

# Serial femtosecond and serial synchrotron crystallography yield data of equivalent quality: a systematic comparison

Mehrabi, P.<sup>1,2,12,13</sup> §\*, Bücker R.<sup>1</sup>, Bourenkov G.<sup>3</sup>, Ginn H.M.<sup>4</sup>, von Stetten D.<sup>3</sup>, Müller-Werkmeister H.M.<sup>4</sup>, Kuo A.<sup>6</sup>, Morizumi, T.<sup>6</sup>, Eger, B.T.<sup>6</sup>, Ou, W.-L.<sup>6</sup>, Oghbaey S.<sup>7</sup>, Sarracini A.<sup>7</sup>, Besaw, J.E.<sup>6,7</sup>, Pare' -Labrosse, O.<sup>1,7</sup>, Meier S.<sup>8</sup>, Schikora H.<sup>9</sup>, Tellkamp F.<sup>9</sup>, Marx A.<sup>1,13</sup>, Sherrell D.A.<sup>4,10</sup>, Axford D.<sup>4</sup>, Owen R.<sup>4</sup>, Ernst, O.P.<sup>6,11</sup>, Pai, E.F.<sup>2,6,13</sup>, Schulz, E.C.<sup>1,13</sup> §\* and Miller, R.J.D.<sup>1,7,8,13</sup>

§These authors contributed equally to this work.

\*corresponding authors: [pedram.Mehrabi@mpsd.mpg.de](mailto:pedram.Mehrabi@mpsd.mpg.de); [eike.schulz@mpsd.mpg.de](mailto:eike.schulz@mpsd.mpg.de);

<sup>1</sup>Department for Atomically Resolved Dynamics, Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

<sup>2</sup>Department of Medical Biophysics, University of Toronto, 101 College Street, Toronto, Ontario, M5G 1L7, Canada,

<sup>3</sup>European Molecular Biology Laboratory (EMBL), Hamburg Outstation c/o Deutsches Elektronen Synchrotron (DESY), Notkestrasse 85, D-22603 Hamburg, Germany

<sup>4</sup>Diamond Light Source, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK

<sup>5</sup>Institute of Chemistry - Physical Chemistry, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam-Golm, Germany

<sup>6</sup>Department of Biochemistry, University of Toronto, 1 King's College Circle, Toronto, Ontario, M5S 1A8, Canada

<sup>7</sup>Departments of Chemistry and Physics, University of Toronto, 80 St. George Street, Toronto, Ontario, M5S 3H6, Canada

<sup>8</sup>Department of Physics, Universität Hamburg, Jungiusstrasse 9, 20355 Hamburg, Germany

<sup>9</sup>Scientific Support Unit Machine Physics, Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

<sup>10</sup>Structural Biology Center, X-ray Science Division, Argonne National Laboratory, Argonne, IL, USA

<sup>11</sup>Department of Molecular Genetics, University of Toronto, 1 King's College Circle, Toronto, Ontario, M5S 1A8, Canada

<sup>12</sup>The Campbell Family Cancer Research Institute, Ontario, Cancer Institute, 101 College Street, Toronto, Ontario, M5G 1L7, Canada,

<sup>13</sup>Center for Free-Electron Laser Science, Luruper Chaussee 149, 22761 Hamburg, Germany

**Keywords:** SSX, SFX, systematic comparison, serial crystallography, XFEL, synchrotron

## Abstract

For the two proteins myoglobin (MB) and fluoroacetate dehalogenase (FACD), we present a systematic comparison of crystallographic diffraction data collected by serial femtosecond (SFX) and serial synchrotron crystallography (SSX). To maximize comparability, we used the same batch of crystals, the same sample delivery device, as well as the same data analysis software. Overall figures of merit indicate that the data of both radiation sources are of equivalent quality. For both proteins reasonable data statistics can be obtained with approximately 5000 room temperature diffraction images irrespective of the radiation source. The direct comparability of SSX and SFX data indicates that diffraction quality is rather linked to the properties of the crystals than to the radiation source. Time-resolved experiments can therefore be conducted at the source that best matches the desired time-resolution.

## Introduction

Structural biology has been highly successful in providing three-dimensional information about the architecture of biomolecules. Usually these structures display equilibrium state conformations, important snapshots that provide insight into functional states of these biomolecules<sup>1</sup>. The advent of X-ray free-electron laser (XFEL) sources promised the possibility to study ultra-fast timescales down to the femtosecond domain to watch conformational changes and even chemical reactions<sup>2</sup>. Addressing irreversible reactions at these ultra-fast timescales and guaranteeing homogeneous reaction initiation required the use of micron-sized protein crystals. However, the extremely intense beam-brightness of XFEL sources will destroy those protein micro-crystals during single exposures. This dilemma was circumvented by serial data collection schemes in which hundreds of thousands of micro-crystals are exploited to each provide a single still diffraction pattern. Due to the *diffraction-before-destruction* approach, which allows the collection of a useful diffraction pattern by outrunning the onset of damage, it has been argued that serial femtosecond crystallography (SFX) data from XFELs are *de-facto* radiation damage free<sup>3-5</sup>.

Soon after the first protein crystal structures were solved from SFX data, the method was adapted for use at synchrotrons giving rise to serial synchrotron crystallography (SSX)<sup>6,7</sup>. The majority of SSX experiments thus far involved static structures. However, since most enzymes have median turnover times in the hundreds of milliseconds range, synchrotrons represent a valid alternative for time-resolved experiments. Historically exploited via Laue-diffraction on single-crystals<sup>8,9</sup>, recently this has been the subject of serial data collection approaches to study enzymes on timescales from milliseconds to many seconds<sup>10-15</sup>. The results of those studies have sparked renewed and expanded interest in serial crystallography methods, and the implementation of time-resolved SSX at a number of synchrotron beamlines can be expected, in particular as improving synchrotron technology will also bring faster time-domains into reach.

Due to the large number of photons compressed into the X-ray laser pulses SFX exposure times can be very short, down to the femtosecond domain<sup>2</sup>. As a consequence of the more limited photon flux at synchrotron sources accordingly longer exposure times are needed in SSX, and consequently radiation damage cannot be fully outrun. This is particularly severe for time-resolved experiments because data need to be collected at room-temperature leading to much increased radiation sensitivity. However, while radiation damage is more pronounced at synchrotron sources, much of its deleterious effects can be mitigated by using low-dose exposure, which still leaves the electron density interpretable with little ambiguity<sup>16,17</sup>. The dose causing a 50% drop in overall diffraction intensity for lysozyme at room temperature SSX has been reported to be 380 kGy<sup>18</sup>. While this represents the upper dose limit for serial crystallography experiments, which should not be exceeded at room temperature, the onset of site-specific damage could be observed at a much lower dose of approximately 80 kGy<sup>18</sup>. Although this value is much lower than doses encountered during typical cryo experiments, at modern microfocus

beamlines it is still possible to collect multiple frames from a single micro-crystal at room temperature within this dose limit<sup>17–19</sup>.

For a systematic comparison of SFX and SSX data, we analyzed two different proteins at both radiation sources. To minimize sources of potential discrepancies, we used the same batch of micron-sized crystals, the same sample delivery device, and the same data analysis software for each of the data sets.

