

Structural knowledge or X-ray damage? Dose dependent case studies on xylose isomerase revealing structural perturbations.

Helena Taberman^{a, b}, Charles S. Bury^b, Mark J. van der Woerd^c, Edward H. Snell^{d, e*} and Elspeth F. Garman^{b*}

^aHelmholtz-Zentrum Berlin, Macromolecular Crystallography (HZB-MX), Albert-Einstein Straße 15, 12489 Berlin, Germany (current)

^bDepartment of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK

^cDepartment of Enterprise Technology Services, 2001 Capitol Avenue, Cheyenne, WY 82001, USA

^dHauptman-Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, NY 14203, USA

^eMaterials Design and Innovation, State University of New York at Buffalo, 700 Ellicott Street, Buffalo, NY 14203, USA

Correspondence email: elspeth.garman@bioch.ox.ac.uk, esnell@hwi.buffalo.edu

PDB ID's 6QRR, 6QRS, 6QRT, 6QRU, 6QRV, 6QRW, 6QRX, 6QRY

Synopsis Dose-dependent X-ray studies show a linear impact on global damage metrics with dose and specific changes in active site residues. Metal motion occurred, but was driven by unit cell expansion and not by direct radiation chemistry effects on the metal site.

Abstract Xylose isomerase (XI) is an industrially important metalloprotein studied for decades. Its reaction mechanism has been postulated to involve movement of the catalytic metal cofactor to several different conformations. Here, a dose dependent approach was used to investigate the radiation damage effects on XI and their potential influence on the reaction mechanism interpreted from the X-ray derived structures. Radiation damage is still one of the major challenges of X-ray diffraction experiments and causes both global and site-specific damage. In this study, consecutive high-resolution data sets from a single XI crystal from the same wedge were collected at 100 K and the progression of radiation damage was tracked over increasing dose (0.13-3.88 MGy). The catalytic metal and its surrounding amino acid environment experience a build-up of free radicals, and the results show radiation damage induced structural perturbations ranging from an absolute metal positional shift to specific residue motions in the active site. The metal shift is an artefact of global damage and the resulting unit cell expansion, but residue motion appears to be driven by the dose. Understanding and identifying radiation-induced damage is an important factor in accurately interpreting the biological conclusions being drawn.

Keywords: metalloprotein; enzyme; radiation damage; specific structural damage; xylose isomerase

1. Introduction

During the process of X-ray diffraction, the crystals used in the experiment also absorb X-ray radiation, causing primary damage and also generating free radicals, which propagate chemical damage through radiation chemistry pathways (Nave, 1995; O'Neill *et al.*, 2002). This is an inherent problem in macromolecular crystallography (MX) even when the samples are held at cryogenic temperatures (~100 K) (Holton, 2009; Garman, 2010). These radicals cause structural perturbations that result in both global and site-specific effects. Global effects can be observed during data collection, for example as a loss of diffraction intensity, increased Wilson *B*-factors and swelling of the unit-cell volume causing increasing non-isomorphism (Ravelli & McSweeney, 2000). In contrast, specific effects can often be seen in electron density maps after refinement and can occur at less than 60 times lower dose before any signs of radiation damage are observed on the diffraction images or are seen from the decreasing data quality (Holton, 2009). Site-specific damage can be detected, for example in the photoreduction of metal centres, as elongation and breakage of disulfide bonds, and decarboxylation of acidic side chains (Ravelli & McSweeney, 2000; Carugo & Carugo, 2005; Weik *et al.*, 2000; Burmeister, 2000). These can distort the biological conclusions drawn from the resultant model (Grabolle *et al.*, 2006) and the various radiation damage effects can also hinder efficient structure solution.

The metal centres of metalloproteins, which constitute 25-50% of all proteins found within an organism (Bowman *et al.*, 2016), have a very large X-ray absorption cross-section (Murray & Garman, 2002) and are highly electron-affinic (Williams, 1985), increasing the probability that the radiation and its subsequent secondary products will affect them. Hence, the metal ion and its surrounding amino acid environment experience a localised build-up of free radicals, which makes them more susceptible to damage than the rest of the protein. This causes, for example, changes in the redox state of the metal sites, which can compromise the active site structure and again have an effect on the biological conclusions being drawn (Corbett *et al.*, 2007).

The 43 kDa metalloprotein, xylose isomerase (XI, EC 5.3.1.5), also referred to as glucose isomerase, holds enormous potential for applications in the pharmaceutical industry and in biotechnology for converting biomass into fuels and chemicals. It is an important enzyme used extensively in the food industry in large-scale production of starch-based high-fructose corn syrup (Bhosale *et al.*, 1996; Pedersen, 1993). XI catalyses the reversible isomerization of a broad range of aldose sugars to their corresponding ketoses using two divalent metal cofactors (designated M1 and M2), such as, Mg^{2+} , Mn^{2+} or Co^{2+} (Carrell *et al.*, 1984; Callens *et al.*, 1986; Jokela *et al.*, 2002), of which XI has the highest affinity for manganese ($K_d = 2.7 \times 10^{-5}$ M) (Schray & Mildvan, 1972). It is thus of great interest to understand the reaction mechanism and thereby potentially increase its suitability for different industrial processes, for example by improving its activity, specificity, thermal stability, and optimum pH.

The exact reaction mechanism of XI has long been a subject of debate fed by increasingly detailed studies, many discussing the mobility of the catalytic metal ion (M2) (Collyer *et al.*, 1990; Jenkins *et al.*, 1992; Allen *et al.*, 1994; Lavie *et al.*, 1994; Whitaker *et al.*, 1995; Asboth & Naray-Szabo, 2000; Fenn *et al.*, 2004; Lee *et al.*, 2017). The overall mechanism is thought to include three major steps: ring opening, isomerisation, and ring closure. The binding of the cyclic α -anomer carbohydrate substrate, including the role of the stabilizing structural metal (M1), is well defined (Asboth & Naray-Szabo, 2000; Blow *et al.*, 1992). However, the subsequent steps in the reaction mechanism, during which the catalytic metal ion M2 is said to adopt as many as three different alternate positions, M2a, M2b, and M2c, remain unclear. Fig. 1 shows a schematic representation of the reaction mechanism and putative role of the alternate M2 positions. Recent X-ray and neutron diffraction studies have established that the ring opening and linearization proceeds through a proton transfer reaction in which the doubly protonated His54 in the active site acts as a proton donor and Lys289 as a proton acceptor, and presumably M2 moves from position M2a to position M2b to bind the linear substrate (Fenn *et al.*, 2004; Kovalevsky *et al.*, 2010; Munshi *et al.*, 2014). The subsequent isomerisation step then occurs through hydride shift from a catalytic water mediated by the essential M2 cofactor, position M2c (Lavie *et al.*, 1994; Whitaker *et al.*, 1995; Fenn *et al.*, 2004; Kovalevsky *et al.*, 2010; Whitlow *et al.*, 1991), after which the bonds to M2 are lost. The acyclic product is then stabilised by a ring closure reaction (Bhosale *et al.*, 1996).

Currently, there are 156 structural models of XI, determined by X-ray or neutron diffraction, or by a combination of both, deposited in the Protein Data Bank (Berman *et al.*, 2000). There is no consensus between the structures regarding the conformation(s) of M2 or its importance and role in the reaction mechanism. In general, neutron diffraction experiments, although limited by the neutron fluxes available, complement X-ray diffraction by being more sensitive to the visualisation of hydrogen atoms (or their deuterium isotopes) in structures, which leads to more detailed definitions of enzymatic reaction mechanisms. Unfortunately, the neutron scattering cross-section of some metal atoms, for example Mg and Mn, is small and they scatter neutrons more weakly than X-rays. Furthermore, the neutron scattering length of the Mg (5.375 fm) or Mn (3.73 fm) atoms is similar to that of a single D (6.61 fm) or H (3.74 fm) atom, making it challenging to distinguish the metals explicitly from water molecules in mobile regions of a protein. Hence, X-ray diffraction studies of metalloproteins are crucial to provide structural and mechanistic detail, even though metal atoms and their immediate environment are highly sensitive to X-ray radiation damage.

In this study, we report the effects of dose dependent radiation damage on XI and examine their impact if any, on identifying the precise reaction mechanism. Fenn *et al.*, (2004), review evidence suggesting mobility of the catalytic metal cofactor, M2. Given the radiation sensitivity of metal sites, we explore this possibility in light of the dose dependent effects we observe. The results highlight the

need to be aware of the radiation damage inflicted during structural studies of metalloproteins, and to understand how global and specific effects may impact the interpretations from resulting models.

2. Materials and methods

2.1. Protein purification and crystallisation

The enzyme used in this study was the thermostable E186Q mutant of *Streptomyces rubiginosus* XI previously supplied by Genencor International (Palo Alto, California, USA) and routinely used in most XI studies. The protein was purified as described previously (Meilleur *et al.*, 2006; Snell *et al.*, 2006). Briefly, the commercial XI was purified with gel-filtration chromatography using a HiLoad 16/600 Superdex 200 column and finally dialysed in 50 mM magnesium sulfate. The purified enzyme was stored at 22°C as a dilute solution (~3 mg/ml) and concentrated to 100 mg/ml immediately before setting up the crystallisation experiments.

The protein was transported to Stanford Synchrotron Radiation Lightsource (SSRL, Stanford, USA) where crystals were grown at the synchrotron by the hanging drop vapour diffusion crystallisation method. Concentrated protein (5 µl at 100 mg/ml) was mixed with an equal volume of reservoir solution and equilibrated against 1 ml of precipitant solution. The successful crystallisation conditions were 8 % (v/v) 2-propanol, 25 % (v/v) ethylene glycol, 50 mM HEPES pH 7.0 and 50 mM magnesium chloride. Orthorhombic crystals grew in less than a day at room temperature (22°C) to dimensions of $0.3 \times 0.3 \times 0.3$ mm³. Inclusion of additional cryoprotectant was not necessary, since the ethylene glycol concentration was sufficient to avoid any ice formation on cooling, and after harvesting with nylon cryo-loops, the crystals were flash-cooled to 100 K directly into the nitrogen cryostream for data collection using an Oxford Instruments CryojetXL (Oxford Instruments, Abingdon, UK).

2.2. Diffraction data collection, processing, structure solution and refinement

X-ray diffraction data were collected at SSRL on beamline 11-1 using an ADSC Quantum 315 CCD detector. Seven consecutive data sets were collected from the same crystal, each starting at the same orientation and position as the first to ensure that the same volume of the crystal was irradiated during each data set. A complete data set was first collected to 1.21 Å resolution (largest inscribed circle on detector) at an X-ray energy of 13.0 keV (0.954 Å wavelength) over 90° of rotation, with a crystal to detector distance of 150 mm, 0.5° oscillation angle and 2 s exposure time per image. An identical 10° high-resolution swathe was collected up to 0.89 Å resolution (largest inscribed circle) between each complete data set on the first 10° of the complete 90° wedge at an X-ray energy of 14.5

keV (wavelength of 0.855 Å) with a crystal to detector distance of 100 mm, and 0.5° oscillation angle and 35 s exposure time per image to monitor the decay of the high-resolution (HR) reflections and to increase the absorbed X-ray dose between the seven complete data sets. While the expected intensity decay was visually observed in the high-resolution swathes, the completeness was approximately 17 %, and thus our study focuses only on the complete 90° data sets and the structural models obtained from them. All data sets were processed with XDS (Kabsch, 2010), which was also used to generate a random 5% R_{free} test set of the unmerged reflections for the first low-dose data set.

