\sim$30 GGy and $\sim$20 GGy for each crystal above and below the Fe absorption K-edge, respectively, calculated using *RADDOSE-3D* (Zeldin *et al.*, 2013). For comparison, we collected a lower dose (0.7 GGy) comparison SFX dataset using 30 fs duration X-ray pulses above the iron K-edge at an irradiance of $\sim$4 $\times$ 10^{17} W cm^{-2}, as well as synchrotron rotation data.

3.1. Overall data quality

Careful evaluation of data quality is very important in radiation damage studies to ensure that the differences observed are indeed due to damage and not to differences in experimental protocols and/or data quality. The former includes measurements performed at different temperatures and/or using significantly different techniques such as conventional synchrotron measurements and SFX measurements.

Two high-intensity SFX datasets were collected from ferredoxin microcrystals in the CXI nanofocus chamber. At 7.36 keV photon energy 1139172 diffraction images were collected (in 156 min), 29586 (2.6%) of which were classified as crystal hits of which 21706 (73.4%) could be indexed with *CrystFEL*, yielding a 2.0 Å resolution dataset. At 6.86 keV photon energy 1503362 diffraction images were collected (in 210 min), 11897 (0.8%) of which were classified as crystal hits, of which 9336 (78.5%) could be indexed with *CrystFEL*, yielding a 2.1 Å resolution dataset. As the transmittance of the post-crystal attenuator varied linearly from 6% at the lowest scattering angle to 15% at the highest scattering angle for the photon energy of 7.36 keV, a linearly varying factor was applied to the merged intensities correcting for the transmission gradient. The factor varied linearly from 1 at the lowest resolution to 0.4 at the highest resolution. Changing the value of the correction factor at the highest resolution by $\pm$20% or not applying it at all had no influence on any of the observations described below, as expected. The statistics of the two datasets are given in Table 1. For comparison, a lower dose SFX dataset was collected from ferredoxin microcrystals in the CXI micro-focus chamber. At 7.23 keV photon energy,

380505 diffraction images were collected (in 53 min), 85705 of which were classified as crystal hits of which 6840 (8%) could be indexed with *CrystFEL*, yielding a 2.5 Å resolution dataset (see Table 1). In addition, room-temperature reference datasets from much larger crystals grown from the same protein batch using very similar crystallization conditions were collected at the SLS using the same photon energies as for the high-dose SFX data. The crystals have the same space group and similar unit-cell constants as the microcrystals used at the LCLS (Table 1). In the SLS case, the dose was $\sim$500 kGy and $\sim$200 kGy for the 2.0 Å resolution datasets collected above (7.36 keV) and below (6.86 keV) the Fe K-edge, respectively. The statistics of the two synchrotron (SLS) datasets are also given in Table 1.

In SFX, each crystal is hit by the FEL beam only fleetingly and does not move during the femtosecond exposure that destroys it. Thus still images are collected, containing only partial reflections. Fully integrated intensities can be obtained by Monte Carlo integration (Kirian *et al.*, 2010, 2011), which merges all measured partial intensities, averaging out fluctuations occurring during data collection (different crystal sizes, incident beam intensities, spectral fluctuations of the FEL beam *etc.*), provided enough individual measurements exist for each unique reflection. This dependence explains the poorer data statistics of the 6.86 keV LCLS dataset compared with the 7.36 keV LCLS dataset, as it contains only half the number of indexed patterns. Indeed, the ratios of signal-to-noise ratios and R_{split} values between the two datasets correspond approximately to the factor of $\sqrt{2}$ expected for a twofold increase in the number of observations. However, the average multiplicity of the 6.86 keV SFX data is $\sim$170, which is in the same range as for other high-resolution SFX structures (*e.g.* Boutet *et al.*, 2012). Despite having similar multiplicities, resolution (2.1 Å in this study and 1.9 Å in Boutet *et al.*) and the same space group ($P4_32_12$), the data statistics are not as good, in particular when considering R_{split} (see Table 1). Also, while the statistics for the low-dose LCLS dataset are poor as well, the Wilson plot of this dataset appears normal and the cumulative intensity distributions do not show an 'averaging' effect as do the high-dose LCLS data (see below). The relatively poor statistics of the high-dose ferredoxin data are even more significant when taking into account that the Boutet *et al.* lysozyme data were collected using an earlier version of the CSPAD detector. The poor statistics are unlikely to be due to non-isomorphism of the microcrystals, since the widths of the unit-cell distributions of the LCLS datasets are the same (not shown for all, see Fig. S3). We scaled the LCLS and SLS data using *XSCALE* (Kabsch, 1988) (see Tables S1–S5) and *Phenix* (see Tables S6–S9). For the 7.36 keV (6.86 keV) data the correlation coefficient was 0.85 (0.86). The correlation coefficient of the low-dose LCLS data and the high-dose LCLS data (and SLS data) is 0.88 (0.94) and 0.87 (0.96) for the data collected above and below the Fe K-edge, respectively.

It is interesting to note that the correlations are of the order of those expected for a heavy atom derivative, and it is thus

Table 1
Data collection statistics.

See Tables S11–S13 for details of the statistics of the SFX datasets.