## Results and discussion

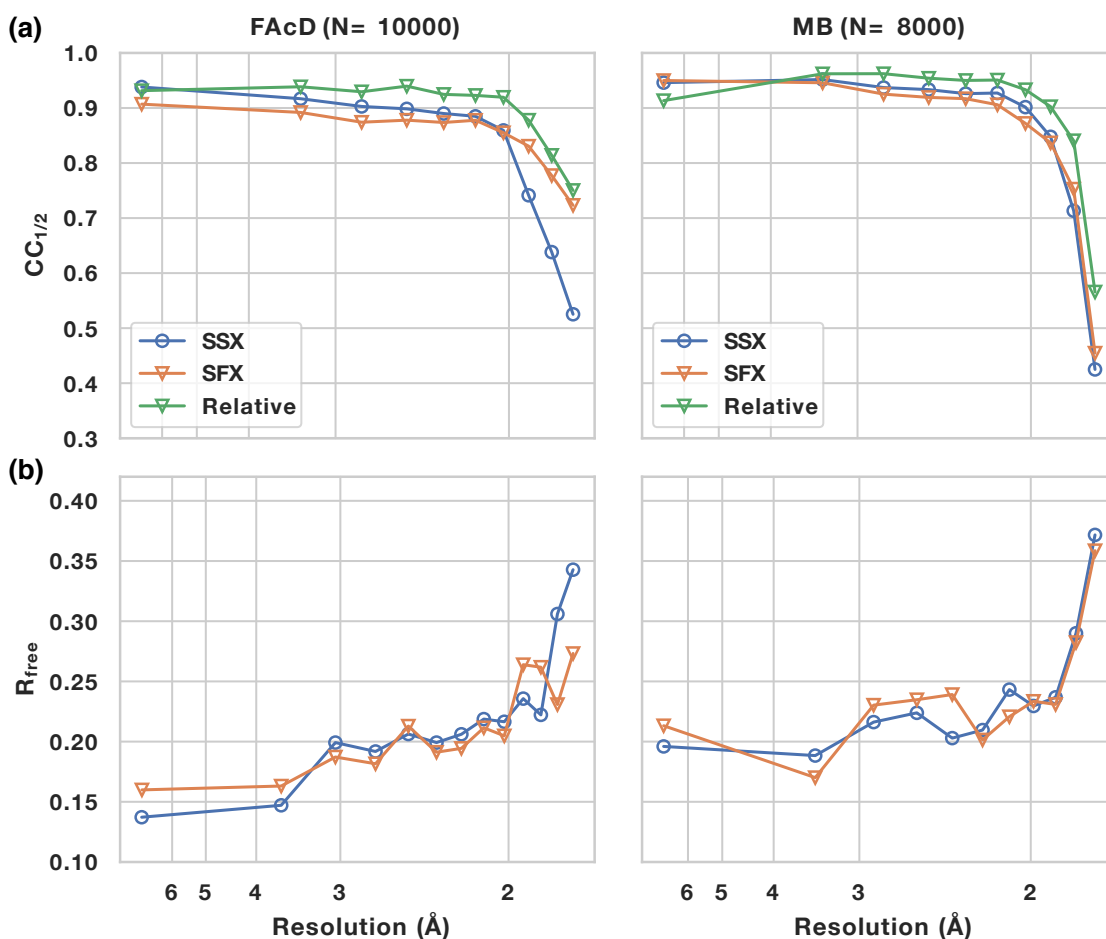
To evaluate potential differences in quality between crystal diffraction data obtained from XFEL (SFX) and those from synchrotron sources (SSX), we exploited two well established protein systems: i) the relatively radiation tolerant fluoroacetate dehalogenase (FACD) and ii) the highly radiation sensitive myoglobin (MB). FACD is a homodimeric protein from the soil-bacterium *Rhodopseudomonas palustris* and one of the very few enzymes that can cleave fluorine carbon bonds. FACD shows half-of-the-sites reactivity and incorporates a covalent substrate-enzyme intermediate into its S<sub>N</sub>2 substitution mechanism<sup>15,20,21</sup>. MB is a monomeric heme protein involved in the oxygen exchange of mammalian muscle tissue. It was the first protein that had its three-dimensional structure determined and is known to be highly radiation sensitive due to the iron center of its heme chromophore<sup>22,23</sup>.

For sample delivery, we made use of our previously described fixed target approach<sup>16,24–27</sup>. Briefly, approximately 25,000 protein micro-crystals are mounted on a silicon support chip in random orientations. These chips are then raster-scanned in the X-ray beam by means of a fast and accurate closed-loop piezo translation stage system. This system allows the collection of data from up to 4 chips (i.e., approximately 100,000 microcrystals) per hour. The modular architecture of this system enables its use at different synchrotron and XFEL end-stations<sup>14,16,19,24–28</sup>. For the present comparison, data were collected at SACLA at the RIKEN SPRING-8 center and the EMBL beamline P14 at the PETRA-III synchrotron at DESY. Importantly, in addition to applying the same data-collection hardware we used protein crystals from the same crystallization batches in both experiments. The micro-crystals of both proteins were isomorphous and of the same physical dimensions. This should reduce any crystal quality differences that may influence overall data quality parameters. Moreover, to reduce any further mismatches in the analysis, the data were processed with the same software package, CrystFEL 0.8.0<sup>29</sup>.

### Global data quality comparison

The global data quality parameters for both proteins indicate close equivalence of the SSX and SFX data over the whole resolution range (**Table I, Figure 1**). For both proteins, signal to noise ratio (SNR), multiplicity,  $R_{\text{split}}$ , and overall completeness are almost identical for SFX and SSX data (**Figure 2**). Only minor resolution differences can be observed, as indicated by the resolution-shell half-set correlation coefficients ( $CC_{1/2}$ ),  $R_{\text{split}}$ , and refinement  $R_{\text{free}}$  values. While the MB dataset shows virtually no differences in B-factors between the SSX and SFX dataset, the B-factors are slightly lower for the SFX

dataset of FAcD (**Table I**). This may correlate with the marginally better data statistics in the high-resolution shell of FAcD as made evident by SNR and  $R_{\text{split}}$  values. It is conceivable that the FAcD crystals would have diffracted to a slightly higher resolution in the XFEL beam relative to the synchrotron beam as indicated by the B-factors. The current resolution cutoff was defined by the detector distance at the XFEL. In addition to the integration statistics, the general quality of the refined structures needs to be assessed. Globally this can be achieved by comparing the free  $R$ -values ( $R_{\text{free}}$ ) of the SSX and SFX structures (**Table I**). The absolute percentage differences between the overall  $R_{\text{free}}$  values for FAcD and MB are negligible, while for the high-resolution shell absolute  $R_{\text{free}}$  value differences of  $\sim 4.6\%$  and  $\sim 1.8\%$  can be observed, respectively. On a per-residue basis, model differences between the SFX and SSX structures can conveniently be assessed via the root-mean-square deviation (r.m.s.d.) and  $B$ -factor values. The all-atom r.m.s.d. of the SFX and SSX structures for FAcD is  $0.12 \text{ \AA}$ , compared to a coordinate error of  $0.20 \text{ \AA}$ . For MB, the all-atom r.m.s.d. of the SFX and SSX structures is  $0.13 \text{ \AA}$ , compared to a coordinate error of  $0.23 \text{ \AA}$ .