The 0.86 Å resolution structure of XI from *Streptomyces olivochromogenes* (PDB code: 1MUW) (Fenn et al., 2002), which has 95% sequence identity with the *S. rubiginosus* XI E186Q mutant, was used as a starting model. The lowest dose data set was used with this starting model and with ARP/wARP (Perrakis et al., 1999; Langer et al., 2008) to produce a new starting model minimizing any phase bias. The structure was initially refined with this starting model using REFMAC5 (Murshudov et al., 2001) followed by phenix.refine (Afonine et al., 2012) from the PHENIX software suite (Adams et al., 2010) and finalised with SHELXL (Sheldrick, 2015). The pdb file from phenix.refine was converted into an ins file with PDBINS. The occupancy and anisotropic B-factors of all atoms were refined.

In between the refinement steps, manual rebuilding of the model took place using COOT (Emsley et al., 2010). The *hkl* data file used for the SHELXL refinement was converted to an mtz file with FREERFLAG from the CCP4 and the existing R_{free} test set was retained. The same test set was also copied to the higher dose data sets when scaling them with AIMLESS (Evans, 2006; Evans, 2011; Evans & Murshudov, 2013) and truncating them all to 1.17 Å resolution. For each of the six subsequent higher dose data sets and the overall summed data set (all seven data sets merged), individual structures were derived using *phenix.refine* with the structure obtained above for the first (low-dose) data set coupled with the observed amplitudes of each later data set. The structures were validated with the integrated MolProbity (Chen et al., 2010) and other structure validation tools within *phenix.refine* and COOT, and using PDB_REDO (Joosten et al., 2014) and the wwPDB Validation Server (Berman et al., 2003). The processing and refinement statistics for each data set are summarized in Table 1. The atomic coordinates and the structure factors for the resulting eight structures have been deposited in the wwPDB (Berman et al., 2003; Gutmanas et al., 2014).

2.3. Dose calculation

RADDOSE-3D (Zeldin et al., 2013b; Bury et al., 2018) was used to calculate the average diffraction weighted doses (DWDs, Zeldin et al., 2013a) for the data sets against which to plot the radiation damage metrics. For the dose calculations, knowledge of the beam (energy, flux, size, and profile) along with prior knowledge of the crystal (size, shape, orientation relative to the incident

beam, number of heavy atoms in solvent and protein, unit cell parameters and the number of amino acid residues) and data collection specifics (oscillation range and exposure time) are required.

The beam shape was determined by 2D fluorescence scanning of a small cobalt crystal ($\sim 10\ \mu\text{m}$) across the full beam, and was found to be top-hat in the horizontal direction and Gaussian in the vertical direction with a FWHM of $200\ \mu\text{m}$. Since a mixture of beam shapes cannot be used in RADDPOSE-3D calculations, the beam profile was approximated as a Gaussian with a very large FWHM value of $2000\ \mu\text{m}$ for the horizontal direction to simulate a top-hat shaped beam.

The beam fluxes were measured through a $200 \times 200\ \mu\text{m}^2$ aperture and were calculated from the ion chamber readings. The same aperture was used for the data collection. The flux was 2.2×10^{11} photons/s (before taking account of the 40 % transmission used which reduced it to 8.8×10^{10} photons/s) for the complete data sets at 13 keV and 1.84×10^{11} photons/s (before taking account of the 48 % transmission used which also gave 8.8×10^{10} photons/s) for the HR swathes at 14.5 keV.

The average DWDs for each complete data set were defined in a single RADDPOSE-3D calculation so that account was taken of accumulation of DWD for the particular wedge used in the experiment. The input file contained all the complete data set wedges each separated by an HR “burn” swathe giving a total accumulated average absorbed dose of 3.88 MGy by the end of the highest dose data set. The calculated accumulated average DWDs for the seven complete data sets were 0.13, 0.76, 1.38, 2.01, 2.63, 3.25, and 3.88 MGy. The metrics used to monitor radiation damage progression were plotted against the resulting DWD values.

2.4. MicroPIXE measurements

Since knowledge of the atomic composition of a protein is a prerequisite for the accurate calculation of the absorbed dose, the metal content of the XI sample was measured using microbeam particle induced X-ray emission (microPIXE), an established technique for such analyses (Garman & Grime, 2005). In microPIXE, the characteristic X-ray emission spectra of heavy atoms ($Z > 10$) in a protein sample are measured, which allows the stoichiometric ratios of these elements to be defined using the sulphur peak from the methionines and cysteines as an internal standard (Garman, 1999). A dried protein sample (approx. volume of solution $0.1\ \mu\text{l}$) was analysed in vacuum with a highly focused ($\sim 1.3\ \mu\text{m}$ diameter) 2.5 MeV ($\sim 300\ \text{pA}$ current) proton beam. The measurements were carried out at the Ion Beam Centre, University of Surrey, UK (Grime *et al.*, 1991).

Firstly, a scan over the dried drop of $750 \times 750\ \mu\text{m}^2$ was collected to determine the position of the protein sample. Secondly, two point spectra were recorded from different places on the sample. The sample thickness and gross matrix composition were defined by detecting the Rutherford backscattered protons, after which the concentrations of the elements in the sample were extracted

from the X-ray emission spectrum with GUPIX (Maxwell *et al.*, 1995) within DAN32 (Grime, 1996) to extract the relative amount of each element, particularly manganese, in the sample.

Unfortunately, the XI E186Q sample used in the measurements was dissolved in a 0.9 M $(\text{NH}_4)_2\text{SO}_4 + 1 \text{ mM MgSO}_4$ solution, which did not allow determination of the stoichiometric ratios of the elements due to the high extra sulphur concentration in the buffer. However, qualitatively from the microPIXE spectra we were able to determine the presence of manganese in the sample, even though no manganese was present in the reagents used in the purification or crystallisation. No other metals heavier than sulphur were present. For the experimental microPIXE setup used here, magnesium was outside the detectable range of the spectrum, but the technique has the capability to detect it if a thinner beryllium filter on the Li(Si) X-ray detector is inserted. The unambiguous atomic signature from microPIXE allowed the XI metal sites to be modelled with manganese. The PIXE spectrum is shown in Fig. 2.

3. Results

3.1. Global radiation damage

The data sets were analysed for global radiation damage and different metrics were plotted against the increasing absorbed average DWD. The data quality decreases strikingly linearly with dose for every metric shown. The results are presented in Fig. 3 and show the unit cell volume, the relative B -factor ($B_{\text{rel}} = B_n - B_1$, Kmetko *et al.*, 2006) and the R_{meas} increasing with dose, and $CC_{1/2}$ and the relative diffraction intensity I_n/I_1 decreasing with dose. Here, I_n is the summed total intensity for a complete data set and I_1 is the total intensity of the first low-dose data set. The total intensities were calculated from the AIMLESS output files as the sum of the measured reflections (N_{meas}) multiplied by the average intensities ($\langle I \rangle$) for each resolution bin.

The coefficient of sensitivity ($s_{AD} = AD / (8\pi^2 \Delta B_{\text{rel}})$ where D is dose) for the XI crystal can be calculated from the gradient of the B_{rel} results, and is $0.007 \text{ \AA}^2 \text{ MGy}^{-1}$, compared with an average value of $0.012 \text{ \AA}^2 \text{ MGy}^{-1}$ determined for four protein crystals at 100 K (lysozyme, thaumatin, catalase and apoferritin, Kmetko *et al.*, 2006). This value implies that the XI crystal has comparatively low sensitivity to global damage. A possible explanation of this is that there are no disulfide bonds within the structure, which are usually the most radiation sensitive species. However, the value for $D_{1/2}$, the dose to half diffraction intensity, obtained by extrapolation of the results shown in Fig 3(e) is approximately 8 MGy which is lower than is generally observed for proteins. Damage to residues involved in intermolecular crystal contacts can drive the collapse of the lattice and thus a decrease in diffraction intensity, as has been observed previously for Arg-Asp contacts (Murray *et al.*, 2005).

3.2. Structure determination

The XI E186Q data sets were indexed in the orthorhombic space group $I222$ with unit cell parameters of $a = 92.5 \text{ \AA}$, $b = 97.5 \text{ \AA}$, $c = 102.5 \text{ \AA}$ for the first (lowest dose) data set. The calculated Matthews coefficient, V_M , was determined to be $2.7 \text{ \AA}^3/\text{Da}$, which corresponded to one monomer in the asymmetric unit and an estimated solvent content, V_S , of 54 % (Matthews, 1968; Kantardjieff & Rupp, 2003). The XI monomer consists of 388 amino acid residues and all but the N-terminal methionine were present in the observed $2F_{obs} - F_{calc}$ electron density map. The final model was thus composed of 387 amino acid residues and 12 ethylene glycol, one 2-propanol, and 528 water molecules, and two ions each of manganese, magnesium and sodium. Many of the amino acid residues, especially at the active site, for example the metal binding Glu217, His220, Asp255, and Asp257, had alternate conformations as described in detail below.

The structure obtained from the first (lowest dose) data set was refined to R_{work}/R_{free} values of 12.1 %/14.7 % respectively. The M1 site was refined partly with an Mn^{2+} ion (occupancy 0.54) and partly a Mg^{2+} ion (occupancy 0.46) as in Munshi *et al.*, 2014. This provided a good fit to the density. The presence of manganese and lack of any other heavy metal species (except the magnesium present in the dialysis buffer and crystallisation solution) was shown with microPIXE. The Mn^{2+} at the M2 site was refined in three different conformations, M2a, M2b, and M2c (Fig. 1 and 4(b)) as found in the structure used as the model building template (PDB code: 1MUW, Fenn *et al.*, 2002).

The alternate metal sites were also associated with two different conformations of residues in the active site, in particular Glu217, His220, Asp255, and Asp257, which are denoted A and B. The seven consecutive complete data sets ($n = 1$ to 7) were truncated to the same resolution of 1.17 \AA in AIMLESS, because of the fall-off in resolution of the highest dose data set. The same R_{free} test set from the lowest dose data set was used for the higher dose data sets. The refinement statistics for the first low dose set and the following six data sets are shown in Table 1.

Following inspection of the difference maps as a function of dose, metal bond distance analysis was carried out using COOT, since subtle movements in active site residues were observed which were plotted against dose (see section 3.3).