	LCLS high dose 7.36 keV†	LCLS high dose 6.86 keV†	LCLS low dose 7.23 keV†	SLS RT 7.36 keV†	SLS RT 6.86 keV†
Software	*CrystFEL* 0.5.3a	*CrystFEL* 0.5.3a	*CrystFEL* 0.5.3a	*XDS*	*XDS*
Data collection					
Wavelength (Å)	1.685 ± 0.001	1.807 ± 0.001	1.715 ± 0.001	1.685	1.807
Photon energy (keV)	7.36	6.86	7.23	7.36	6.86
Pulse energy (mJ)	1.4 ± 0.1	1.6 ± 0.1	1.3 ± 0.1	–	–
Photon flux at the sample position‡	$\sim 4.8 \times 10^{11}$ photons pulse^{-1}‡	$\sim 4.8 \times 10^{11}$ photons pulse^{-1}‡	$\sim 4.4 \times 10^{11}$ photons pulse^{-1}‡	9.8×10^{10} photons s^{-1}	3.0×10^{10} photons s^{-1}
Pulse duration (fs)§	80 ± 3	80 ± 3	30 ± 3	–	–
Focus	$\sim 0.2\ \mu m \times 0.2\ \mu m$	$\sim 0.2\ \mu m \times 0.2\ \mu m$	$\sim 2.0\ \mu m \times 2.0\ \mu m$	$45\ \mu m \times 80\ \mu m$	$45\ \mu m \times 80\ \mu m$
Power density (W cm^{-2})‡	1.8×10^{19}	1.7×10^{19}	4.2×10^{17}	–	–
Crystals					
Dimensions (μm)	$1.6 \pm 0.5 \times 1.6 \pm 0.5 \times 17 \pm 7.5$	$1.6 \pm 0.5 \times 1.6 \pm 0.5 \times 17 \pm 7.5$	$1.0 \pm 0.2 \times 1.0 \pm 0.2 \times 5 \pm 1.0$	$30 \times 30 \times 80$	$30 \times 30 \times 80$
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
a (Å)	34.1 ± 0.1	34.1 ± 0.1	34.7 ± 0.1	34.6	33.6
b (Å)	34.1 ± 0.1	34.1 ± 0.1	34.7 ± 0.1	34.6	33.6
c (Å)	74.5 ± 0.1	74.5 ± 0.1	74.5 ± 0.1	74.9	74.9
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	31.03–2.0 (2.1–2.0)	31.03–2.1 (2.2–2.1)	31.51–2.5 (2.6–2.5)	17.48–2.0 (2.1–2.0)	17.48–2.0 (2.1–2.0)
Dose (GGy)	~ 30	~ 20	~ 0.7	5.4×10^{-5}	2.0×10^{-5}
Statistics					
Total # observations	2281705 (82905)	840310 (23885)	380505 (12866)	16472 (1887)	14206 (815)
Number of hits	29586	11897	85705	–	–
Indexed patterns	21706	9336	6840	120	160
Indexing rate (%)	73.4	78.5	8.0	100.0	100.0
Multiplicity	398.5 (147.5)	170.8 (57.6)	203.6. (79.4)	18.2 (16.3)	17.4 (2.7)
Unique reflections	5717 (562)	4857 (415)	1839 (162)	5541 (668)	3084 (304)
Completeness (%)	99.9 (99.5)	98.3 (85.6)	99.9 (99.4)	94.2 (83.6)	90.3 (68.2)
$I/\sigma(I)$	8.7 (2.3)	5.5 (0.5)	4.6 (1.0)	13.8 (10.5)	15.4 (7.2)
R_{split} (%)¶	19.4 (62.0)	32.2 (101.4)	37.7 (107.5)	–	–
$CC_{1/2}$ (%)	82.3 (29.1)	73.2 (27.8)	77.9 (33.9)	99.3 (98.7)	99.6 (98.5)
CC^* (%)	93.5 (67.1)	90.4 (66.0)	91.6 (71.1)	–	–
CC_{ano} (%)	18.7 (0.0)	–	11.0 (0.0)	78.0 (70.0)	–
Wilson B (Å^2)††	3.8	15.8	54.1	7.7	7.1

† The SFX xtc files and the integrated intensities of the SLS datasets will be deposited in the cxidb.org database. ‡ Taking into account a beamline transmission of 40%. § Standard deviation of the X-ray pulse duration was estimated from measured electron bunch lengths. ¶ $R_{split} = (1/\sqrt{2})(\sum_{hkl} |I_{hkl}^{even\_images} - I_{hkl}^{odd\_images}|)/[\sum_{hkl}(1/2)(I_{hkl}^{even\_images} + I_{hkl}^{odd\_images})]$. †† Calculated by *TRUNCATE*.

not inconceivable that the low correlations at least in part also reflect a degree of local radiation damage to the iron–sulfur clusters.

When we analysed the intensity distribution of the ferredoxin diffraction data by generating Wilson and cumulative intensity distribution plots for the LCLS and synchrotron datasets, it became apparent that the LCLS high-dose SFX datasets differ from the expected distributions, which are exemplified by the low-dose SFX and synchrotron rotation data (see Fig. S4). In particular, both strong and weak intensities are under-represented in the high-dose FEL data. This effect is independent of the resolution limits used (not shown) and so is unlikely to be the result of difficulties with the integration of weak spots at the edge of the detector, a common problem in SFX. Moreover, the low-dose SFX data show a normal intensity distribution.

To investigate whether the unusual intensity distribution of the high-dose FEL data could be due to local damage to the iron–sulfur clusters, which are responsible for a considerable fraction of the total scattering, structure factors (F_{calc}) were calculated for ferredoxin with and without coordinates of the two [4Fe–4S] clusters included, identifying a number of reflections that differ significantly in magnitude. The ratio of intensities of these reflections between the synchrotron data (corresponding to the situation with full cluster) and FEL data (corresponding to the situation with a potentially damaged cluster) were then analysed, and compared with ratios of randomly chosen pairs of reflections. No correlation between the calculated and observed pairs of reflections was observed (see Fig. S5). Next the effects of Bragg termination on the Wilson and intensity distribution statistics were simulated by modifying the low-dose synchrotron intensities using formula (3) from (Barty *et al.*, 2012), assuming average atomic displacements of either 10 Å [the maximum value expected according to Chapman *et al.* (2014)] or 40 Å (chosen to clarify the effect) at the end of the pulse. This, however, reproduced neither the observed Wilson plots nor the cumulative intensity distributions observed for the high-dose LCLS data (see

Fig. S6). This can be explained by assuming that under the conditions of the current experiment the ionization effects are so strong (as outlined below) that ionization will cause termination of the Bragg diffraction (Caleman *et al.*, 2015), rather than the increasing disorder in the structure as described by Barty *et al.* (2012).

In fact, in the case of the cumulative intensity distribution for the 40 Å displacement Bragg termination simulation, the curve for acentric reflections even lies above the expected distribution rather than below it, as observed for the LCLS ferredoxin data. However, it is noteworthy that several SFX datasets of other proteins indeed show such an effect (data not shown), which thus could be explained by Bragg termination.