**Figure 1: Global data quality parameters as function of resolution.**

**(a)** half-set correlation coefficient for SSX (blue, open circles) and SFX (orange, open triangles) for FAcD and Myoglobin (MB) data sets. The green curve displays the relative correlation coefficient between SFX and SSX data. The datasets were limited to the same number of diffraction images (10,000 for FAcD and 8,000 for myoglobin) and the same resolution cutoff was applied to SSX and SFX data. **(b)** corresponding refinement  $R_{\text{free}}$  values.

**Table I: Data collection and refinement statistics.**

	<b>FAcD-SSX</b>	<b>FAcD-SFX</b>	<b>MB-SSX</b>	<b>MB-SFX</b>
<b>accession code</b>	7A42	7A43	7A44	7A45
<b>resolution range</b>	33.08 - 1.75 (1.813 - 1.75)	33.08 - 1.75 (1.813 - 1.75)	31.47 - 1.75 (1.813 - 1.75)	31.47 - 1.75 (1.813 - 1.75)
<b>space group</b>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
<b>unit cell</b>	41.9 79.9 84.8 90 103.4 90	41.9 79.9 84.8 90 103.4 90	37.9 47.9 83.5 90 90 90	37.9 47.9 83.5 90 90 90
<b>unique reflections</b>	54827 (5456)	54828 (5457)	15863 (1492)	15891 (1520)
<b>multiplicity</b>	139.52 (74.0)	101.92 (66.0)	114.79 (65.4)	134.40 (76.2)
<b>completeness (%)</b>	100.00 (100.00)	100.00 (100.00)	99.53 (95.08)	99.72 (97.19)
<b>SNR</b>	4.18 (1.19)	3.88 (2.06)	5.54 (1.62)	5.65 (1.61)
<b>Wilson B-factor</b>	18.73	13.83	22.21	24.83
<b>R-split</b>	19.08 (85.41)	21.70 (48.78)	14.27 (77.14)	13.78 (74.6)
<b>CC<sub>1/2</sub></b>	95.34 (52.52)	93.10 (72.32)	97.33 (50.03)	97.35 (40.54)
<b>reflections used in refinement</b>	54827 (4396)	54828 (4873)	15866 (1231)	15332 (1280)
<b>reflections used for R-free</b>	1852 (151)	1838 (172)	1522 (125)	1529 (127)
<b>R<sub>work</sub></b>	0.1570 (0.2889)	0.1627 (0.2232)	0.1730 (0.2770)	0.1683 (0.2551)
<b>R<sub>free</sub></b>	0.1929 (0.3064)	0.1995 (0.2605)	0.2099 (0.3208)	0.2085 (0.3025)
<b>number of non-hydrogen atoms</b>	5093	5051	1371	1384
<b>macromolecules</b>	4791	4743	1247	1248
<b>ligands</b>	2	2	50	50
<b>solvent</b>	300	306	74	86
<b>protein residues</b>	595	595	154	154
<b>RMS (bonds)</b>	0.009	0.012	0.016	0.016
<b>RMS (angles)</b>	1.08	1.25	1.32	1.34
<b>Ramachandran favored (%)</b>	97.80	97.97	98.03	98.03
<b>Ramachandran allowed (%)</b>	2.20	2.03	1.97	1.97
<b>Ramachandran outliers (%)</b>	0.00	0.00	0.00	0.00
<b>rotamer outliers (%)</b>	0.41	0.63	1.55	1.54
<b>clashscore</b>	3.92	3.01	4.59	4.21
<b>average B-factor macromolecules</b>	22.87	16.53	26.50	28.98
<b>ligands</b>	22.28	15.87	25.97	28.32
<b>solvent</b>	35.54	28.12	24.17	25.87
	32.22	26.72	37.11	40.36

*Statistics for the highest-resolution shell are shown in parentheses*

By collecting room temperature data from the same batch of micron-sized crystals, using the same sample delivery device, as well as the same software for data-analysis, we have attempted to maximize

comparability of data collection at XFEL and synchrotron sources. Indeed, this experimental equivalence is reflected in very similar data statistics for the SFX and SSX data for both protein systems. The data quality indicators obtained (such as  $CC_{1/2}$ ,  $R_{\text{free}}$  and r.m.s.d.) are nearly identical for both radiation sources. However, thus far this statement only holds true for crystals of suitable size (i.e. on the order of a few micrometers), that enable data collection at synchrotrons. Clearly XFELs can make use of much higher flux densities, thereby accommodating much smaller crystals, down to the nanometer scale, which currently is outside the range of synchrotrons<sup>30</sup>. In summary, by normalizing all experimental parameters, this comparison demonstrates that in (static) serial X-ray crystallography data quality is predominantly a crystal-dependent property. Our observations reveal that neither obtainable resolution nor  $CC_{1/2}$  or  $R_{\text{free}}$ -values strongly depend on the radiation source. As a consequence, the diffraction properties of protein microcrystals can be explored at synchrotron sources to test their suitability for XFEL experiments. This would relieve XFEL beamlines of the high burden of screening time as well as increase opportunities for testing crystals as beamtime at synchrotrons is much more readily available. This not only includes static data collection, but of course also extends to time-resolved experiments, where crystals can be used for experiments at synchrotron sources for observing somewhat slower time domains.

### Scaling with data set size

Subsequently, we correlated the quality of the datasets with the number of crystals used for structure determination, randomly picking sub-sets of the total acquired datasets. Analysis of correlation coefficients,  $R_{\text{free}}$  values, SNR,  $R_{\text{split}}$ , and completeness again shows a high degree of similarity between SSX and SFX data. Although the SFX data were recorded with a bandwidth approximately 2 orders of magnitude larger than available at monochromatic synchrotron beams, the number of diffraction patterns needed to converge is similar for both datasets. While true Laue serial diffraction data require fewer images for convergence, this criterion is apparently not met by the increased bandwidth of XFEL beams compared to monochromatic synchrotron beams<sup>31</sup>.

In agreement with our previous observations, the datasets of both proteins show almost full completeness, that is over 95% in the highest resolution shell, even for a low number (< 1000) of diffraction patterns<sup>32</sup>. This indicates that for serial data collection, completeness should be the first data quality metric to be assessed before turning to other quality indicators such as  $CC_{1/2}$  or  $R_{\text{split}}$ . We argue that if the data are not fully complete other data-quality indicators are probably meaningless. For both FAcD and MB, the overall quality indicators  $CC_{1/2}$  and  $R_{\text{free}}$  start to converge to reasonable levels at approximately 5,000 diffraction images, by standard assessments (**Figure 2**). While there are only minor improvements for overall data quality indicators, the high resolution-shell quality indicators still undergo substantial improvements if more than 5,000 diffraction patterns are included. The FAcD and MB crystals used in this analysis were of monoclinic and orthorhombic symmetry, respectively – representing the two most commonly encountered crystal systems in the PDB. This suggests that for the

vast majority of cases reasonable data statistics should be achievable with a similar number of diffraction images. For triclinic symmetry (representing less than 5% of the structures in the PDB), however, a higher number of diffraction patterns would be required.