The resulting XI monomer model has a bi-domain structure, as was expected from previously deposited models (Fig. 4(a)). It is formed of an N-terminal parallel $(\alpha/\beta)_8$ -barrel fold (residues 2-322) and a long C-terminal loop (residues 323-387) with five α -helical segments. The active site is located at the C-terminal side of the β -barrel where both of the metal cofactors are bound in octahedral coordination in a crevice formed between the β -sheets β_6 , β_7 and β_8 . XI forms a dimer by packing the active sites of two monomers together so that the barrels are coaxial. Homotetrameric quaternary structure is then formed as a dimer of dimers, in which the C-terminal extensions wrap around the adjacent monomers. Both, the dimer and the tetramer have been shown to be biologically active

multimers (Rangarajan *et al.*, 1992). Fig. 4(b) presents a close-up of the active site and the refined metal positions with their $2F_{obs}-F_{calc}$ electron density map contoured at 1σ ($0.41 \text{ e}^-/\text{\AA}^3$).

3.3. Specific damage observations

B_{net} values for the seven structures were calculated with the RABDAM program (Shelley *et al.*, 2018). B_{net} is a derivative of the B_{damage} metric (Gerstel *et al.*, 2015), which is derived from the atomic B -factor corrected for packing density. It summarises in a single value the total extent of site-specific radiation damage from the coordinate model associated with a single PDB entry. A linear increase in B_{net} for the structures with increasing dose was observed (Fig. 5), corresponding with the global damage effects seen in Fig. 3.

The site-specific radiation damage was analysed using the Radiation-Induced Density Loss program, RIDL (Bury *et al.*, 2017; Bury & Garman, 2018). RIDL calculates per-atom metrics to describe changes in electron density between structures derived from consecutive data sets. Here the D_{loss} metric, calculated as the maximum electron density loss compared with the lowest dose structure in the local region around each atom in XI for each higher data set's $F_{obs,n} - F_{obs,1}$ Fourier difference map, has been used as a per-atom indicator of site-specific damage with increasing dose.

In Fig. 6 the top 25 D_{loss} sites calculated by RIDL are presented as red spheres, with the radius of each sphere being proportional to the magnitude of D_{loss} . Fig. 6(a) and 6(b) present the top 25 specific damage sites identified by the $F_{obs,2} - F_{obs,1}$ and $F_{obs,7} - F_{obs,1}$ maps respectively. The increase in damage with dose, at the active site and the surroundings of the catalytic metals, can clearly be seen. For $F_{obs,2} - F_{obs,1}$ (see Fig. 6(a)), of the 25 most damaged residues, the greatest indication of damage was associated with the M2 site, with the majority of the other sites shown related to Met residues and those residues having multiple conformations.

For the 25 most damaged damage residues seen in the $F_{obs,7} - F_{obs,1}$ map (Fig. 6(b)), the sites are those expected to be damaged, with the greatest damage again appearing at the active site. For multiple conformations, RIDL removes all but the highest occupancy conformation and performs its analysis on that conformation. Because two conformations of the active site exist, the lower occupancy B conformation was examined by removing the A conformation before running RIDL a second time. The D_{loss} metric in RIDL did not appear sensitive to the very small structural changes in the active site and so the D_{loss} analysis was only performed for the highest occupancy state.

The top locations of site-specific damage for the $F_{obs,7} - F_{obs,1}$ map include the catalytic metal, the sulphurs of methionine side chains, the oxygen atoms of aspartic and glutamic acids around the active site and the multimer forming surfaces of XI. Furthermore, oxygen atoms of water molecules, ethylene glycols, and the side chains of glutamine, serine, and threonine appear on the list of damaged

sites. XI has no disulfide bonds. Even for radiation-insensitive atoms the damage is observed to increase with dose due to the increasing noise in $F_{obs,n} - F_{obs,l}$ map caused by the decreasing diffraction data quality.

Isomorphous difference Fourier maps (calculated by RIDL) for the active site are presented for $F_{obs,2} - F_{obs,l}$ in Fig. 7(a) and 7(b) (negative/positive difference density respectively), and for $F_{obs,7} - F_{obs,l}$ in Fig. 7(c) and 7(d) (negative/positive difference density respectively). Both positive and negative difference density is observed around the three M2 metal sites and clear negative density on the oxygen atoms of the glutamates and aspartates, in line with usual observations of specific damage in MX. The single conformation of the residues in the active site shows the A conformation with the lower occupancy B conformation being removed as part of the RIDL processing.

For the lowest dose interval, the $F_{obs,2} - F_{obs,l}$ electron density data, Fig. 7(a) and 7(b) show electron density loss around Glu217 (OE2 and CD atoms), as well as Asp257 (OD1 and OD2 atoms). This corresponds to the decarboxylation of the acidic side chains noted in many studies on radiation damage. There is no corresponding positive density for the side chains, largely ruling out any coordinated shift in atom positions. The catalytic metal M2 in position M2a shows a reduction in electron density, and a gain around positions M2b and M2c. This is not reflected in an occupancy change (see Table 2). There is loss of electron density around the O1 and O2 atoms of ethylene glycol and the water molecule next to M2a. For the highest dose interval, the $F_{obs,7} - F_{obs,l}$ electron density maps in Fig. 7(c) and 7(d) show both local gain and loss of density. Both OD1 and OD2 atoms of Asp245, Asp255, and Asp287 see a decrease in density, as well as OE2 of Glu181 and OE1 of Glu217, and the bond between His220 NE2 and CE1 atoms. It is also noteworthy that the M1 position and the water molecule adjacent to it lose electron density.

As detailed in Table 2, only slight changes are seen in the occupancies of the alternate conformations of the catalytic metal M2 and the active site residues, indicating that the changes in the electron density as a function of dose are not a consequence of conformational changes. Analysis of the M2 isotropic B -factor shows an increase for M2a from 6.7 to 14.6 Å², for M2b from 7.2 to 12.1 Å² and M2c a decrease from 10.4 to 7.7 Å². This suggests disorder increasing in the M2a and M2b sites with order in the M2c site improving with dose.

No movement in metal position is observed as a function of dose. Table 3 shows changes in the absolute position of the M2 metal site. M2a undergoes a shift of 0.17 Å, M2b of 0.08 Å and M2c of 0.10 Å. However, when the absolute coordinates are converted to fractional coordinates the changes in metal positions are negligible from lowest to highest dose, being 0.008, 0.004 and 0.004 for M2a, M2b, and M2c respectively, on a normalized scale. These absolute changes are compatible with the 0.13 Å, 0.18 Å and 0.14 Å expansion in the a , b , and c parameters of the unit cell over a dose of 3.88

MGy and might partly explain the electron density loss seen around M2a and gain around M2b and M2c (see Fig. 7, produced without consideration of the cell parameter change).

Residues associated with the M2 metal position are Glu217, His220, Asp255, and Asp257. M2a, M2b, and M2c maintain a fairly constant bond distance to Glu217 in both the A and B conformations. The metal occupancies did not vary significantly as a function of dose with M2a, M2b and M2c having average occupancies of 0.42, 0.53 and 0.05 respectively and standard deviations over the set of 0.02, 0.02 and 0.01. Similarly, both the A and B conformations of Glu217, His220, Asp255 and Asp277 maintained fairly constant occupancies during the dose series. For conformation A, this was 0.14, 0.33, 0.38, and 0.36 and for conformation B, 0.86, 0.67, 0.62 and 0.64 respectively. Glu217 occupancy in the B conformation is closely related to the sum of the M2a and M2b occupancies, whereas those of conformation A for His220, Asp255 and Asp257 closely match that of M2a, and B conformations of the same residues are a closer match to the M2b occupancy. Glu217 in the A conformation seems more closely associated with the occupancy of M2c. This can be seen in Fig. 8. The residues involved in the proton transfer reaction, His54 and Lys289 are not bonded to either M1 or M2.

Metal bonding distances are shown in Table 4 with their standard deviations for high-resolution protein structures (Zheng *et al.*, 2008), as well as the error in the atomic coordinate (Kumar *et al.*, 2015) for the bonding atoms. Bond distance decreases from the lowest to highest dose for Asp225 OD2 conformation A to M2a, but no significant change is seen for conformation B to M2b or M2c. The Asp257 bond distance to M2a, M2b and M2c is fairly stable for both A and B conformations. It can be seen visually that M2a matches conformation A and M2b matches conformation B. His220 appears to shift closer to M2b and closest to M2c in the A conformation.

Active site residue atom distances are plotted as a change in length as a function of dose in Fig. 9. The figure breaks down the distances to M2a, M2b and M2c and residue A and B conformations. Fig. 9(a) shows that Asp255 OD2 increases in distance from M2a while OD1 shifts toward M2a in conformation A. This is a subtle change which appears to be due to rotation of the terminal part of the residue. A more pronounced change is evident for His220 NE2 which shifts almost 0.3 Å towards M2a. Given that the unit cell is expanding as a function of dose, any motion that reduces distances is likely to be a result of specific radiation chemistry effects rather than of global damage. Glu217 and Asp257 show an increase in distance from M2a in conformation A. For conformation B, (see Fig. 9(b)), the distance between His220 and M2a is still decreasing with dose, whereas Asp257 OD1 is increasing in distance and all other residue atoms remain at a fairly constant distance.

For M2b (see Fig. 9(c)), there are small changes seen for conformation A, but less than the almost 0.3 Å for His220. The results are similar for M2c and conformation A, (see Fig. 9e). For M2b and M2c, there are minimal distance changes for residue conformation B (see Figs. 9(d) and (f)). The

structural model is an average of all the unit cells in the crystal and the occupancy table thus describes the overall composition of that crystal. The distance and occupancy results suggest that the A residue conformation is driven to the B conformation by dose and that the B conformation is associated with M2b and M2c. Glu217 distances appear to be constant in both the A and B conformations. M2a is not bonded to His220 and has a distance of 3.46 Å at the highest dose but is bonded to M2b in the B conformation and very closely bonded to M2c. H220 NE2 shows an almost linear decrease in distance to M2a with dose.

RIDL also outputs a metric called D_{neg} which is essentially a measure of the average $F_{obs,n} - F_{obs,l}$ negative density surrounding each atom, where the contribution of each voxel's value to the D_{neg} metric is weighted by the relative proportion of actual electron density corresponding to the atom at that position. This provides an indication of the magnitude of electron density lost with increasing dose for each atom (similar to the D_{loss} metric) but it up-weights negative density changes close the centre of mass of each atom and down-weights negative density changes further from the centre of mass of each atom. As such, it is more sensitive to the small spatial changes indicated for active site residue conformations A and B.

Fig. 10 shows D_{neg} for the residues associated with M2 for both the A (Fig. 10(a)) and B (Fig. 10(b)) conformations. The damage seems similar with the only noticeable difference being the increased sensitivity of Asp257 OD2 in the B conformation. Asp257 shows a significant twist in the B conformation and this damage indicator may explain that twist. His220 is twice as sensitive to damage in the A conformation than when in the B position. Glu217 and Asp255 show similar sensitivity in both the A and B conformations. The OD2 of Asp257 is about three times as sensitive in the B conformation as the A conformation. For comparison, in Fig. 10(c) Asp80, one of the residues involved in the tetramer crystal contacts is shown on the same scale. The main chain atoms show little damage and the side chain shows an almost linear effect with dose.