For high-dose SFX data from ferredoxin, however, the cumulative intensity distributions do not only lie below the expected curve but are actually sigmoidal, such as those expected for the cumulative intensity distribution from a twinned crystal. This can only be caused by an averaging of different structure factor amplitudes, not by global damage effects such as an increased *B*-factor or Bragg termination, which do not affect the structure factors but modulate the intensities in a continuous resolution-dependent fashion. A possible scenario explaining this could be different degrees of damage for different crystals, or different parts of the crystals being damaged to a different extent. This is in line with the experimental setup, which involves crystals that are significantly larger than the intense X-ray nanofocus. Thus, only small parts of the crystals will be exposed to the focused beam (or even the focused part of the beam may have missed the crystal entirely) while significantly larger parts of the crystals will be grazed by the 'halo' around the focused X-ray beam. Given the strong scattering power of the crystals, even very weak parts of the X-ray beam are likely to result in appreciable diffraction signal. To test this hypothesis, structure factors calculated from five ferredoxin models were averaged, to which random shifts of the atoms of up to 1.5 Å had been applied. The resulting averaged structure factor amplitudes indeed showed the expected sigmoidal cumulative intensity distribution (see Fig. S7).

Given the unusual cumulative intensity distribution, the information content of the SFX intensities became of interest. Therefore, simulated annealed omit maps were calculated for the 7.36 keV synchrotron and high-dose FEL data. As can be seen in Fig. S8, omitting the coordinates of a tyrosine residue from the model results in excellent difference electron density in the case of the synchrotron data and reasonable density obtained with the low-dose SFX data in contrast to the situation for the high-dose SFX data, where no interpretable difference density was apparent. This shows that the information content of the high-dose SFX data is rather low. This is remarkable, given that the multiplicity of our SFX data far exceeds that of, for example, the SFX data of the *B. viridis* photosynthetic reaction centre, which was ~27, and that difference maps calculated from those data clearly showed electron density for cofactors *etc.* (Johansson *et al.*, 2013). This observation, as well as the finding that the low-dose SFX data yield reasonable electron density maps, makes it very unlikely

that the low information content and low data quality of the high-dose SFX data are due to poor convergence of the Monte Carlo integration.

3.2. The iron–sulfur clusters

C. acidiurici ferredoxin contains two iron–sulfur clusters, cluster 1 [consisting of S1, S2, S3, S4, Fe1-SG8, Fe2-SG11, Fe3-SG14, Fe4-SG47, same nomenclature as used by Dauter *et al.* (1997) with SG indicating the thiol sulfur of the coordinating cysteine residues] and cluster 2 [S1, S2, S3, S4, Fe1-SG37, Fe2-SG40, Fe3-SG43, Fe4-SG18]. The occupancy of the two iron–sulfur clusters can vary between different protein preparations, depending on both the growth conditions of the bacterial cells as well as the details of the protein purification protocol. To avoid any influence on the data from these purely biochemical factors, the same protein sample was used to obtain the crystals utilized for measurements at both the LCLS and the SLS. Importantly, the SLS crystals were grown after the high-dose LCLS beam time, eliminating protein degradation as a possible source for the observation of lower electron density in the high-dose FEL data. Also, since both the high-dose LCLS data and the SLS data were collected using the same photon energy and similar temperature, any difference in the electron density of the clusters is likely to be related to radiation damage and not due to using different data collection techniques *per se*. Note that we have shown previously that undamaged SFX data of lysozyme microcrystals agree well with the corresponding synchrotron rotation data of large crystals (Boutet *et al.*, 2012) as well as with synchrotron serial crystallography data of small crystals (Botha *et al.*, 2015).

In order to obtain unbiased electron densities of the [4Fe–4S] clusters, ten cycles of restrained refinement were performed in *Refmac5* (Murshudov *et al.*, 2011) with the clusters omitted from the 2fdn model (Dauter *et al.*, 1997) used as the starting model for refinement, before calculating $2mF_{obs} - DF_{calc}$ and $mF_o - DF_c$ maps (Read, 1986) for the two ferredoxin datasets collected at the LCLS and the SLS, both above and below the iron *K*-shell absorption edge. The results are shown in Fig. 1 (LCLS high-dose data), Fig. 2 (LCLS low-dose data) and Fig. 3 (SLS data) with the cluster coordinates of the 2fdn model superimposed as a reference for the positions of Fe and S. Despite the fact that the iron–sulfur clusters contribute a very significant fraction to the scattering power of the relatively small ferredoxin molecule, clearly defined density is observed for the clusters in the omit maps (Figs. 1–3). Moreover, as shown below, there is significant anomalous difference density (see Fig. S9) at the iron positions of the clusters. However, closer inspection of the omit maps calculated from the high-dose SFX data shows that the electron density of cluster 1 is significantly weaker than that of cluster 2 which is less defined (Fig. 1, see also Table S10 for integrated electron densities). Interestingly, the former is also the case for the low-dose SFX data (Fig. 2, Table S10) whereas neither feature is apparent in the synchrotron data (Fig. 3, Table S10). Another feature in the difference maps is that
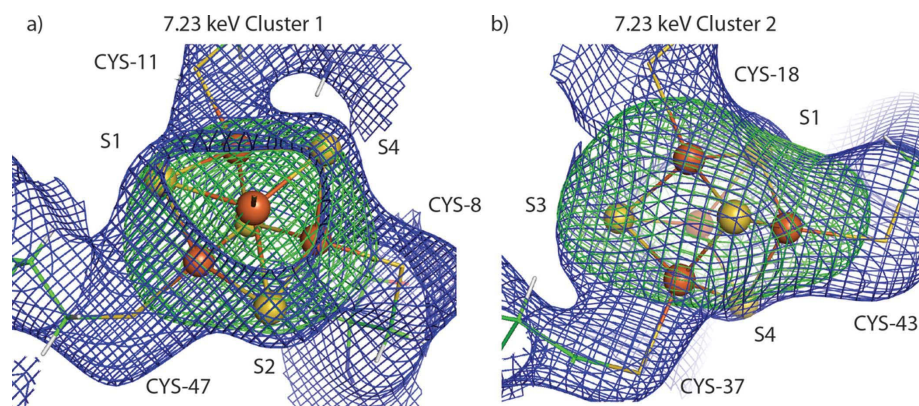
Figure 2
The [4Fe–4S] clusters in the 7.23 keV low-dose SFX dataset. The $2mF_{obs} - DF_{calc}$ (blue, 1.0σ) and $F_{obs} - DF_{calc}$ (green, 2.5σ) maps were calculated after ten cycles of restrained refinement in *Refmac5*. The coordinates of the [4Fe–4S] clusters were removed from the model prior to refinement. The [4Fe–4S] clusters are superposed as a reference for the original positions of Fe and S atoms in the clusters. The width of the electron density reflects the lower resolution compared with the data shown in Figs. 1 and 3.