This corroborates our previous findings that protein structures and even protein small-molecule complexes obtained by SFX and SSX can be solved from ~5,000 diffraction patterns - or fewer if high symmetry permits - yielding reasonable electron density<sup>32,33</sup>. However, merging and refinement statistics, as well as the quality of electron density, clearly improve if further diffraction patterns are added. This can be especially important in time-resolved analyses to reveal low occupancy details or to distinguish subtle electron density differences from the noise.

Thus, as a rule of thumb, the present data suggest to obtain approximately 5,000 diffraction patterns, as at this point the overall dataset quality has mostly converged to reasonable quality indicators. Provided micron-sized crystals are used, SFX and SSX appear to be equivalent alternatives not only regarding the quality of the data but also with respect to the required number of diffraction patterns. Exciting developments in nano-focus X-ray beamlines combined with the increased brilliance of next generation synchrotron sources may narrow the gap of useful crystal sizes in the future, thereby also enabling the use of nano-crystals at synchrotrons<sup>34,35</sup>. Since microsecond time-resolution can also be obtained at current generation synchrotron sources, experiments on XFELs can therefore be focused on ultrafast time-resolution experiments, the time domain on which XFELs are essential.



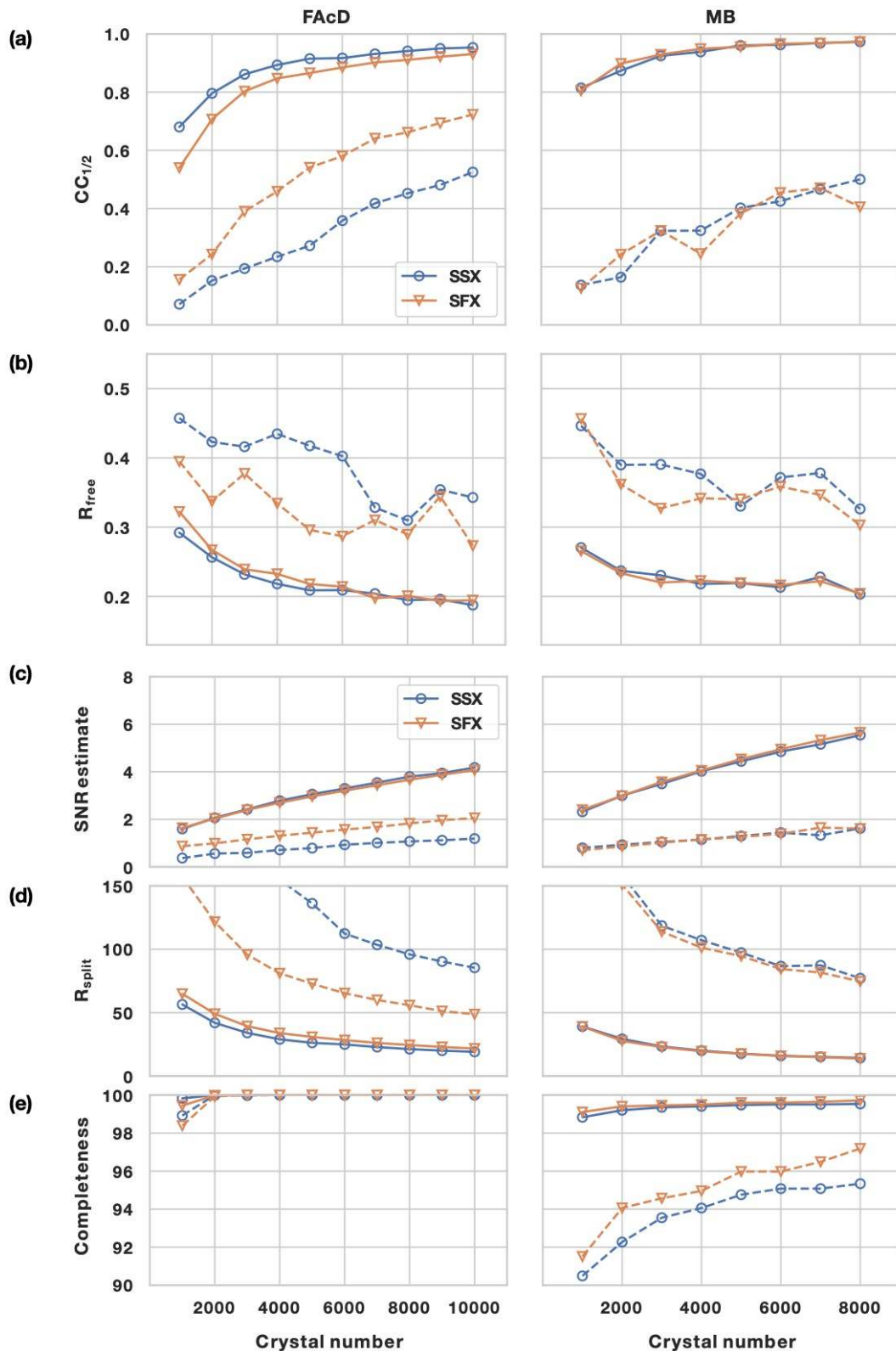


Figure 2: Global data quality parameters as function of data set size (number of crystals).

Global data quality parameters for SSX and XFEL data sets: (a)  $CC_{1/2}$  values, (b)  $R_{free}$  values, (c) SNR estimate, (d)  $R_{split}$  values, and (d) completeness. SSX data are displayed as blue, open circles, SFX data as orange, open triangles. Overall values are shown as solid lines, highest resolution shell values as dashed lines.



## Radiation damage

For traditional crystallographic experiments at synchrotrons, radiation damage is a key limiting factor in obtaining high resolution structures even at cryogenic temperatures. Damage is caused by the interaction of the X-ray photons with the electrons in the protein crystals and the energy they deposit in the sample, either caused by the photoelectric effect or inelastic scattering (Compton scattering) of X-rays leading to primary, secondary, and tertiary damage<sup>36,37</sup>. The traditional Garman limit for cryo-temperature structures is 30 MGy<sup>38</sup>. The situation is worse at room-temperature, typically encountered during serial crystallography experiments. Radiation damage manifests either globally as a loss of diffraction intensity, or specifically via chemical modification of protein residues, i.e. breaking of disulfide bonds<sup>36,39</sup>. Recently, dose limits for room-temperature SSX were inferred from lysozyme crystals, indicating a half-diffraction dose of 380 kGy and site specific damage starting to appear at approximately 80 kGy<sup>18</sup>. Notably, the authors implied that the onset of site-specific damage may occur at much lower doses for radiation sensitive proteins. This agrees well with previous findings that the dose-limit at room temperature is highly sample-dependent varying over an order of magnitude for 15 different structures<sup>40</sup>. In comparing synchrotron and XFEL sources radiation, one needs to further distinguish between direct and indirect damage. While direct damage occurs via the primary X-ray absorption at an atom in the protein, indirect damage occurs via X-ray absorption in the bulk solvent<sup>36</sup>. The latter leading to the formation of reactive oxygen species on the ns-ms timescale that can diffuse through the crystal<sup>41</sup>. Thus, the advantages of XFELs with respect to radiation damage originate in their short pulse lengths, which outrun those damage mechanisms caused by the diffusion of radicals encountered during longer exposure times at synchrotrons<sup>42,43</sup>. Due to the short pulse length (10 fs) and high intensity of the SACLA beam, established dose criteria cannot be applied to SFX data in an unmodified manner. For reference, we therefore used an updated version of RADDDOSE (RADDDOSE-XFEL), which now also offers to correctly calculate XFEL as well as SSX doses. We estimated the average dose in the exposed region, indicating that the 80 kGy dose limit was exceeded by a factor of ~2 for the SSX data (Table II)<sup>44</sup>.