4. Discussion

As the results presented above indicate, both the global and site-specific radiation damage to the XI crystal progress linearly with dose over the dose range studied here (0.13 – 3.88 MGy). Considering that the highest dose is only 3.88 MGy, the site-specific damage is significant: the difference density map calculated for the lowest dose data sets $F_{obs,2} - F_{obs,1}$ (Figs. 7(a) and (b)) already shows negative difference density around the M2a position, acidic side chains, and water and ethylene glycol molecules at the active site. The difference density significantly increases with dose (Fig. 7) showing clearly, for example, the decarboxylation of acidic side chains taking place around Asp245, Asp255, Asp257, Asp287 and Glu181.

The increase in site-specific damage is observable from inspection of the top 25 damage sites (Fig. 6), despite the fact that the dose is only 3.88 MGy. This is far lower than the experimental radiation dose limit of 30 MGy determined for cryo-cooled protein crystals (Owen *et al.*, 2006). XI is a tetramer in the unit cell with Asp80, Thr82, Arg331, Asp336 and Arg340 forming crystal contacts between the separate tetramers. The crystal contacts associated with Asp may increase the sensitivity of the crystal lattice to damage, although the trends seen in Fig. 10(c) suggests that this damage takes place at a lower rate than that occurring in the active site.

The detailed analysis of the changes XI active site with dose presented above raise two pertinent questions in relation to the interpretation of active site changes during the mechanistic and structural analysis of metalloproteins.

Firstly, several papers in the literature report a movement of the XI M2 metal positions which was hypothesised to be associated with the catalysis mechanism (Collyer *et al.*, 1990; Jenkins *et al.*, 1992; Bogumil *et al.*, 1997; Asbóth & Náray-Szabó, 2000; Lavie *et al.*, 1994; Fenn *et al.*, 2004). Here, we indeed observed an absolute movement of the metals with increasing dose. However, once the radiation induced unit cell expansion has been taken into account by conversion of the metal positions to fractional coordinates, this disappears within error. We thus conclude that the metals do not shift.

Note that in our work, the M2 site occupancies do not change with dose either, and that our values are at variance with those of Fenn *et al.* (2002) where for the ligand free model 1MUW, M2a, M2b and M2c have occupancies of 0.69, 0.48, 0.16 compared to those here of 0.42, 0.53 and 0.05. In other XI structural models, there are occupancies of less than unity for the M2 position. It is possible that lack of knowledge of the exact identity of the metal or metals in the M2 site may explain some of these results or that the reduction of the metal by free radicals produced from X-ray irradiation could impact occupancy calculations. Given the importance of the metal site in the enzymatic mechanism and the observation that many structures have over-occupied or under-occupied states, accurate knowledge of both the identity of the metal and radiation chemistry effects is a critical element of understanding mechanism.

Secondly, although the unit cell volume expands with dose, we observe that two distances to the M2a metal site decrease by small but significant amounts, and that they behave differently for conformations A and B: those of His220 NE2 (3.73 Å to 3.46 Å) and of Asp255 OD2 to (1.98 Å to 1.85 Å), see Fig. 8. Asp255 OD2 is explained by a rotation of the residue but His220 shifts toward M2a. The D_{neg} metric from RIDL is sensitive to the His220 and Asp255 rotation. The bond lengths for other residues in the active site remain reasonable static with the possible exception of His220 to M2b with a small reduction going from 2.29 Å to 2.22 Å.

A study of four active site mutants of the His220 residue (His220Ser, His220Glu, His220Asn and His220Lys) in the XI (Cha *et al.*, 1994) M2 octahedral coordination sphere concluded that His220 residue was vital for enzyme activity. The k_{cat} values decreased to 0.3-0.5% of the wild type values for His220Ser, His220Asn and His220Glu, and the K_m increased 30-40-fold. His220Lys was inactive. The metal binding was perturbed by replacing His220 and the mechanism disrupted. Thus, His220 appears to fulfil a pivotal role for the XI enzyme. Here, we observe it moving towards the M2a site.

An explanation of this could be hypothesised from the radiation chemistry. Metals are reduced by the X-ray beam and become negatively charged. The crystallisation conditions have a pH of 7.0 and the pKa of His is 6.0. At zero dose, the His is uncharged. Free radical production will lower the pH and if it decreases to the pKa value or below, a positive charge will be generated on the His and thereby produce an attractive force toward the reduced metal. The rotation of the Asp is probably a combination of the metal reduction and decarboxylation of the residue. Thus, radiation chemistry could be driving these changes in the enzymatic active site.

Our study shows that the M2c position is real, and is seen in the unliganded protein at low occupancy. The native protein has three M2 positions and two alternate residue conformations in the active site. M2a is associated with one of these and M2b and M2c with the other. Glu217 remains at a constant distance from both M2a and M2b. It is possible that radiation chemistry is mimicking mechanism but without substrate and product studied in the same manner, it is not possible to conclude this.

The mechanism of action of the XI enzyme has been the subject of discussion for many years and involves the three catalytic metal sites. From the analysis reported above, our hypothesis is that the elusiveness of the mechanism has been confounded by radiation chemistry. Our results showed a motion of the metals in absolute space. This could be interpreted as mechanistically important when comparing models determined with different effective X-ray doses. However, a fractional coordinate analysis showed that it was simply driven by the expansion of the unit cell, a well-known result of global radiation damage.

In summary, this study highlights once again the need to be aware of, and take into account, the various effects of radiation damage when interpreting structures mechanistically, particularly for metalloproteins where even a low dose structure can be perturbed by the radiation chemistry induced around the metal binding site. While the changes seen are small, it shows that both global and specific damage can have a potential impact on structural interpretation.

FIGURE LEGENDS:

Figure 1 A schematic presentation of the reaction mechanism of XI converting xylose to xylulose with the putative mobility of the catalytic metal ion with conformations M2a, M2b and M2c. Mechanism adapted from Asbóth & Náray-Szabó, (2000), Fenn *et al.*, (2004) and Munshi *et al.*, (2014).

Figure 2 Spectrum of the elemental analysis of XI measured with microPIXE showing the presence of manganese, calcium, chlorine, and sulphur. Note that only elements heavier than magnesium were detected in this experiment. The peak between the Ca and Mn corresponds to twice the sulphur peak energy and is due to pile-up in the Si(Li) X-ray detector of the sulphur signals.

Figure 3 Radiation damage metrics and data quality indicators for the seven consecutive complete data sets as a function of dose. a) The unit cell volume expands with increasing absorbed dose, as does b) the relative B -factor and c) R_{meas} , but d) $CC_{1/2}$ and e) the relative diffraction intensity decrease with dose.

Figure 4 The structure of XI. a) XI monomer showing the bi-domain arrangement: N-terminal parallel $(\alpha/\beta)_8$ -barrel fold and a long C-terminal loop. b) A close-up of the XI active site showing the refined metal cofactors (M1 and M2 as purple spheres, waters in red) with their $2F_{\text{obs}} - F_{\text{calc}}$ electron density map contoured at 1σ ($0.41 \text{ e}^-/\text{\AA}^3$).

Figure 5 B_{net} (Shelley *et al.*, 2018) for the refined structures as a function of dose. B_{net} , a derivative of the B_{damage} metric (Gerstel *et al.*, 2015), which is calculated from the atomic B -factors corrected for packing density, summarises in a single value the total extent of specific radiation damage and increases linearly with absorbed dose for the consecutive XI data sets.

Figure 6 Top 25 specific damage sites calculated by RIDL (Bury & Garman, 2018) and presented as red spheres. The radius of each sphere is proportional to the D_{loss} metric calculated by RIDL, which indicates the maximum electron density loss in the local region of each atom in XI for the $F_{\text{obs},n} - F_{\text{obs},1}$ Fourier difference map associated with each higher dose data set n . a) the specific damage sites for $F_{\text{obs},2} - F_{\text{obs},1}$ and b) for $F_{\text{obs},7} - F_{\text{obs},1}$. The increase in the damage with increasing dose is clearly visible, especially at the active site and the surroundings of the catalytic (M2) and stabilising (M1) metals which are shown as violet spheres. For those residues with dual conformations, the highest occupancy positions are shown.

Figure 7 Fourier difference density maps of the active site of XI calculated using FFT (Ten Eyck, 1973) through the RIDL pipeline (Bury & Garman, 2018). Here, each $F_{\text{obs},n} - F_{\text{obs},1}$ map has been calculated using the observed structure factor amplitudes $|F_{\text{obs},n}|$ and $|F_{\text{obs},1}|$ associated with dataset n and 1, using the calculated phase set corresponding to the 0.13 MGy low-dose refined coordinate model. The $F_{\text{obs},n} - F_{\text{obs},1}$ maps represent the electron density loss between the low dose data set and each higher dose data set $n > 1$. a) and b) are difference density maps for $F_{\text{obs},2} - F_{\text{obs},1}$ contoured at $\pm 3\sigma$ ($0.07 \text{ e}^-/\text{\AA}^3$), c) and d) for $F_{\text{obs},7} - F_{\text{obs},1}$ (negative/positive difference densities in red/green respectively). The increase in damage with increasing dose, especially at the active site and the surroundings of the catalytic metals is clear. The difference density maps also reflect visually how the electron density around the catalytic metal ion changes. The conformation with the

highest occupancy, A, are used by RIDL but as mentioned in the text, the RIDL analysis is not sensitive enough to the radiation damage associated with the small structural changes seen between A and B conformations.

Figure 8 Details of the three M2 positions and the alternate residue conformations, A and B, shown in green and violet respectively. Waters are also shown in red. Residues with no alternate conformations are coloured grey. The three metal positions M2a, M2b and M2c are coloured green, violet and orange respectively.

Figure 9 Relative change in distance between residues in the active site and M2 metal positions in conformation A and conformation B for M2a (a, b), M2b (c, d) and M2c (e, f). All graphs are plotted on the same scale so that significant distance changes are easily seen.

Figure 10 Plots of D_{neg} for the A and B residue conformations, plots a) and b) respectively, in the M2 active site shown compared to Asp80 in plot c), one of the residues associated with the XI tetramer crystal contacts.

Table 1 Data processing and refinement statistics

Values in parentheses are for highest resolution shell. The lowest dose data set was cut to the resolution limit of the highest dose so that an identical set of reflections could be used for calculation of R_{free} . All the data were collected at SSRL beamline 11-1. The crystal was in space group $I222$. The number of non-H atoms=4168 (and was not adjusted for each model), protein atoms=3572, water molecules=528, and ligands=68.