there are distinct positive difference density peaks close to the sulfur positions but on the outside of the clusters (Fig. 1). In particular, sulfur atoms S1, S2, S4 of cluster 1 and sulfur atoms S1, S3 of cluster 2 are not at the 2fdn atom locations but have moved away from the centre of the clusters. This effect is large enough to come out of any model bias after simulated annealing, as it is also visible when the clusters are part of the model (data not shown). This feature is strongest for the high-

Figure 3
The [4Fe–4S] clusters in the SLS data at (*a*, *b*) 7.36 keV and (*c*, *d*) 6.86 keV. The $2mF_{obs} - DF_{calc}$ (blue, 1.0σ) and $F_{obs} - DF_{calc}$ (green, 2.5σ) maps were calculated after ten cycles of restrained refinement in *Refmac5*. The coordinates of the [4Fe–4S] clusters were removed from the model prior to refinement. The [4Fe–4S] clusters are superposed as a reference for the original positions of Fe and S atoms in the clusters.

dose SFX data, but also apparent for the data collected at the SLS [with data collected at room temperature producing a stronger effect than data collected at cryogenic temperature with the same dose (data not shown)]. The reason for this is unknown, given that the experimental parameters used for the acquisition of the 2fdn data are not published. Therefore the absorbed dose and thus the extent of damage is unknown. The observation of a possible radiation-damage-induced movement of sulfur atoms in synchrotron derived structures is, however, reminiscent of the findings by Weik *et al.* (2002). Interestingly, moreover, recent molecular dynamics simulations show significant correlated movement of the sulfur atoms of the [4Fe–4S] clusters in ferredoxin nanocrystals exposed to intense FEL pulses of the same pulse length, photon and pulse energy (Hau-Riege & Bennion, 2015) as used in the experiment reported here.

Photoabsorption depends strongly on atomic number, rendering metal centres such as the iron–sulfur clusters in ferredoxin highly susceptible to radiation damage. Since the crystals used for the FEL and synchrotron measurements are largely isomorphous, it is possible to calculate difference electron density maps [$F_{obs}(SLS) - F_{obs}(LCLS)$, see *Methods* for details, and Section S3 of the supporting information]. One has to bear in mind, however, that the effects of the loss of electrons through ionization, increased temperature factors due to disorder and the motion of cluster atoms, as well as effects caused by different experimental protocols or sample preparations can all manifest themselves as difference peaks. Therefore, the experimental conditions for the high-dose FEL and SLS datasets were kept as close to identical as possible. $F_{obs}(SLS) - F_{obs}(LCLS)$ difference electron density maps were calculated for the datasets collected above and below the iron K-absorption edge, respectively, using the ferredoxin model obtained after molecular replacement (with the clusters omitted) for phasing. The maps show significant positive peaks at the iron positions of the [4Fe–4S] clusters, in line with a strong ionization of the iron atoms (Fig. 4).

Given the difference in the photoionization cross section ratios of Fe and S above ($\sigma_{Fe}/\sigma_S = 5.5$) and below

Figure 4
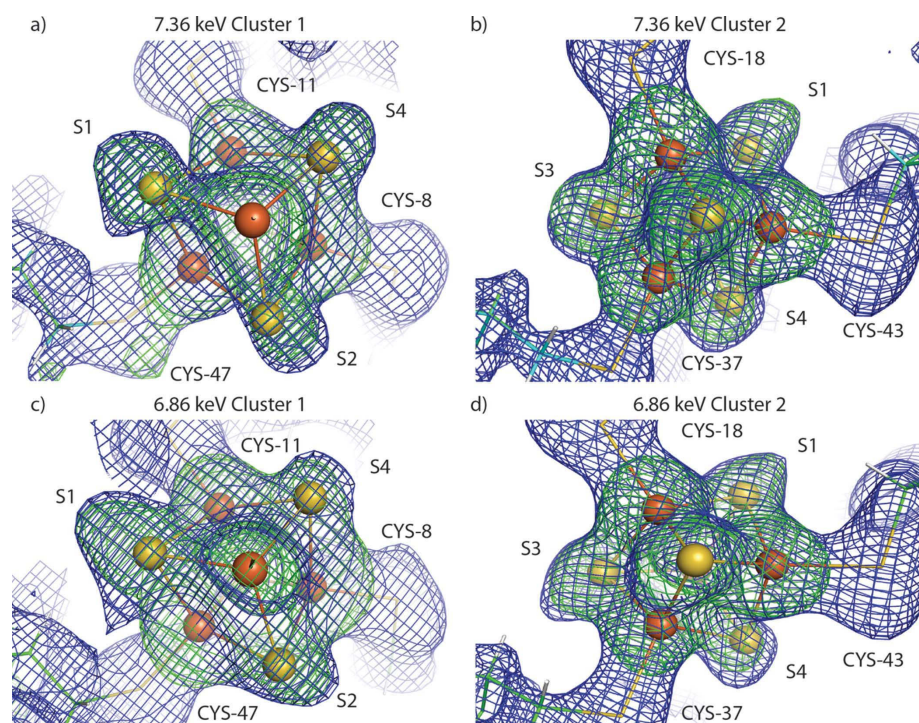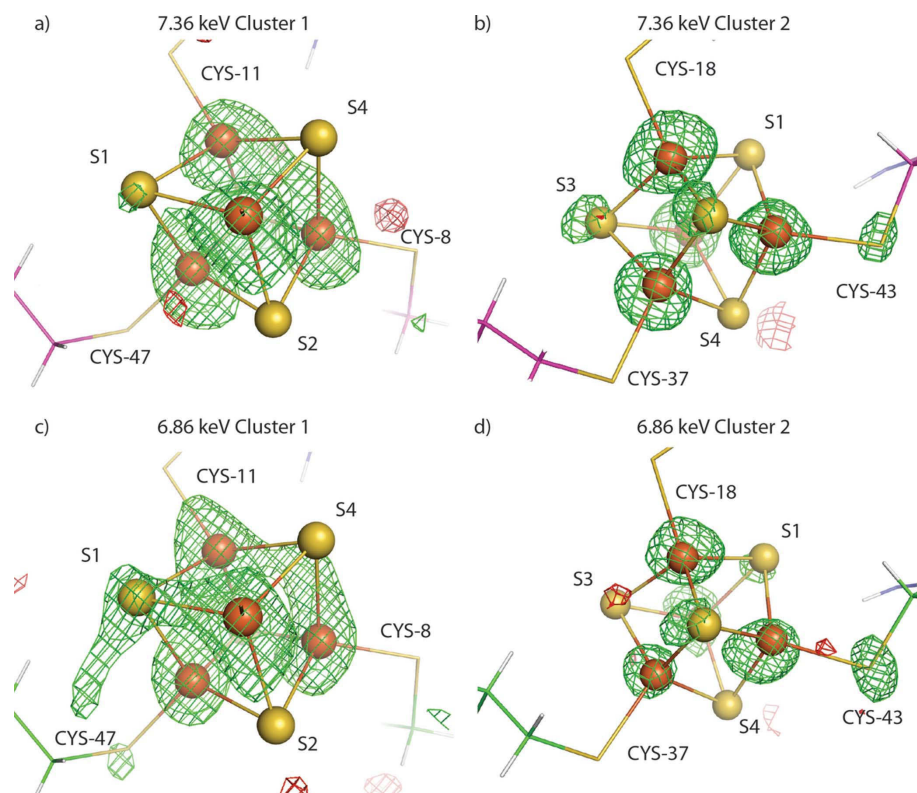Difference electron density maps $F_{obs}(SLS) - F_{obs}(LCLS)$ contoured at $\pm 3\sigma$ (positive, green mesh; negative, red mesh) calculated for the SLS and LCLS high-dose datasets measured at (a, b) 7.36 keV and (c, d) 6.86 keV.