**Table II: Results from RADDDOSE-XFEL**

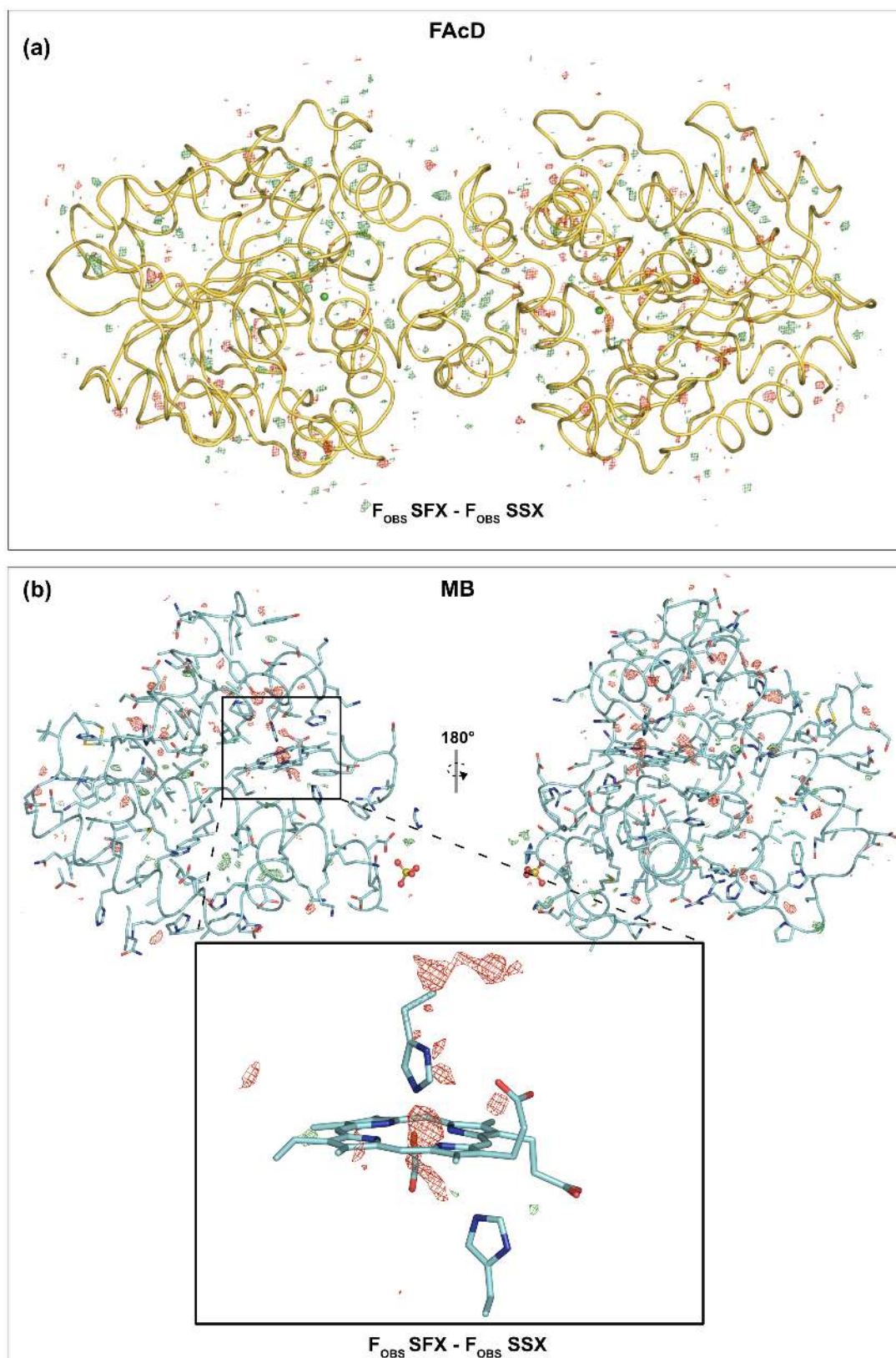
	average diffraction weighted dose (MGy)	RADDDOSE-XFEL average dose exposed region (MGy)	RADDDOSE-3D style average dose exposed region (MGy)
<b>MB-SSX</b>	0.252	-	0.184
<b>MB-SFX</b>	-	0.200	1.022
<b>FAcD-SSX</b>	0.229	-	0.167
<b>FAcD-SFX</b>	-	0.163	0.959

No global structural differences were observed in either FAcD or MB. To address the site-specific differences between the SFX and SSX data we utilized a resolution-shell scaling tool to obtain isomorphous differences densities<sup>45</sup> (*Ginn et al. – under review*). The obtained difference density maps

do not display the absolute damage but rather indicate relative differences between the two datasets (**Figure 3**). For FAcD the difference density peaks are evenly distributed over the whole structure, indicating no preferred site of radiation damage. By contrast for MB, damage is less homogeneously dispersed throughout the structure; rather local differences are primarily concentrated around certain areas such as the radiation sensitive heme-centre. This agrees with canonical radiation damage mechanisms induced by the reduction of the iron centre<sup>23</sup> (**Figure 3**). These findings further support the notion that SFX and SSX data are largely comparable and no major differences with respect to radiation damage can be observed. While SSX data are clearly not radiation damage free, the serial data collection approach apparently allows to collect data that show only minimal differences to SFX-data, which outrun the most deleterious radiation damage effects – even for very radiation damage-sensitive proteins such as myoglobin. This implies that only in very critical cases radiation damage is a crucial criterion for the choice of the radiation source. While the present SSX data were collected at doses of approximately 170 kGy, with new rapid acquisition detectors (Eiger) it is possible to collect SSX data with < 2 ms exposure time thus at much lower doses (< 5kGy), further mitigating the differences between SFX and SSX data<sup>18</sup>.

## Conclusions

Data obtained via serial crystallography approaches at synchrotron and XFEL sources can be of equivalent quality. Therefore, data quality is a crystal-dependent property and does not depend on the radiation source. We conclude that as a rule of thumb ~5,000 diffraction patterns are required to obtain data statistics and electron density maps of reasonable quality; higher numbers may be required to tease out lowly populated electron density states. Further, our findings suggest that it is possible to tailor SSX data collection, to largely avoid the effects of radiation damage thereby preserving the biological interpretability of the results. Due to the overlapping data quality, the choice of the radiation source should primarily be guided by the required time resolution for time-resolved experiments.