Data set	1 (uncut)	1 (cut 1.17 Å)	2	3	4	5	6	7	8 (merged data)
DWD (MGy)	0.13	0.13	0.76	1.38	2.01	2.63	3.25	3.88	3.88
PDB code	6QRR	-	6QRS	6QRT	6QRU	6QRV	6QRW	6QRX	6QRY
Unit cell [a (Å)]	92.48	92.48	92.50	92.52	92.54	92.57	92.59	92.61	92.35
[b (Å)]	97.46	97.46	97.49	97.52	97.54	97.58	97.61	97.64	97.28
[c (Å)]	102.46	102.49	102.50	102.52	102.54	102.56	102.58	102.60	102.3
Resolution (Å)	38.69-1.096 (1.135-1.096)	38.63-1.17 (1.19-1.17)	38.69-1.17 (1.19-1.17)	38.70-1.17 (1.19-1.17)	38.71-1.17 (1.19-1.17)	38.72-1.17 (1.19-1.17)	39.79-1.17 (1.19-1.17)	38.74-1.17 (1.19-1.17)	39.67-1.17 (1.19-1.17)
No. of observed reflections	589885 (22614)	537565 (17750)	540859 (17728)	540378 (17781)	540698 (17818)	540867 (17745)	542096 (17827)	538263 (17654)	3770952 (125015)
No. of unique reflections	180192 (12983)	153154 (7271)	154106 (7391)	154212 (7401)	154295 (7413)	154466 (7422)	154610 (7420)	154713 (7399)	153877 (7311)
Multiplicity	3.3 (1.7)	3.5 (2.4)	3.5 (2.4)	3.5 (2.4)	3.5 (2.4)	3.5 (2.4)	3.5 (2.4)	3.5 (2.4)	24.5 (17.1)
Completeness (%)	95.7 (69.6)	99.4 (96.6)	99.4 (96.4)	99.4 (96.5)	99.4 (96.4)	99.4 (96.5)	99.4 (96.3)	99.4 (96.5)	99.8 (97.1)
Wilson B -factor (Å ²)	10.3	7.8	8.2	8.5	8.8	9.1	9.5	9.9	8.6
R_{meas} (%)	7.2 (43.8)	8.5 (37.2)	8.5 (42.2)	8.7 (47.6)	8.6 (53.9)	9.0 (62.3)	9.0 (73.4)	9.0 (84.7)	10.2 (55.9)
$CC_{1/2}$ (%)	99.3 (67.9)	99.4 (86.7)	99.4 (82.9)	99.3 (79.6)	99.3 (74.6)	99.3 (69.3)	99.3 (61.4)	99.3 (54.5)	99.8 (94.9)
$I/\sigma I$	10.7 (1.8)	12.2 (3.2)	11.7 (2.8)	11.3 (2.5)	11.1 (2.3)	10.5 (1.9)	10.2 (1.7)	9.9 (1.4)	23.3 (5.5)
R_{work} (%) / R_{free} (%)	12.1/14.7	9.8/11.5	9.9/11.7	10.0/11.8	10.2/12.0	10.5/12.6	10.7/12.6	10.9/12.8	9.3/11.0
Average B -factor (Å ²)	17.1	16.1	16.6	17.1	17.7	18.1	18.6	19.2	16.6
Protein	13.7	12.7	13.2	13.7	14.1	14.5	15.0	15.4	13.1
Ligands	21.7	21.6	24.2	25.3	27.4	28.0	30.0	31.3	26.3
rmsd bond lengths (Å)	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
rmsd bond angles (°)	2.3	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
Ramachadran plot (%)									
Favoured	96.7	96.9	96.7	97.4	97.1	97.4	97.1	97.7	97.4
Additionally allowed	2.9	2.9	2.9	2.3	2.6	2.3	2.6	2.1	2.3

Table 2 Occupancies of the catalytic metal M2 and alternate conformations of active site amino acid residues as a function of dose

DWD (MGy)	M2a	M2b	M2c	Glu 217		His 220		Asp 255		Asp 257	
				A	B	A	B	A	B	A	B
0.13	0.41	0.55	0.04	0.13	0.87	0.32	0.68	0.42	0.58	0.38	0.62
0.76	0.44	0.52	0.04	0.14	0.86	0.36	0.64	0.37	0.63	0.35	0.65
1.38	0.41	0.55	0.04	0.13	0.87	0.33	0.67	0.39	0.61	0.33	0.67
2.01	0.41	0.54	0.05	0.14	0.86	0.31	0.69	0.40	0.60	0.35	0.65
2.63	0.40	0.55	0.05	0.14	0.86	0.33	0.67	0.33	0.65	0.37	0.63
3.25	0.42	0.51	0.07	0.13	0.87	0.31	0.69	0.37	0.63	0.34	0.66
3.88	0.46	0.50	0.04	0.14	0.86	0.33	0.67	0.38	0.62	0.38	0.62

Table 3 Absolute and fractional coordinates of the M2a, M2b, and M2c positions as a function of dose

Fractional dose is given to three decimal places to increase slight changes.

Absolute position				Fractional position		
M2a						
DWD (MGy)	x	Y	c	X	y	c
0.13	7.55	35.19	43.79	0.082	0.380	0.474
0.76	7.57	35.21	43.86	0.082	0.381	0.474
1.38	7.57	35.24	43.87	0.082	0.381	0.474
2.01	7.59	35.26	43.93	0.082	0.381	0.475
2.63	7.59	35.26	43.93	0.082	0.381	0.475
3.25	7.62	35.33	44.04	0.082	0.382	0.476
3.88	7.64	35.33	44.05	0.082	0.382	0.476
M2b						
0.13	7.80	35.57	44.47	0.080	0.365	0.456
0.76	7.80	35.63	44.52	0.080	0.365	0.457
1.38	7.79	35.63	44.55	0.080	0.365	0.457
2.01	7.80	35.67	44.58	0.080	0.366	0.457
2.63	7.80	35.69	44.59	0.080	0.366	0.457
3.25	7.81	35.72	44.63	0.080	0.366	0.457
3.88	7.77	35.75	44.62	0.080	0.366	0.457
M2c						
0.13	9.05	35.21	44.85	0.088	0.344	0.438
0.76	9.00	35.27	44.91	0.088	0.344	0.438
1.38	9.02	35.29	44.92	0.088	0.344	0.438
2.01	8.99	35.34	44.96	0.088	0.345	0.438
2.63	9.03	35.33	44.95	0.088	0.344	0.438
3.25	9.01	35.33	44.98	0.088	0.344	0.438
3.88	8.98	35.35	44.96	0.088	0.345	0.438

Active site metal bonding distances for the lowest and highest dose structures with the coordinate error (CE) (Kumar *et al.*, 2015) for the bonding atoms. Bonds within one standard deviation of expectations (Zheng *et al.*, 2008) are shown with green highlighting, those with two standard deviations in yellow and finally, those within three standard deviations, unlikely to be bonds, not highlighted at all. A difference is clearly seen between low and high doses for M2a and M2c, compared with M2b. The differences are significant when compared to the errors indicated by the precision of the atomic coordinates.

Table 4

Distances, including the coordinate error (CE), are given in Å. No B conformation is given, if the distances are above 3 standard deviations from an expected bond length.

	Lowest DWD 0.13 MGy						Highest DWD 3.88 MGy					
	M1Mn	M1Mg	M2a	M2b	M2c	CE	M1Mn	M1Mg	M2a	M2b	M2c	CE
Conformation A												
Glu181(OE2)	2.08	2.01				0.013	2.09	2.00				0.017
Asp245(OD2)	1.95	2.05				0.014	1.92	2.07				0.017
Asp287(OD2)	2.01	2.05				0.013	2.02	2.09				0.017
EDO1(O2)	2.17	2.23				0.014						0.019
HOH148	2.44	2.33				0.016	2.36	2.24				0.022
Glu217(OE1)	1.72	1.68				0.011	1.78	1.55				0.013
Glu217(OE2)			2.09	2.15	2.17	0.010			2.13	2.27	2.22	0.013
His220(NE2)			3.73	2.99	2.21	0.011			3.46	2.85	2.15	0.014
Asp255(OD2)			1.98	1.66	2.62	0.013			1.85	1.54	2.51	0.017
Asp257(OD1)			1.86	2.59	2.69	0.014			1.90	2.51	3.50	0.015
HOH38			1.94	2.23	2.02	0.015			1.92	2.29	1.96	0.019
Conformation B												
Glu217(OE1)	2.13	2.05				0.013	2.25	1.97				0.016
Glu217(OE2)			2.04	2.00	2.33	0.012			2.04	2.01	2.30	0.015
His220(NE2)			3.06	2.29	1.83	0.012			2.87	2.22	1.81	0.015
Asp255(OD1)			2.61	2.26	3.41	0.013			2.59	2.21	3.37	0.015
Asp255(OD2)			2.62	2.24	2.90	0.014			2.61	2.27	2.91	0.016
Asp257(OD1)			1.6	2.30	3.44	0.014			1.80	2.34	3.43	0.017

Acknowledgements Kristin A. Sutton, Ana González, Kathryn L. Shelley and Joshua L. Dickerson are thanked for the various contributions to this work. Stanford Synchrotron Radiation Lightsource (SSRL, Stanford, USA) and the staff of beamline 11-1 are gratefully acknowledged. Use of the SSRL, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH. We also gratefully acknowledge the Surrey Ion Beam Centre and Geoff Grime for assistance with the PIXE measurements, and thank the UK Engineering and Physical Sciences Research Council (grant No. EP/G03706X/1) for studentship funding in the Systems Biology Programme of the University of

Oxford Doctoral Training Centre (CBS). This work was supported by the Finnish Cultural Foundation and the NSF BioXFEL Center (grant 1231306).