($\sigma_{Fe}/\sigma_S = 0.7$) the Fe K-edge, significantly higher ionization of iron had been expected in the 7.36 keV dataset. However, neither the difference electron density map calculated between the SLS [$F_{obs}(SLS_{7.36\,keV}) - F_{obs}(SLS_{6.86\,keV})$] nor the LCLS [$F_{obs}(LCLS_{7.36\,keV}) - F_{obs}(LCLS_{6.86\,keV})$] data above and below the iron K-absorption edge show strong features at the iron positions (see Fig. S10). A possible reason for this observation in the synchrotron data is that the iron–sulfur clusters are highly electrophilic, attracting some of the several hundred electrons released by a single X-ray absorption event which results in reduction of the clusters. Although most of the electrons in the nascent plasma in the SFX high-dose experiments may have too high a kinetic energy to be captured by electrophilic atoms more predominantly than other atoms and ions, the electron temperature distribution will include a population of slow electrons that indeed may preferentially be attracted to the iron–sulfur clusters. Also, even though the conditions are markedly different, photoreduction to the all-ferrous state has been reported for the [2Fe–2S] cluster in putidaredoxin crystals kept at 110 K during exposure which resulted in a 1 MGy dose (Corbett *et al.*, 2007).

3.3. Ultrafast damage processes

Radiation damage processes in proteins are strongly modified by the presence of atoms with high atomic numbers (Z), rendering metal centres such as the iron–sulfur clusters in ferredoxin particularly susceptible to ionization, since such atoms will lose proportionally more electrons by direct photoionization than will lighter atoms. Damage is initiated by inner-shell photoabsorption. The excited atoms quickly relax through non-radiative Auger decay, which is more probable than radiative decay even for Fe. Auger decay occurs much more rapidly in Fe (0.55 fs) and S (1.3 fs) than in C (10 fs) (Campbell & Papp, 2001), and can be considered to be practically instantaneous for the 80 fs pulse length used in this study. With typical velocities between 60 and 500 Å fs^{-1}, the Auger and photoelectrons generate electron cascades through electron impact ionization, leading to the emission of secondary electrons with velocities below 30 Å fs^{-1}. It has been estimated that one 6 keV photoelectron will lead to 300 secondary electrons, and each Auger electron will give 10–20 secondary electrons (Caleman *et al.*, 2011). Most cascades are completed within tens of femtoseconds from the initial photoionization event, which is well within the pulse length. On a large scale, the electrons quickly equilibrate among themselves and will couple with the ions through collisions, leading to a hydrodynamic expansion of the molecule. In addition, the ions are expected to reconfigure on a local scale, which depends on the local atomic arrangement. Inhomogeneous distributions of atomic elements in molecules, in particular the presence of strongly absorbing Fe, is expected to significantly alter the local damage response.