**Figure 3: Isomorphous difference maps for FAcD and MB, respectively.**

**Resolution shell scaled  $F_{\text{OBS SFX}} - F_{\text{OBS SSX}}$  difference maps of FAcD (a) and MB (b). Difference map peaks are homogeneously distributed over FAcD, while difference map peaks are more concentrated around the radiation-sensitive heme center in MB. Proteins are shown as cartoon representations, all maps are shown at  $\pm 3$  sigma.**

## Methods

### Protein expression, purification and crystallization

Recombinant fluoroacetate dehalogenase (FAcD) was purified from *Escherichia coli* BL21(DE3) as described<sup>20</sup>. FAcD was extracted from *E. coli* cell-free lysate using Ni-chromatography with subsequent cleavage of the His<sub>6</sub>-tag using TEV protease. Size exclusion chromatography was completed using 50 mM Tris-H<sub>2</sub>SO<sub>4</sub>, pH 8.5 and 150 mM NaCl and buffer exchanged to remove the NaCl as a final step in purification. FAcD crystals were grown in crystallization buffer (18-20 % (w/v) PEG3350, 200 mM CaCl<sub>2</sub>, and 100 mM Tris-HCl, pH 8.5). From these crystals, a microseed stock was generated at a 4-8% higher PEG3350 concentration, using a seed bead kit from Hampton research (HR2-320). Microcrystals were produced using batch crystallization; 100-200  $\mu$ l of seed stock and an equal volume of 0.5 mM FAcD solution were mixed. Within 24-72 hours crystals grew to approximately 20 x 20 x 10  $\mu$ m<sup>3</sup> in size. CO-bound sperm whale myoglobin was prepared and crystallized as described previously<sup>16,25</sup>.

### Serial synchrotron crystallography data collection

Single crystalline silicon chips provide a suitable scaffold material for fixed target crystallography. They hold randomly oriented crystals in precisely defined, bottomless wells (also called features). 100-200  $\mu$ l of a suspension of crystals were loaded onto a single chip by applying vacuum suction as described previously. Chip fabrication and sample loading process are described in detail in<sup>25-27</sup>. Serial synchrotron crystallographic (SSX) diffraction data were collected at room temperature (294 K) at EMBL beamline P14 at DESY, Hamburg, using our previously described fixed-target setup<sup>16,24</sup>. Crystals were not rotated during an exposure; still images were recorded at a wavelength of 0.976 Å and an exposure time of 37 ms using an Eiger 16M detector (Dectris, Switzerland). SSX diffraction data were processed using CrystFEL 0.8.0<sup>29</sup>.

### Serial femtosecond crystallography data collection

Serial femtosecond crystallography (SFX) diffraction data of FAcD and MB were collected at room temperature (294 K) at SACLA at SPRING-8, Japan, using the same setup as described above. SFX diffraction data were processed using CrystFEL 0.8.0<sup>29</sup>.

### Molecular replacement and refinement

The structures were determined by molecular replacement in PHASER<sup>46</sup> using PDB IDs 5K3D and 5JOM as search models for FAcD and myoglobin, respectively. Structure refinement was completed by iterative cycles of refinement in *phenix.refine*<sup>47</sup> and manual model building in COOT<sup>48,49</sup>.

### Dose and radiation damage

Doses for each dataset were calculated using the software RADDOSE-3D (v. 4.0) and its subprogram RADDOSE-XFEL, respectively<sup>44</sup>. Cuboid crystal forms with dimensions of 20 x 20 x 10  $\mu$ m<sup>3</sup> for MB, and 20 x 20 x 15  $\mu$ m<sup>3</sup> for FAcD were applied. For the SSX data, a Gaussian beam profile with a flux of 4<sup>12</sup> ph/s and a full width half maximum (FWHM) beam of 5 x 10  $\mu$ m<sup>2</sup>, with a rectangular collimation of

15 x 30  $\mu\text{m}^2$  was considered at an energy of 12.703 keV for an exposure time of 37 ms and a rotation wedge of  $0^\circ$ . RADDPOSE-3D was run with a resolution of 5 pixels per micron and default angular resolution for a wedge size of  $0^\circ$ . For the XFEL data, a Gaussian beam profile, with an uncollimated FWHM size of 4.7 x 3.4  $\mu\text{m}^2$  and an energy of 10.2 keV, a pulse energy of 0.4 mJ and an energy distribution of 0.5% (FWHM) as well as an exposure time of 10 fs was applied. RADDPOSE-XFEL was run with a resolution of 0.5 pixels per micron and simulated 1,000,000 photons. Results from 3 runs were averaged and are summarized in **Table II**. Isomorphous difference maps were calculated in analogy to a previously published approach scaling the datasets to their resolution shells<sup>45</sup>. The SFX datasets were scaled to the reference SSX datasets, by dividing the limiting resolution range into 20 equal volume bins containing similar numbers of reflections. The reflection amplitudes in each bin were multiplied by a constant scale factor, calculated to set the average amplitude in each bin equal to that of the reference. The code for shell-scaling is part of the Vagabond software suite (Ginn et al., unpublished). The resulting scaled datasets were plugged into sftools (CCP4 software suite) and difference structure factors were calculated  $F_{\text{obs-SFX}} - F_{\text{obs-SSX}}$  for both FAcD and MB respectively<sup>50</sup>.

## Data analysis

Molecular images were generated in PyMOL (Schrödinger LLC).