References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. -, Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* **D66**, 213-221.
- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* **D68**, 352-367.
- Allen, K. N., Lavie, A., Glasfeld, A., Tanada, T. N., Gerrity, D. P., Carlson, S. C., Farber, G. K., Petsko, G. A. & Ringe, D. (1994). *Biochemistry*. **33**, 1488-1494.
- Asbóth, B. & Náray-Szabó, G. (2000). *Curr. Protein Pept. Sci.* **1**, 237-254.
- Berman, H., Henrick, K. & Nakamura, H. (2003). *Nat. Struct. Mol. Biol.* **10**, 980.
- Berman, H., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalov, I. & Bourne, P. (2000). *Nucleic Acids Res.* **28**, 235-242.
- Bhosale, S. H., Rao, M. B. & Deshpande, V. V. (1996). *Microbiol. Rev.* **60**, 280-300.
- Blow, D. M., Collyer, C. A., Goldberg, J. D. & Smart, O. S. (1992). *Faraday Discuss.* **93**, 67-73.
- Bogumil, R., Kappl, R., Hüttermann, J. & Witzel, H. (1997). *Biochemistry*. **36**, 2345-2352.
- Bowman, S. E., Bridwell-Rabb, J. & Drennan, C. L. (2016). *Acc. Chem. Res.* **49**, 695-702.
- Burmeister, W. P. (2000). *Acta Cryst.* **D56**, 328-341.
- Bury, C. S. & Garman, E. F. (2018). *J. Appl. Cryst.* **51**, 952-962.
- Bury, C. S., Brooks-Bartlett, J. C., Walsh, S. P. & Garman, E. F. (2018). *Protein Sci.* **27**, 217-228.
- Bury, C. S., Carmichael, I. & Garman, E. F. (2017). *J. Synchrotron Rad.* **24**, 7-18.
- Callens, M., Kersters-Hilderson, H., Van Opstal, O. & De Bruyne, C. K. (1986). *Enzyme Microb. Technol.* **8**, 696-700.
- Carrell, H. L., Rubin, B. H., Hurley, T. J. & Glusker, J. P. (1984). *J. Biol. Chem.* **259**, 3230-3236.
- Carugo, O. & Carugo, K. D. (2005). *Trends Biochem. Sci.* **30**, 213-219.
- Cha, J., Cho, Y., Whitaker, R. D., Carrell, H. L., Glusker, J. P., Karplus, P. A. & Batt, C. A. (1994). *J. Biol. Chem.* **269**, 2687-2694.
- Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12-21.
- Collyer, C., Henrick, K. & Blow, D. (1990). *J. Mol. Biol.* **212**, 211-235.
- Corbett, M. C., Latimer, M. J., Poulos, T. L., Sevrioukova, I. F., Hodgson, K. O. & Hedman, B. (2007). *Acta Cryst.* **D63**, 951-960.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486-501.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72-82.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282-292.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204-1214.
- Fenn, T. D., Ringe, D. & Petsko, G. A. (2002). doi:10.2210/pdb1MUW/pdb.
- Fenn, T. D., Ringe, D. & Petsko, G. A. (2004). *Biochemistry*. **43**, 6464-6474.
- Garman, E. (1999). *Structure*. **7**, R291-R299.
- Garman, E. F. (2010). *Acta Cryst.* **D66**, 339-351.
- Garman, E. F. & Grime, G. W. (2005). *Prog. Biophys. Mol. Biol.* **89**, 173-205.
- Gerstel, M., Deane, C. M. & Garman, E. F. (2015). *J. Synchrotron Rad.* **22**, 201-212.
- Grabolle, M., Haumann, M., Muller, C., Liebisch, P. & Dau, H. (2006). *J. Biol. Chem.* **281**, 4580-4588.
- Grime, G. (1996). *Nucl. Instrum. Methods Phys. Res. B*. **109**, 170-174.
- Grime, G., Dawson, M., Marsh, M., McArthur, I. & Watt, F. (1991). *Nucl. Instrum. Methods Phys. Res. B*. **54**, 52-63.
- Gutmanas, A., Alhroub, Y., Battle, G. M., Berrisford, J. M., Bochet, E., Conroy, M. J., Dana, J. M., Fernandez Montecelo, M. A., van Ginkel, G., Gore, S. P., Haslam, P., Hatherley, R., Hendrickx, P. M., Hirshberg, M., Lagerstedt, I., Mir, S., Mukhopadhyay, A., Oldfield, T. J., Patwardhan, A., Rinaldi, L., Sahni, G., Sanz-Garcia, E., Sen, S., Slowley, R. A., Velankar, S., Wainwright, M. E. & Kleywegt, G. J. (2014). *Nucleic Acids Res.* **42**, D285-91.
- Holton, J. M. (2009). *J. Synchrotron Rad.* **16**, 133-142.
- Jenkins, J., Janin, J., Rey, F., Chiadmi, M., Van Tilbeurgh, H., Lasters, I., De Maeyer, M., Ven Belle, D. & Wodak, S. J. (1992). *Biochemistry*. **31**, 5449-5458.

Jokela, J., Pastinen, O. & Leisola, M. (2002). *Enzyme Microb. Technol.* **31**, 67-76.

Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. (2014). *IUCrJ*. **1**, 213-220.

Kabsch, W. (2010). *Acta Cryst.* **D66**, 125-132.

Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865-1871.

Kovalevsky, A. Y., Hanson, L., Fisher, S. Z., Mustyakimov, M., Mason, S. A., Forsyth, V. T., Blakeley, M. P., Keen, D. A., Wagner, T. & Carrell, H. (2010). *Structure*. **18**, 688-699.

Kmetko, J., Husseini, N. S., Naides, M., Kalinin, Y. & Thorne, R. E. (2006). *Acta Cryst.* **D62**, 1030-1038.

Kumar, K. D., Gurusaran, M., Satheesh, S. N., Radha, P., Pavithra, S., Thulaa Tharshan, K. P. S., ... & Sekar, K. (2015). *J. Appl. Cryst.* **48**, 939-942.

Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nat. Protoc.* **3**, 1171-1179.

Lavie, A., Allen, K. N., Petsko, G. A. & Ringe, D. (1994). *Biochemistry*. **33**, 5469-5480.

Lee, M., Rozeboom, H. J., de Waal, P. P., de Jong, R. M., Dudek, H. M. & Janssen, D. B. (2017). *Biochemistry*. **56**, 5991-6005.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491-497.

Maxwell, J., Campbell, J. & Teesdale, W. (1989). *Nucl. Instrum. Methods Phys. Res. B*. **43**, 218-230.

McCoy, A. J., Oeffner, R. D., Wrobel, A. G., Ojala, J. R. M., Tryggvason, K., Lohkamp, B. & Read, R. J. (2017). *Proc. Natl. Acad. Sci. USA* **114**, 3637-3641.

Meilleur, F., Snell, E. H., Van der Woerd, Mark J, Judge, R. A. & Myles, D. A. (2006). *Eur. Biophys. J.* **35**, 601-609.

Munshi, P., Snell, E. H., Van Der Woerd, Mark J, Judge, R. A., Myles, D. A., Ren, Z. & Meilleur, F. (2014). *Acta Cryst.* **D70**, 414-420.

Murray, J. & Garman, E. (2002). *J. Synchrotron Rad.* **9**, 347-354.

Murray, J. W., Rudino-Pinera, E., Owen, R. L., Grninger, M., Ravelli, R. B. & Garman, E. F. (2005). *J. Synchrotron Rad.* **12**, 268-275.

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.* **D67**, 355-367.

Nave, C. (1995). *Radiat. Phys. Chem.* **45**, 483-490.

O'Neill, P., Stevens, D. L. & Garman, E. F. (2002). *J. Synchrotron Rad.* **9**, 329-332.

Owen, R. L., Rudino-Pinera, E. & Garman, E. F. (2006). *Proc. Natl. Acad. Sci. USA*. **103**, 4912-4917.

Pedersen, S. (1993). *Bioprocess Technol.* **16**, 185-208.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nat. Struct. Mol. Biol.* **6**, 458-463.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comp. Chem.* **25**, 1605-1612.

Rangarajan, M., Asboth, B. & Hartley, B. S. (1992). *Biochem. J.* **285 (Pt 3)**, 889-898.

Ravelli, R. B. & McSweeney, S. M. (2000). *Structure*. **8**, 315-328.

Rould, M. A. & Carter, C. W. (2003). *Meth. Enzymol.* **374**, 145-163.

Schray, K. J. & Mildvan, A. S. (1972). *J. Biol. Chem.* **247**, 2034-2037.

Sheldrick, G. M. (2015). *Acta Cryst.* **C71**, 3-8.

Shelley, K. L., Dixon, T. P., Brooks-Bartlett, J. C. & Garman, E. F. (2018). *J. Appl. Cryst.* **51**, 552-559.

Snell, E. H., Van der Woerd, Mark J, Damon, M., Judge, R. A., Myles, D. A. & Meilleur, F. (2006). *Eur. Biophys. J.* **35**, 621-632.

Ten Eyck, L. F. (1973). *Acta Cryst.* **A29**, 183-191.

Weik, M., Ravelli, R. B., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl. Acad. Sci. USA*. **97**, 623-628.

Whitaker, R. D., Cho, Y., Cha, J., Carrell, H. L., Glusker, J. P., Karplus, P. A. & Batt, C. A. (1995). *J. Biol. Chem.* **270**, 22895-22906.

Whitlow, M., Howard, A. J., Finzel, B. C., Poulos, T. L., Winborne, E. & Gilliland, G. L. (1991). *Proteins*. **9**, 153-173.

Williams, R. J. (1985). *FEBS J.* **150**, 231-248.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* **D67**, 235-242.

Zeldin, O. B., Brockhauser, S., Bremridge, J., Holton, J. M. & Garman, E. F. (2013a). *Proc. Natl. Acad. Sci. USA*. **110**, 20551-20556.

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

Zheng, H., Chruszcz, M., Lasota, P., Lebioda, L. & Minor, W. (2008). *J. Inorg. Biochem.* **102**, 1765-1776.