Both global and local effects of radiation damage are clearly observed in the ferredoxin high-dose SFX data as expected given the experimental conditions of very high flux density, long FEL pulse length and the associated high dose absorbed by each microcrystal. Using plasma simulations (Scott, 2001; Caleman *et al.*, 2011), the underlying effects of ionization and displacement of the atoms, electrons and ions, respectively, and their influence on the global damage can be explored. Fig. 5(a) shows the average ionization for C, S and Fe as a function of time during the pulse, for photon energies below and above the Fe K-edge. The simulations show that all atoms rapidly lose electrons at the beginning of the exposure, reaching a high ionization state by the end of the 80 fs pulse. The average ionization of Fe below and above the absorption edge are similar (Fig. 5a). Considering only the direct photoionization this is difficult to explain, since the photon ionization cross sections differ significantly below and above the edge. Instead this suggests that the ionization of Fe is dominated by the electron impact ionization. To estimate how the rapid change in ionization will change the diffraction signal,
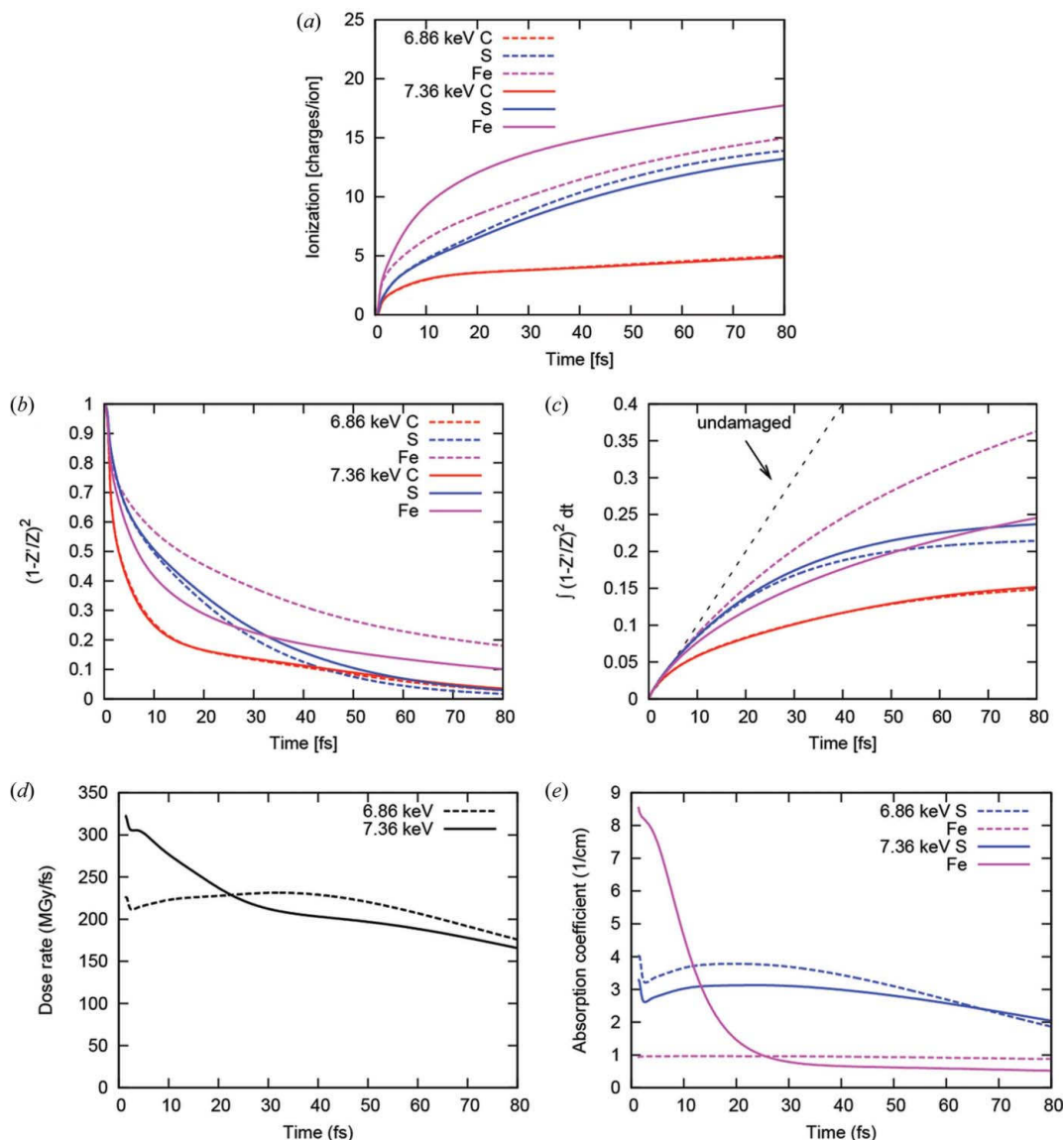
Figure 5
Ionization dynamics simulated with the plasma code *CRETIN*. (*a*) Average ionization per C, S and Fe atoms, as a function of time during the pulse, for two photon energies (6.86 keV and 7.36 keV). Ionization saturates faster for light atoms. For Fe it is dominated by collisional ionization. (*b*) Scattering power for C, S and Fe ions as a function of time, estimated from the number of bound electrons (without atomic form factors). (*c*) Expected accumulated signal during the pulse due to the loss of scattering power for C, S and Fe ions, compared with the neutral undamaged atoms (black dashed line). (*d*) Calculated dose rate (MGy fs^{-1}) in the sample as a function of time, for two photon energies. The dose rate decreases during exposure due to saturation in ionization and changes in the photoionization cross sections. (*e*) Absorption coefficient for Fe and S in the sample as a function of time. The drastic change in absorption for Fe above the K-edge in the first 20 fs is due to a rapid loss of electrons and a lowering of the ionization potential in the plasma environment.

the scattering was taken to be proportional to the squared number of electrons left in the atoms. Fig. 5 shows the instantaneous scattering power of each atomic species (C, S, Fe) as a function of time (Fig. 5*c*) and also the integrated signal as it accumulates during the pulse (Fig. 5*c*). The results suggest that most of the diffraction signal comes from the first 10–30 fs depending on the element, while the choice of wavelength only affects the contribution from the Fe.

Barty *et al.* (2012) describe the decay of the Bragg signal as a function of atomic displacements only, using a disorder factor (the root mean square displacement) derived from plasma calculations that approximately follows a $t^{3/2}$ time-

dependence. Based on the simulations (Fig. 5) described here, ionization is expected to play an important role in the present study and, as a result, a model based solely on atomic displacements should not reproduce the data, as indeed appears to be the case (Fig. S6). We have previously calculated atomic displacements for light atoms (Barty *et al.*, 2012; Chapman *et al.*, 2014) based on a model of ion diffusion in homogeneous solid density plasma. Chapman *et al.* (2014) show the calculated root mean square deviation (RMSD) of the C atoms as a function of dose. In our model, RMSD scales with $1/\sqrt{m}$ and the displacement for Fe is expected to be a factor of 2 lower than for C. This RMSD is found to scale

roughly with time as $t^{3/2}$ during the pulse, which implies that Fe ions will reach the same disorder as C delayed by a factor 1.5 in time. With this scaling, if C atoms reach 1 Å disorder in 20 fs for a dose of about 400 MGy (Chapman *et al.*, 2014) then Fe ions would reach the same 1 Å random displacement in 30 fs, which gives us a good estimation of how fast the disordering in the crystal occurs. This is due both to variations in atomic scattering factors from the randomly distributed ionized atoms and to random displacements of atoms. The measurements of Bragg counts are thus only sensitive to this initial portion of the pulse, before heavier elements have moved substantially.

Iron atoms are expected to ionize faster than the lighter elements under these experimental conditions (see. Fig. 5) and thus a decrease in the associated electron density of the iron atoms in the [4Fe–4S] clusters is expected (as measured in the time window before ionization-induced disorder sets in and as compared with the overall decrease in electron density). This is what is observed (Fig. 4). What is unexpected, however, is that the overall electron density of the two clusters differs in the SFX data (Fig. 1, Fig. 2, Table S10). At least for the high-dose SFX data (Fig. 1) this is not due to a difference in cluster incorporation during biosynthesis since the two clusters have the same electron density in the SLS data (Fig. 3, Table S10) which were collected using crystals grown from the same batch of protein. This would suggest that the difference in electron density is due to radiation damage. Given the same chemical composition of the two clusters and thus the same photo-absorption and collisional ionization probabilities it is very surprising to see that they differ in the extent to which they are damaged. Therefore, a difference was suspected in their immediate environment causing, for example, non-equal relaxation dynamics. However, this is not the case since the two binding sites evolved from a gene duplication event, making them extremely similar. There is one difference though, which is the mobile loop containing amino acids 26–28 (Fig. S11). It could not be modelled in the electron density of data collected from a room-temperature crystal [PDB code 1fdn (Duée *et al.*, 1994)], while two conformations were modelled for the data of a cryocooled crystal [PDB code 2fdn (Dauter *et al.*, 1997)]. Since this loop is in the vicinity of cluster 1, it is conceivable that it might influence the damage response. Another possibility is differences in the clusters themselves. Since the resolution of our structures is not high enough to address this question it is very fortunate that the ferredoxin structure had been determined to 0.94 Å resolution previously (Dauter *et al.*, 1997), and was published including a detailed comparison of the two clusters. Interestingly, significant differences in the stereochemistry of the two clusters were observed.