## References

- <sup>1</sup> S.P. Muench, S. V. Antonyuk, and H. Samar, *IUCrJ* **6**, 167 (2019).
- <sup>2</sup> H.N. Chapman, *Annu. Rev. Biochem.* **88**, 35 (2019).
- <sup>3</sup> R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, *Nature* **406**, 752 (2000).
- <sup>4</sup> I. Schlichting, *IUCrJ* **2**, 246 (2015).
- <sup>5</sup> I. Schlichting and J. Miao, *Curr. Opin. Struct. Biol.* **22**, 613 (2012).
- <sup>6</sup> C. Gati, G. Bourenkov, M. Klinge, D. Rehders, F. Stellato, D. Oberthür, O. Yefanov, B.P. Sommer, S. Mogk, M. Duszynko, C. Betzel, T.R. Schneider, H.N. Chapman, and L. Redecke, *IUCrJ* **1**, 87 (2014).
- <sup>7</sup> F. Stellato, D. Oberthür, M. Liang, R. Bean, C. Gati, O. Yefanov, A. Barty, A. Burkhardt, P. Fischer, L. Galli, R.A. Kirian, J. Meyer, S. Panneerselvam, C.H. Yoon, F. Chervinskii, E. Speller, T.A. White, C. Betzel, A. Meents, and H.N. Chapman, *IUCrJ* **1**, 204 (2014).
- <sup>8</sup> K. Moffat, *Annu Rev Biophys Chem* **18**, 309 (1989).
- <sup>9</sup> J. Hajdu, K.R. Acharya, D.I. Stuart, P.J. McLaughlin, D. Barford, N.G. Oikonomakos, H. Klein, and L.N. Johnson, *EMBO J.* **6**, 539 (1987).
- <sup>10</sup> T. Weinert, N. Olieric, R. Cheng, S. Brunle, D. James, D. Ozerov, D. Gashi, L. Vera, M. Marsh, K. Jaeger, F. Dworkowski, E. Panepucci, S. Basu, P. Skopintsev, A.S. Dore, T. Geng, R.M. Cooke, M.N. Liang, A.E. Prota, V. Panneels, P. Nogly, U. Ermler, G. Schertler, M. Hennig, M.O. Steinmetz, M.T. Wang, and J. Standfuss, *Nat. Commun.* **8**, (2017).
- <sup>11</sup> J.M. Martin-Garcia, C.E. Conrad, G. Nelson, N. Stander, N.A. Zatsepin, J. Zook, L. Zhu, J. Geiger, E. Chun, D. Kissick, M.C. Hilgart, C. Ogata, A. Ishchenko, N. Nagaratnam, S. Roy-Chowdhury, J. Coe, G. Subramanian, A. Schaffer, D. James, G. Ketwala, N. Venugopalan, S. Xu, S. Corcoran, D. Ferguson, U. Weierstall, J.C.H. Spence, V. Cherezov, P. Fromme, R.F. Fischetti, and W. Liu, *IUCrJ* **4**, (2017).
- <sup>12</sup> K.R. Beyerlein, D. Dierksmeyer, V. Mariani, M. Kuhn, I. Sarrou, A. Ottaviano, S. Awel, J. Knoska, S. Fuglerud, O. Jönsson, S. Stern, M.O. Wiedorn, O. Yefanov, L. Adriano, R. Bean, A. Burkhardt, P. Fischer, M. Heymann, D.A. Horke, K.E.J. Jungnickel, E. Kovaleva, O. Lorbeer, M. Metz, J. Meyer, A. Morgan, K. Pande, S. Panneerselvam, C. Seuring, A. Tolstikova, J. Lieske, S. Aplin, M. Roessle, T.A. White, H.N. Chapman, A. Meents, and D. Oberthuer, *IUCrJ* **4**, 769 (2017).
- <sup>13</sup> T. Weinert, P. Skopintsev, D. James, F. Dworkowski, E. Panepucci, D. Kekilli, A. Furrer, S. Brünle, S. Mous, D. Ozerov, P. Nogly, M. Wang, and J. Standfuss, *Science* **365**, 61 (2019).
- <sup>14</sup> E.C. Schulz, P. Mehrabi, H.M. Müller-Werkmeister, F. Tellkamp, A. Jha, W. Stuart, E. Persch, R. De Gasparo, F. Diederich, E.F. Pai, and R.J.D. Miller, *Nat. Methods* **15**, 901 (2018).
- <sup>15</sup> P. Mehrabi, E.C. Schulz, R. Dsouza, H.M. Müller-Werkmeister, F. Tellkamp, R.J.D. Miller, and E.F. Pai, *Science* **365**, 1167 (2019).
- <sup>16</sup> R.L. Owen, D. Axford, D.A. Sherrell, A. Kuo, O.P. Ernst, E.C. Schulz, R.J.D. Miller, and H.M. Mueller-Werkmeister, *Acta Crystallogr. Sect. D, Struct. Biol.* **73**, 373 (2017).

- <sup>17</sup> A. Ebrahim, M. V. Appleby, D. Axford, J. Beale, T. Moreno-Chicano, D.A. Sherrell, R.W. Strange, M.A. Hough, and R.L. Owen, *Acta Crystallogr. Sect. D, Struct. Biol.* **75**, 151 (2019).
- <sup>18</sup> E. de la Mora, N. Coquelle, C.S. Bury, M. Rosenthal, J.M. Holton, I. Carmichael, E.F. Garman, M. Burghammer, J.-P. Colletier, and M. Weik, *Proc. Natl. Acad. Sci. U. S. A.* **117**, 4142 (2020).
- <sup>19</sup> A. Ebrahim, T. Moreno-Chicano, M. V. Appleby, A.K. Chaplin, J.H. Beale, D.A. Sherrell, H.M.E. Duyvesteyn, S. Owada, K. Tono, H. Sugimoto, R.W. Strange, J.A.R. Worrall, D. Axford, R.L. Owen, and M.A. Hough, *IUCrJ* **6**, 543 (2019).
- <sup>20</sup> P.W. Chan, A.F. Yakunin, E.A. Edwards, and E.F. Pai, *J Am Chem Soc* **133**, 7461 (2011).
- <sup>21</sup> T.H. Kim, P. Mehrabi, Z. Ren, A. Sljoka, C. Ing, A. Bezginov, L. Ye, R. Pomès, R.S. Prosser, and E.F. Pai, *Science* **355**, eaag2355 (2017).
- <sup>22</sup> J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, and D.C. Phillips, *Nature* **181**, 662 (1958).
- <sup>23</sup> D.C. Blake, C. & Phillips, in *Proc. Symp. Biol. Eff. Ioniz. Radiat. Mol. Lev.* (Int. Atomic Energy Agency, Vienna, 1962), pp. 183–191.
- <sup>24</sup> D.A. Sherrell, A.J. Foster, L. Hudson, B. Nutter, J. O’Hea, S. Nelson, O. Pare-Labrosse, S. Oghbaey, R.J.D. Miller, and R.L. Owen, *J. Synchrotron Radiat.* **22**, 1372 (2015).
- <sup>25</sup> S. Oghbaey, A. Sarracini, H.M. Ginn, O. Pare-Labrosse, A. Kuo, A. Marx, S.W. Epp, D.A. Sherrell, B.T. Eger, Y. Zhong, R. Loch, V. Mariani, R. Alonso-Mori, S. Nelson, H.T. Lemke, R.L. Owen, A.R. Pearson, D.I. Stuart, O.P. Ernst, H.M. Mueller-Werkmeister, and R.J.D. Miller, *Acta Crystallogr. Sect. D Struct. Biol.* **72**, 944 (2016).
- <sup>26</sup> C. Mueller, A. Marx, S.W. Epp, Y. Zhong, A. Kuo, A.R. Balo, J. Soman, F. Schotte, H.T. Lemke, R.L. Owen, E.F. Pai, A.R. Pearson, J.S. Olson, P.A. Anfinrud, O.P. Ernst, and R.J.D. Miller, *Struct. Dyn.* **2**, (2015).
- <sup>27</sup> P. Mehrabi, H.M. Müller-Werkmeister, J.-P. Leimkohl, H. Schikora, J. Ninkovic, S. Krivokuca, L. Andriček, S.W. Epp, D. Sherrell, R.L. Owen, A.R. Pearson, F. Tellkamp, E.C. Schulz, and R.J.D. Miller, *J. Synchrotron Radiat.* **27**, 1 (2020).
- <sup>28</sup> P. Mehrabi, E.C. Schulz, M. Agthe, S. Horrell, G. Bourenkov, D. von Stetten, J. Leimkohl, H. Schikora, T.R. Schneider, A.R. Pearson, F. Tellkamp, and R.J.D. Miller, *Nat. Methods* **16**, 979 (2019).
- <sup>29</sup> T.A. White, R.A. Kirian, A. V. Martin, A. Aquila, K. Nass, A. Barty, and H.N. Chapman, *J. Appl. Crystallogr.* **45**, 335 (2012).
- <sup>30</sup> H.N. Chapman, P. Fromme, A. Barty, T.A. White, R.A. Kirian, A. Aquila, M.S. Hunter, J. Schulz, D.P. Deponte, U. Weierstall, R.B. Doak, F.R.N.C. Maia, A. V. Martin, I. Schlichting, L. Lomb, N. Coppola, R.L. Shoeman, S.W. Epp, R. Hartmann, D. Rolles, A. Rudenko, L. Foucar, N. Kimmel, G. Weidenspointner, P. Holl, M. Liang, M. Barthelmess, C. Caleman, S. Boutet, M.J. Bogan, J. Krzywinski, C. Bostedt, S. Bajt, L. Gumprecht, B. Rudek, B. Erk, C. Schmidt, A. Hömke, C. Reich, D. Pietschner, L. Ströder, G. Hauser, H. Gorke, J. Ullrich, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.U. Kühnel, M. Messerschmidt, J.D. Bozek, S.P. Hau-Riege, M. Frank, C.Y. Hampton, R.G.