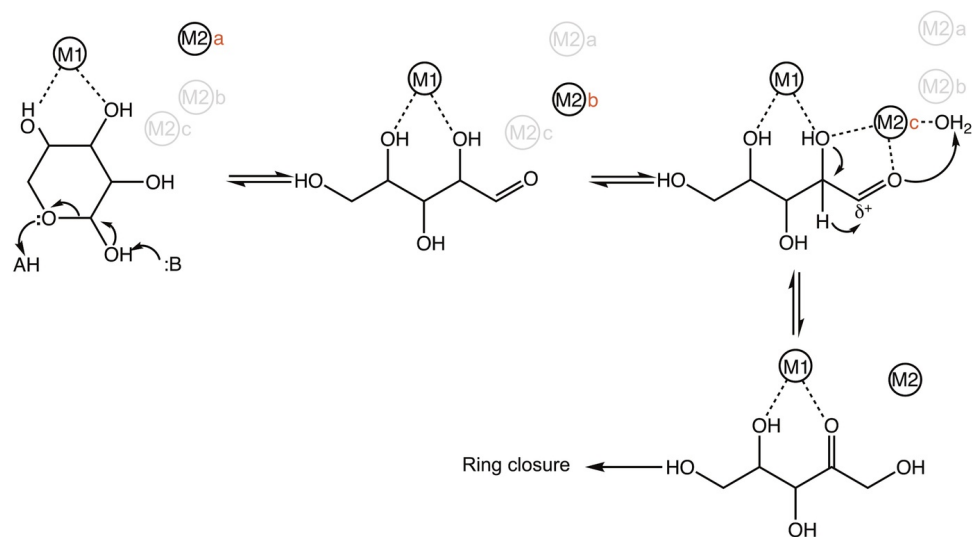


Figure 1

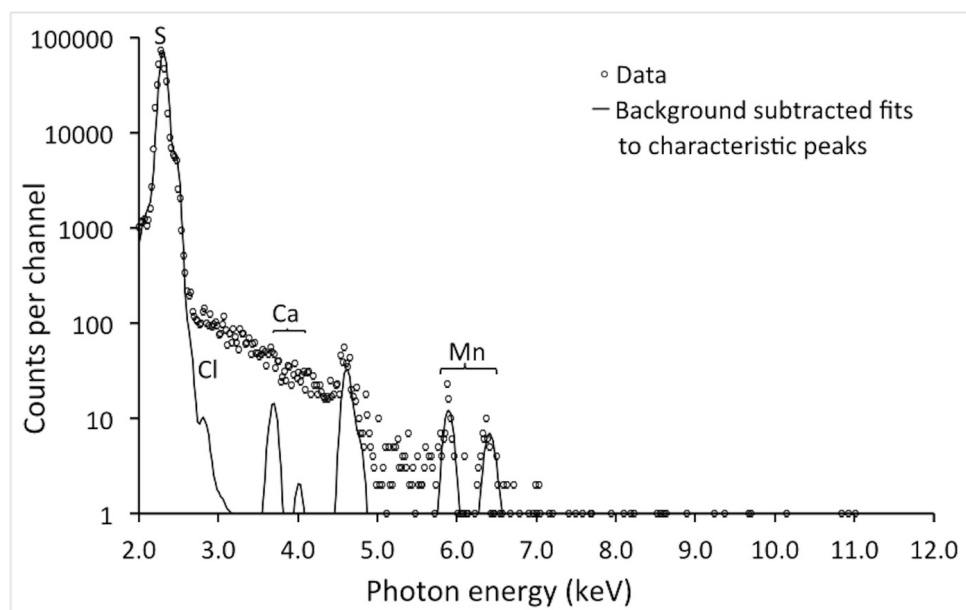


Figure 2

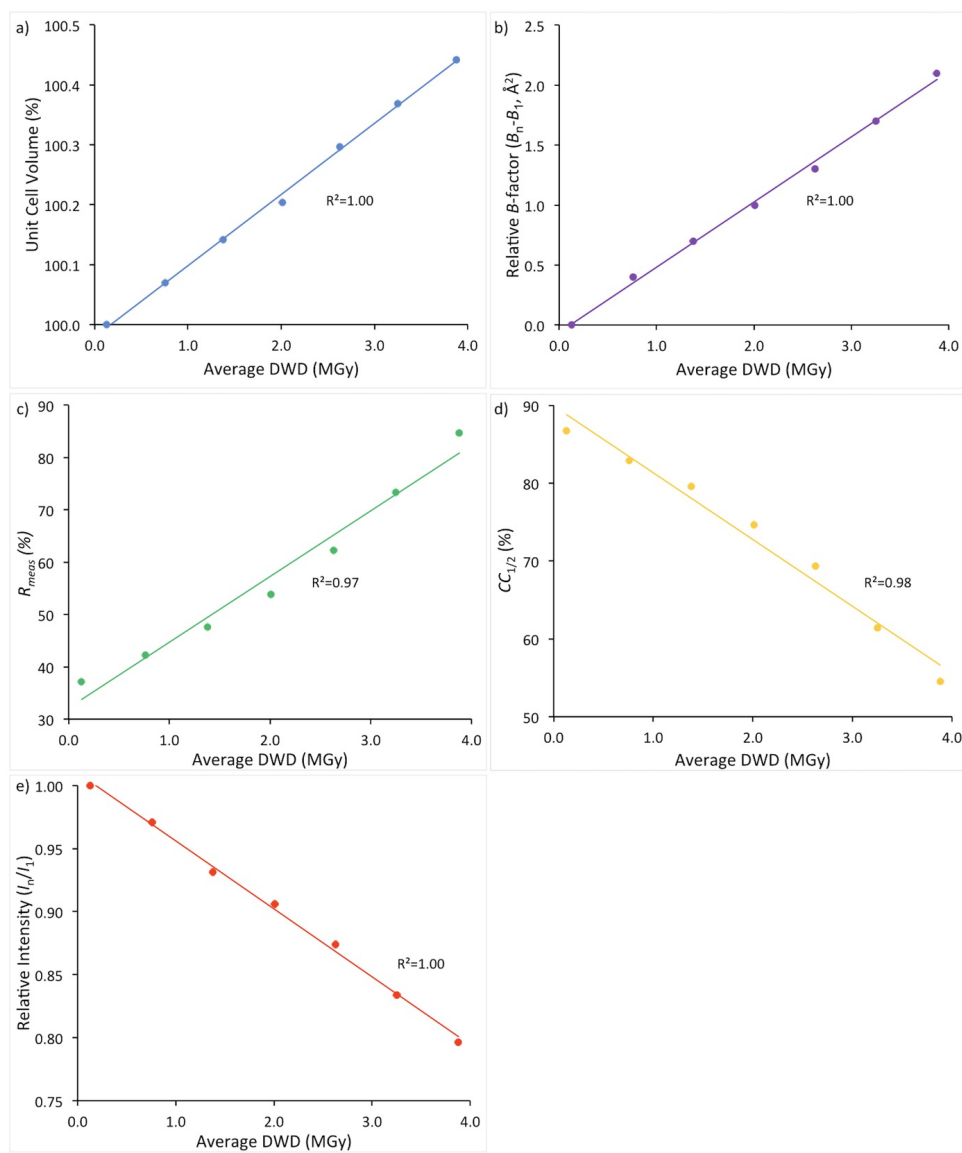


Figure 3

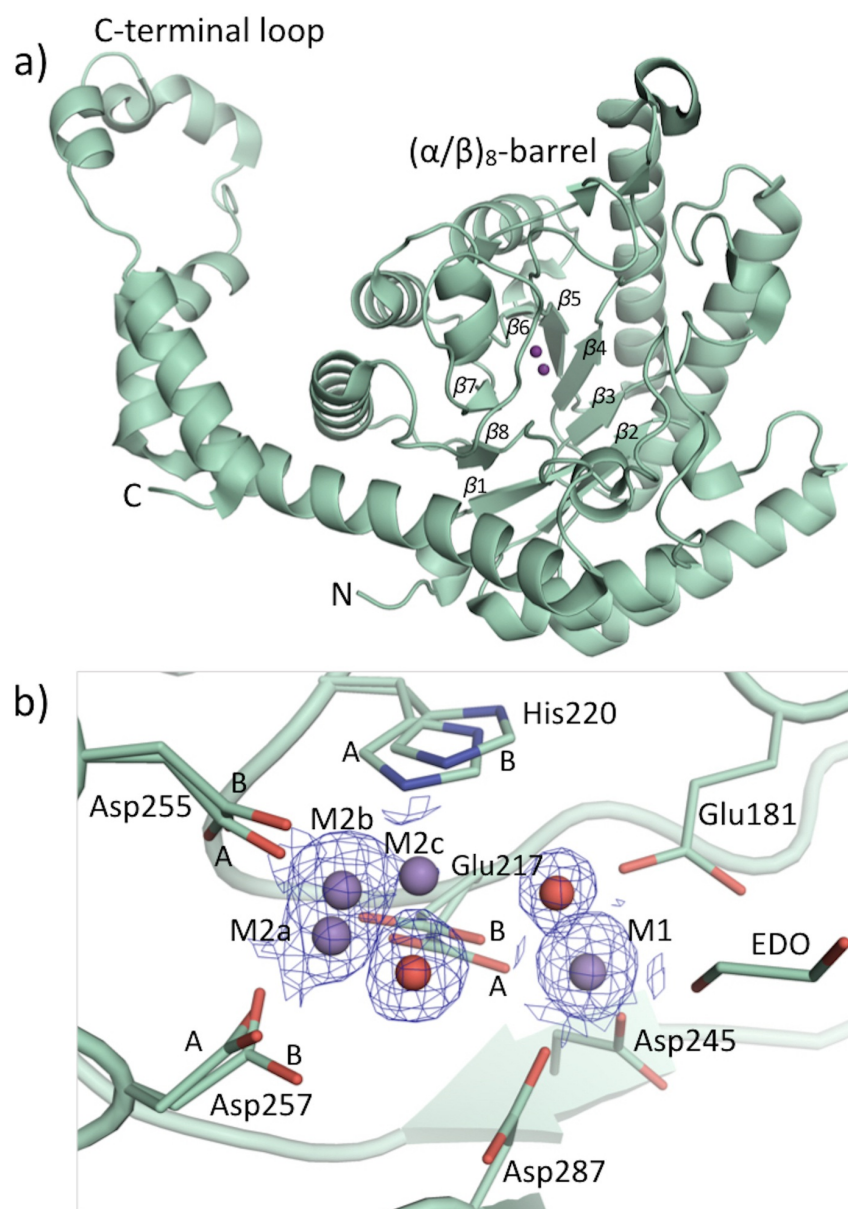


Figure 4

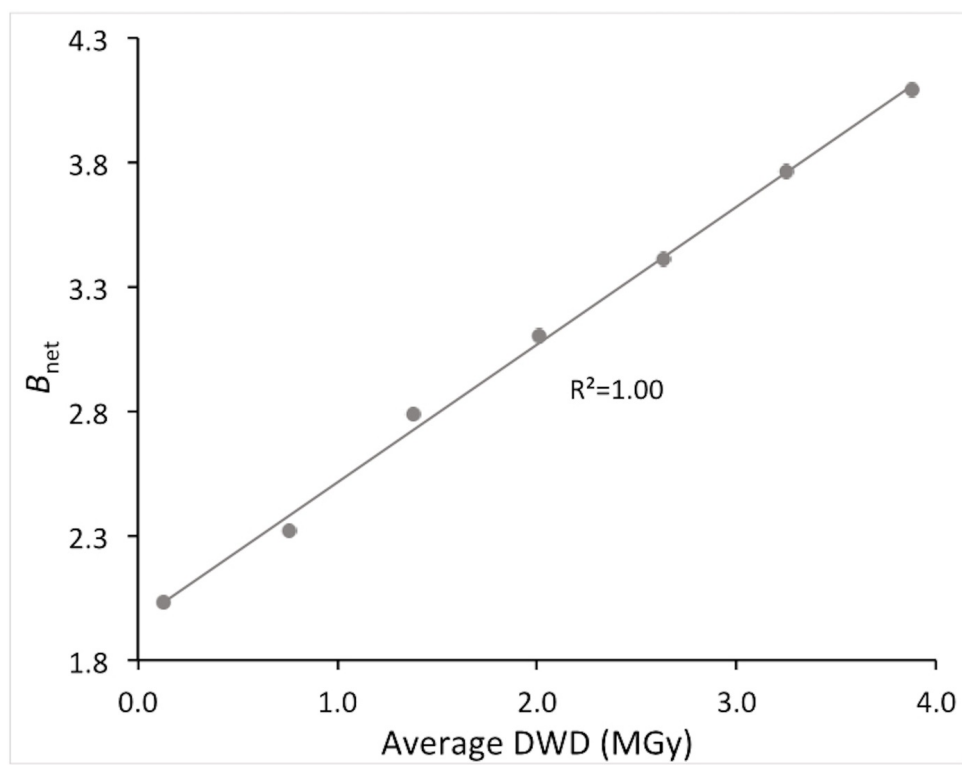


Figure 5