Similarly to most [4Fe–4S] clusters, ferredoxin's cluster 2 displays a tetragonal geometry compressed along one pseudo-fourfold axis, resulting in a local symmetry close to the D_{2d} point group (Dauter *et al.*, 1997). The compression axis passes through the (Fe2, Fe4, S1, S3) face of the cubane core containing the iron atoms coordinated by Cys40 and Cys18 on one side and the (Fe1, Fe3, S2, S4) face on the opposite side. In contrast, cluster 1 does not display any clear-cut symmetry, meaning that it could be conformationally strained. This strain may promote dissociation and thus disorder and/or charge migration and motion of atoms during the pulse, which might explain the higher loss of electron density in cluster 1. In contrast to the Bijvoet-averaged electron density of the clusters (Fig. 1, Table S10), the Bijvoet-difference Fourier (anomalous scattering density) map of the two clusters is similar (Fig. S9), suggesting that it contains mainly contributions from the very beginning of the pulse when both clusters were pristine. However, even with a large amount of ionization, and a large variation in ionization and electronic excited states of the iron atoms, the change in the atomic cross section is such that the anomalous term is not altogether obliterated. This was the prediction of Son *et al.* (2011*b*) who carried out detailed calculations of the resulting scattering factors under such conditions.

There are several effects that could be causing these observations concerning the [4Fe–4S] clusters. One cause might be an ultrafast charge rearrangement between the ionized iron and the coordinating sulfur. Such an effect has been observed recently in methylselenol molecules exposed to intense ($>10^{17}$ W cm^{-2}) ultra-short 2 keV FEL pulses. In that case an ultrafast charge redistribution to the inner-shell ionized selenium from the other atoms in the molecule takes place, as well as significant displacement of the atomic constituents upon multiple ionization during the pulse (Erk *et al.*, 2013). Detailed calculations are needed to explore this further.

4. Conclusions

It is well known in conventional crystallography that the sites of high-Z atoms are prone to higher levels of radiation damage. Our results here suggest that a similar situation is encountered in the high-intensity FEL regime. This regime is attractive for collecting data from yet smaller protein microcrystals, and for the possibility to take advantage of atom-specific changes to allow phasing crystallographic data (Son *et al.*, 2011*b*) in a somewhat analogous fashion to radiation-damage-induced phasing (Ravelli & McSweeney, 2000). However, to utilize these opportunities and optimize experimental conditions requires a detailed understanding of the processes taking place. The present results suggest that even small differences in metal coordination and environment may affect the extent of radiation damage in the high-intensity regime at FELs. While the conditions chosen for the experiment reported here are unlikely to be those chosen for realistic data collection schemes (given the very long pulse lengths), the results nevertheless demonstrate that FEL-induced radiation damage is far from being understood. The simulations indicate that pulse lengths of less than 20 fs are required to appreciably reduce damage effects, and that indeed most of the diffraction may come from the first 30 fs of longer pulses, depending on the pulse irradiance.

We cannot completely exclude the possibility that the results may in part be compromised by non-damage-related