Sierra, D. Starodub, G.J. Williams, J. Hajdu, N. Timneanu, M.M. Seibert, J. Andreasson, A. Rocker, O. Jönsson, M. Svenda, S. Stern, K. Nass, R. Andritschke, C.D. Schröter, F. Krasniqi, M. Bott, K.E. Schmidt, X. Wang, I. Grotjohann, J.M. Holton, T.R.M. Barends, R. Neutze, S. Marchesini, R. Fromme, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, I. Andersson, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, and J.C.H. Spence, *Nature* **470**, 73 (2011).

<sup>31</sup> A. Meents, M.O. Wiedorn, V. Srajer, R. Henning, I. Sarrou, J. Bergtholdt, M. Barthelmess, P.Y.A. Reinke, D. Dierksmeyer, A. Tolstikova, S. Schaible, M. Messerschmidt, C.M. Ogata, D.J. Kissick, M.H. Taft, D.J. Manstein, J. Lieske, D. Oberthuer, R.F. Fischetti, and H.N. Chapman, *Nat. Commun.* **8**, (2017).

<sup>32</sup> T. Moreno-Chicano, A. Ebrahim, D. Axford, M. V. Appleby, J.H. Beale, A.K. Chaplin, H.M.E. Duyvesteyn, R.A. Ghiladi, S. Owada, D.A. Sherrell, R.W. Strange, H. Sugimoto, K. Tono, J.A.R. Worrall, R.L. Owen, and M.A. Hough, *IUCrJ* **6**, 1074 (2019).

<sup>33</sup> E.C. Schulz, J. Kaub, F. Busse, P. Mehrabi, H.M. Müller-Werkmeister, E.F. Pai, W.D. Robertson, and R.J.D. Miller, *J. Appl. Crystallogr.* **50**, 1773 (2017).

<sup>34</sup> E. V. Beale, A.J. Warren, J. Trincão, J. Beilsten-Edmands, A.D. Crawshaw, G. Sutton, D. Stuart, and G. Evans, *IUCrJ* **7**, 500 (2020).

<sup>35</sup> D. Laundry, K. Sawhney, I. Nistea, S.G. Alcock, I. Pape, J. Sutter, L. Alianelli, and G. Evans, *Rev. Sci. Instrum.* **87**, 051802 (2016).

<sup>36</sup> E.F. Garman, *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 339 (2010).

<sup>37</sup> R. Henderson, *Q. Rev. Biophys.* **28**, 171 (1995).

<sup>38</sup> R.L. Owen, E. Rudiño-Piñera, and E.F. Garman, *Proc. Natl. Acad. Sci.* **103**, 4912 (2006).

<sup>39</sup> J.M. Holton, *J. Synchrotron Radiat.* (2009).

<sup>40</sup> R.M.F. Leal, G. Bourenkov, S. Russi, and A.N. Popov, *J. Synchrotron Radiat.* **20**, 14 (2013).

<sup>41</sup> S. Le Caer, *Water* **3**, 235 (2011).

<sup>42</sup> J. Tenboer, S. Basu, N. Zatsepin, K. Pande, D. Milathianaki, M. Frank, M. Hunter, S. Boutet, G.J. Williams, J.E. Koglin, D. Oberthuer, M. Heymann, C. Kupitz, C. Conrad, J. Coe, S. Roy-Chowdhury, U. Weierstall, D. James, D. Wang, T. Grant, A. Barty, O. Yefanov, J. Scales, C. Gati, C. Seuring, V. Srajer, R. Henning, P. Schwander, R. Fromme, A. Ourmazd, K. Moffat, J.J. Van Thor, J.C.H. Spence, P. Fromme, H.N. Chapman, and M. Schmidt, *Science* **346**, 1242 (2014).

<sup>43</sup> K. Nass, *Acta Crystallogr. Sect. D Struct. Biol.* **75**, 211 (2019).

<sup>44</sup> J.L. Dickerson, P.T.N. McCubbin, and E.F. Garman, *J. Appl. Crystallogr.* **53**, 549 (2020).

<sup>45</sup> H.M. Ginn and D.I. Stuart, *Acta Crystallogr. Sect. D Struct. Biol.* (2016). Volume 72| Part 6| June 2016| Pages 817-822

<sup>46</sup> A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read, *J. Appl. Crystallogr.* **40**, 658 (2007).

<sup>47</sup> P.D. Adams, P. V Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C.

Richardson, J.S. Richardson, T.C. Terwilliger, and P.H. Zwart, *Acta Crystallogr D Biol Crystallogr* **66**, 213 (2010).

<sup>48</sup> P. Emsley and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126 (2004).

<sup>49</sup> P. Emsley, B. Lohkamp, W.G. Scott, and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486 (2010).

<sup>50</sup> M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, and K.S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235 (2011).

## Author Contributions

E.C.S and P.M. initiated the project; P.M., T.M. and A.K. prepared the protein and crystals; E.C.S, P.M., D.v.S, H.M.-W., A.M., H.S, H.G. D.S., D.A., R.O., F.T., A.K., J.E.B., T.M, W.-L.O., J.B., O.P.E., E.F.P., R.J.D.M. collected the data; E.C.S, P.M., R.B., H.G., B.T.E., and A.M. processed the data; E.C.S, P.M., R.B., D.v.S. and G.B. analyzed the data; E.C.S, P.M. and R.B. wrote the manuscript. All authors discussed and corrected the manuscript.

## Acknowledgements

The SSX data was collected at beamlines P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We would like to thank Drs. A.R. Pearson (University of Hamburg), S. Horrell (Diamond Light Source), M. Agthe, and T.R. Schneider (EMBL Hamburg) for their assistance in using the beamline, and for their continuous support and helpful discussions in the implementation and improvement of SSX instruments at EMBL Hamburg. We acknowledge Dr. T. White for most helpful discussions on data processing. The XFEL experiments were performed at the BL3 of SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2016A8036 and 2016B8052). We thank the beamline scientists for their excellent support.

## Funding

Support was provided by the Max Planck Society and by the Cluster of Excellence 'The Hamburg Centre for Ultrafast Imaging' of the Deutsche Forschungsgemeinschaft (DFG) - EXC 1074 - project ID 194651731 (R.J.D.M.), the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007–2013) under REA grant no. 623994 (H.M.M.-W.), the Joachim Herz Foundation (Biomedical Physics of Infection) (E.C.S., R.J.D.M.), grant RGPIN-2015-04877 from the National Science and Engineering Council of Canada, and the Canada Research Chairs Program (E.F.P.), the Canada Excellence Research Chairs Program (O.P.E.). O.P.E. holds the Anne and Max Tanenbaum Chair in Neuroscience at University of Toronto, and O.P.E. and R.J.D.M. are CIFAR fellows.

## Competing interests

The author(s) declare no competing interests.