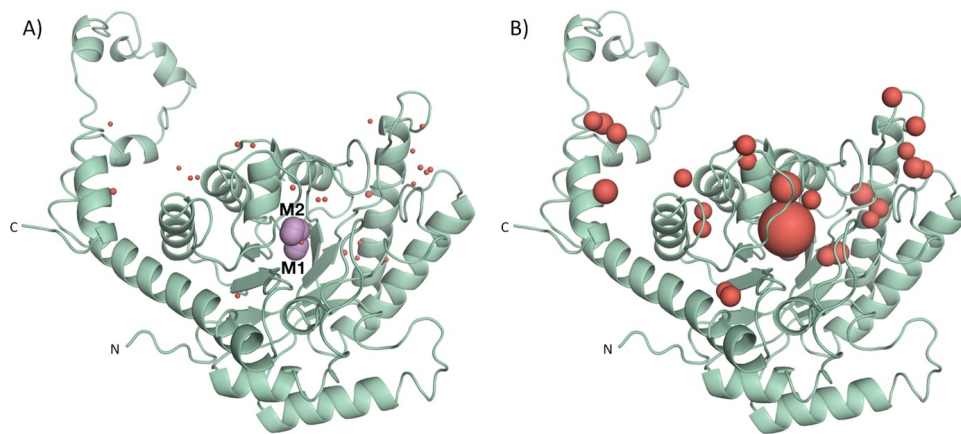


Figure 6

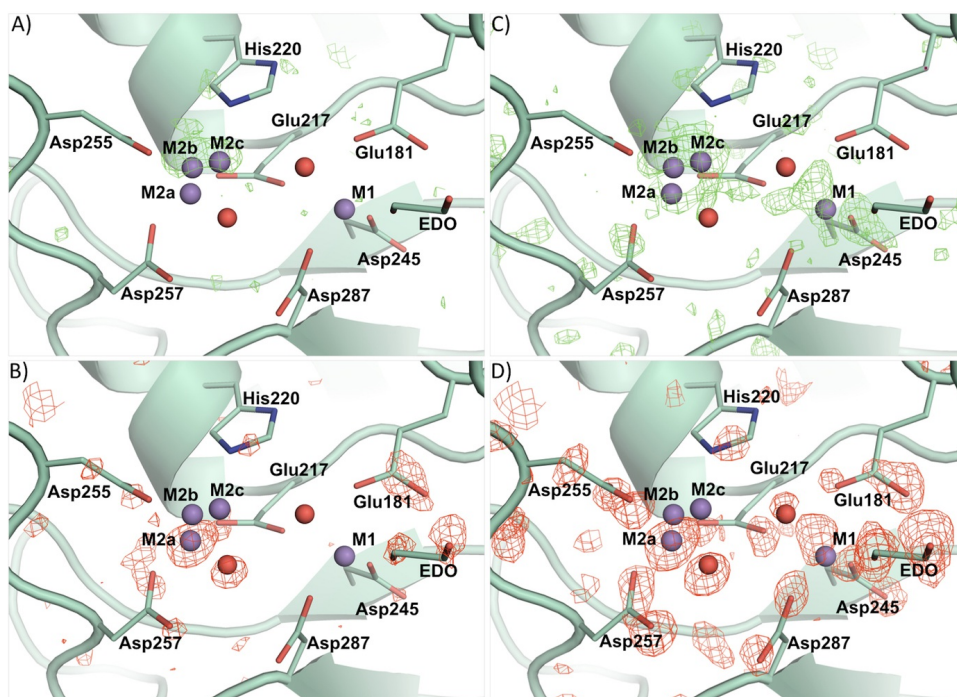


Figure 7

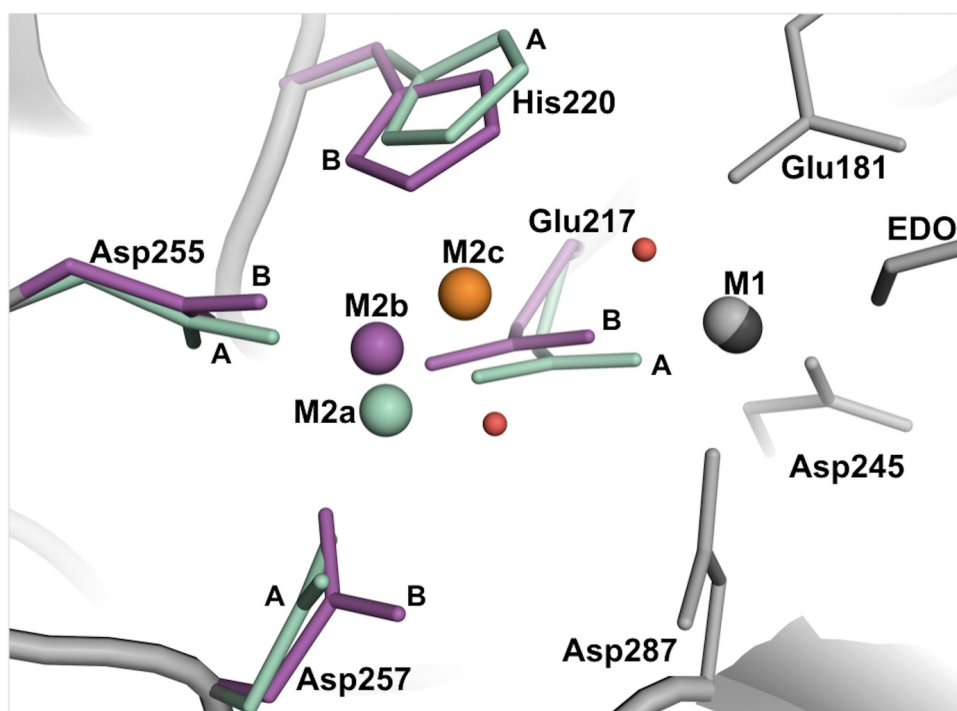


Figure 8

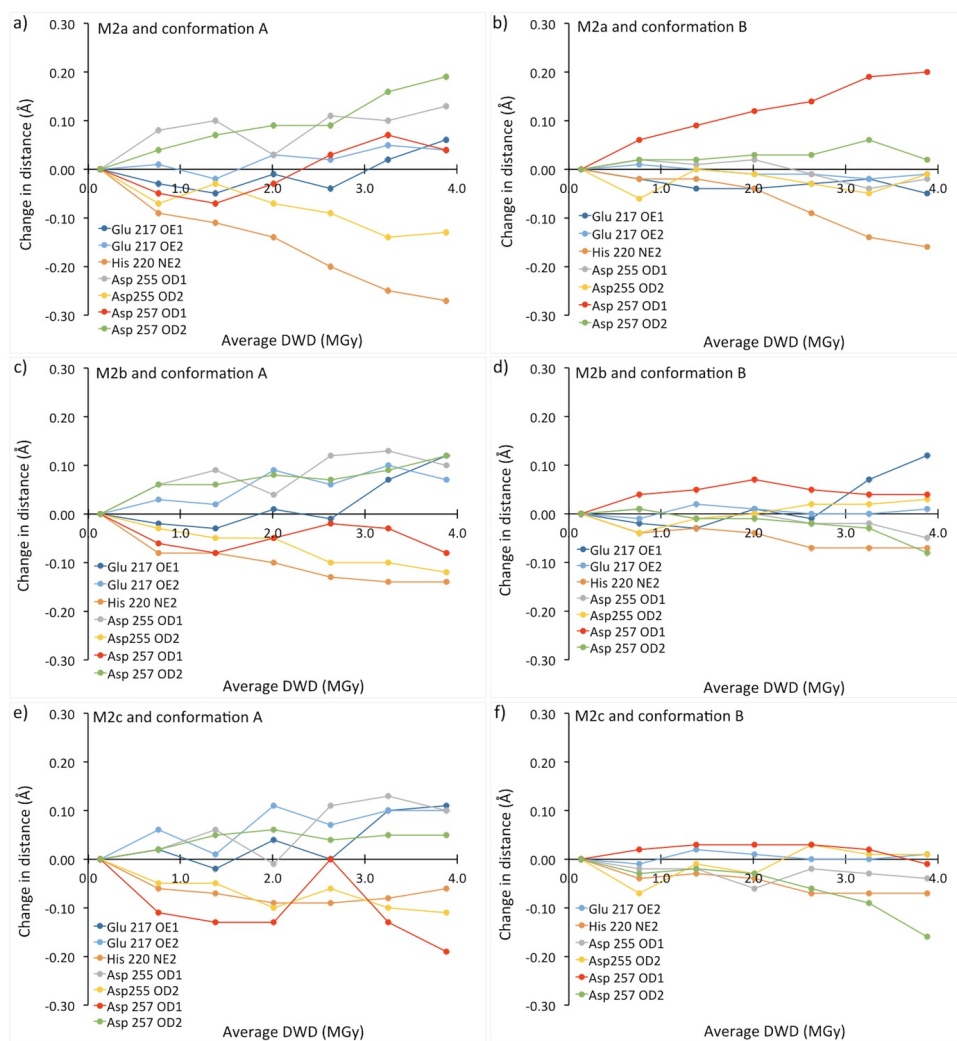


Figure 9

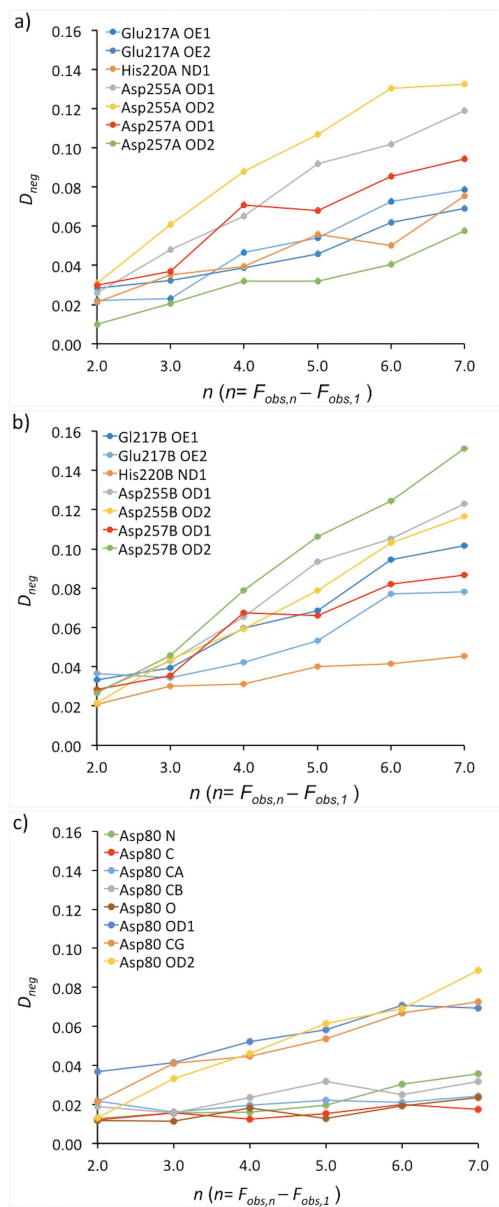


Figure 10