differences between the FEL data and the synchrotron data used as a reference. Thus, while caution is appropriate, this study provides a first tantalizing glimpse of local damage processes on femtosecond timescales in the extremely high intensity regime. More experiments, systematically probing the FEL pulse length and fluence and thus dose and dose rate are therefore needed, as well as detailed calculations and new theoretical tools exploring the effects described here.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Andreasson, J. *et al.* (2011). *Phys. Rev. E*, **83**, 016403.
Barends, T. R. *et al.* (2013). *Acta Cryst.* D**69**, 838–842.
Barty, A. *et al.* (2012). *Nat. Photon.* **6**, 35–40.
Behrens, C., Decker, F.-J., Ding, Y., Dolgashev, V. A., Frisch, J., Huang, Z., Krejcik, P., Loos, H., Lutman, A., Maxwell, T. J., Turner, J., Wang, J., Wang, M.-H., Welch, J. & Wu, J. (2014). *Nat. Commun.* **5**, 3762.
Bergh, M., Timneanu, N., Hau-Riege, S. P. & Scott, H. A. (2008). *Phys. Rev. E*, **77**, 026404.
Botha, S., Nass, K., Barends, T., Kabsch, W., Latz, B., Dworkowski, F., Foucar, L., Panepucci, E., Wang, M., Shoeman, R. L., Schlichting, I. & Doak, R. B. (2015). *Acta Cryst.* D**71**, 387–397.
Boutet, S. *et al.* (2012). *Science*, **337**, 362–364.
Boutet, S. & Williams, G. J. (2010). *New J. Phys.* **12**, 035024.
Burmeister, W. P. (2000). *Acta Cryst.* D**56**, 328–341.
Caleman, C., Huldt, G., Maia, F. R. N. C., Ortiz, C., Parak, F. G., Hajdu, J., van der Spoel, D., Chapman, H. N. & Timneanu, N. (2011). *ACS Nano*, **5**, 139–146.
Caleman, C., Tîmneanu, N., Martin, A. V., Jönsson, H. O., Aquila, A., Barty, A., Scott, H., White, T. A. & Chapman, H. N. (2015). *Opt. Express*, **23**, 1213–1231.
Campbell, J. L. & Papp, T. (2001). *At. Data Nucl. Data Tables*, **77**, 1–56.
Chapman, H. N., Caleman, C. & Timneanu, N. (2014). *Philos. Trans. R. Soc. B*, **369**, 20130313.
Chapman, H. N. *et al.* (2011). *Nature (London)*, **470**, 73–77.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Corbett, M. C., Latimer, M. J., Poulos, T. L., Sevrioukova, I. F., Hodgson, K. O. & Hedman, B. (2007). *Acta Cryst.* D**63**, 951–960.
Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J. & Moulis, J. M. (1997). *Biochemistry*, **36**, 16065–16073.
DeLano, W. L. (2002). *The pyMOL Molecular Graphics System.* DeLano Scientific, Palo Alto, CA, USA.
DePonte, D. P., Weierstall, U., Schmidt, K., Warner, J., Starodub, D., Spence, J. C. H. & Doak, R. B. (2008). *J. Phys. D*, **41**, 195505.
Diederichs, K. (2006). *Acta Cryst.* D**62**, 96–101.
Duée, E. D., Fanchon, E., Vicat, J., Sieker, L. C., Meyer, J. & Moulis, J. M. (1994). *J. Mol. Biol.* **243**, 683–695.
Erk, B. *et al.* (2013). *Phys. Rev. Lett.* **110**, 053003.
Foucar, L. (2015). To be published.
Foucar, L., Barty, A., Coppola, N., Hartmann, R., Holl, P., Hoppe, U., Kassemeyer, S., Kimmel, N., Küpper, J., Scholz, M., Techert, S., White, T. A., Strüder, L. & Ullrich, J. (2012). *Comput. Phys. Commun.* **183**, 2207–2213.
Galli, L., Barends, T., Son, S.-K., White, T. A., Barty, A., Botha, S., Boutet, S., Caleman, C., Doak, R. B. M. N., Nass, K., Shoeman, R. L., Timneanu, N., Santra, R., Schlichting, I. & Chapman, H. N. (2015). *IUCrJ.* Submitted.
Hau-Riege, S. P. (2007). *Phys. Rev. A*, **76**, 042511.
Hau-Riege, S. P. & Bennion, B. J. (2015). *Phys. Rev. E.* Accepted.
Hau-Riege, S. P. *et al.* (2007). *Phys. Rev. Lett.* **98**, 145502.
Hirata, K. *et al.* (2014). *Nat. Methods*, **11**, 734–736.
Holton, J. M. & Frankel, K. A. (2010). *Acta Cryst.* D**66**, 393–408.
Hong, J. S. & Rabinowitz, J. C. (1970). *J. Biol. Chem.* **245**, 4982–4987.
Johansson, L. C. *et al.* (2013). *Nat. Commun.* **4**, 2911.
Jönsson, H. O., Tîmneanu, N., Östlin, C., Scott, H. A. & Caleman, C. (2015). *J. Synchrotron Rad.* **22**, 256–266.
Jurek, Z. & Faigel, G. (2009). *Europhys. Lett.* **86**, 68003.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kern, J. *et al.* (2013). *Science*, **340**, 491–495.
Kirian, R. A., Wang, X., Weierstall, U., Schmidt, K. E., Spence, J. C. H., Hunter, M., Fromme, P., White, T., Chapman, H. N. & Holton, J. (2010). *Opt. Express*, **18**, 5713–5723.
Kirian, R. A., White, T. A., Holton, J. M., Chapman, H. N., Fromme, P., Barty, A., Lomb, L., Aquila, A., Maia, F. R. N. C., Martin, A. V., Fromme, R., Wang, X., Hunter, M. S., Schmidt, K. E. & Spence, J. C. H. (2011). *Acta Cryst.* A**67**, 131–140.
Krumrey, M. & Ulm, G. (2001). *Nucl. Instrum. Methods Phys. Res. A*, **467–468**, 1175–1178.
Kühnel, K., Derat, E., Terner, J., Shaik, S. & Schlichting, I. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 99–104.
Liu, W. *et al.* (2013). *Science*, **342**, 1521–1524.
Lomb, L. *et al.* (2011). *Phys. Rev. B*, **82**, 214111.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature (London)*, **406**, 752–757.
Owen, R. L., Axford, D., Nettleship, J. E., Owens, R. J., Robinson, J. I., Morgan, A. W., Doré, A. S., Lebon, G., Tate, C. G., Fry, E. E., Ren, J., Stuart, D. I. & Evans, G. (2012). *Acta Cryst.* D**68**, 810–818.
Owen, R. L., Paterson, N., Axford, D., Aishima, J., Schulze-Briese, C., Ren, J., Fry, E. E., Stuart, D. I. & Evans, G. (2014). *Acta Cryst.* D**70**, 1248–1256.
Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4912–4917.
Philipp, H. T., Koerner, L. J., Hromalik, M. S., Tate, M. W. & Gruner, S. M. (2010). *IEEE Trans. Nucl. Sci.* **57**, 3795–3799.
Ravelli, R. B., Leiros, H. K., Pan, B., Caffrey, M. & McSweeney, S. (2003). *Structure*, **11**, 217–224.
Ravelli, R. B. G. & McSweeney, S. M. (2000). *Structure*, **8**, 315–328.

radiation damage

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Rubiano, J. G., Florido, R., Bowen, C., Lee, R. W. & Ralchenko, Y. (2007). *High. Energ. Density Phys.* **3**, 225–232.

Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringe, D., Petsko, G. A. & Sligar, S. G. (2000). *Science*, **287**, 1615–1622.

Scott, H. A. (2001). *J. Quant. Spectrosc. Radiat. Transfer*, **71**, 689–701.

Son, S.-K., Chapman, H. N. & Santra, R. (2011*b*). *Phys. Rev. Lett.* **107**, 218102.

Son, S.-K., Young, L. & Santra, R. (2011*a*). *Phys. Rev. A*, **83**, 033402.

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H. & Shen, J. R. (2015). *Nature (London)*, **517**, 99–103.

Weierstall, U., Spence, J. C. & Doak, R. B. (2012). *Rev. Sci. Instrum.* **83**, 035108.

Weik, M., Bergès, J., Raves, M. L., Gros, P., McSweeney, S., Silman, I., Sussman, J. L., Houée-Levin, C. & Ravelli, R. B. G. (2002). *J. Synchrotron Rad.* **9**, 342–346.

White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A. & Chapman, H. N. (2012). *J. Appl. Cryst.* **45**, 335–341.

Yano, J., Kern, J., Irrgang, K. D., Latimer, M. J., Bergmann, U., Glatzel, P., Pushkar, Y., Biesiadka, J., Loll, B., Sauer, K., Messinger, J., Zouni, A. & Yachandra, V. K. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 12047–12052.

Zeldin, O. B., Gerstel, M. & Garman, E. F. (2013). *J. Appl. Cryst.* **46**, 1225–1230.