Kilohertz droplet-on-demand serial femtosecond crystallography at the European XFEL station FXE

Special Collection: Celebrating the work and achievements of Keith Moffat

Samuel Perrett ; Alisia Fadini; Christopher D. M. Hutchison ; Sayantan Bhattacharya ;
Cade Morrison ; Oleksii Turkot; Mads Bregenholt Jakobsen ; Michael Größler; José Licón-Saláiz;
Florian Griese; Samuel Flewett; Joana Valerio; Joachim Schulz ; Mykola Biednov; Yifeng Jiang ;
Huijong Han; Hazem Yousef; Dmitry Khakhulin; Christopher Milne ; Anton Barty; Jasper J. van Thor ✉

Check for updates

View Online

Export Citation

Structural Dynamics

Special Topic:

Artificial Intelligence and Structural Science

Guest Editors: Charles Carter and George Phillips

Submit Today!

co-published by

AIP Publishing

Kilohertz droplet-on-demand serial femtosecond crystallography at the European XFEL station FXE

Samuel Perrett,[1] (ID) Alisia Fadini,[1,a)] Christopher D. M. Hutchison,[1,2] (ID) Sayantan Bhattacharya,[1] (ID) Cade Morrison,[1] (ID) Oleksii Turkot,[3] Mads Bregenholt Jakobsen,[4] (ID) Michael Größler,[4] José Licón-Saláiz,[4] Florian Griese,[4,5,6] Samuel Flewett,[4] Joana Valerio,[3] Joachim Schulz,[3] (ID) Mykola Biednov,[3] Yifeng Jiang,[3] (ID) Huijong Han,[3] Hazem Yousef,[3] Dmitry Khakhulin,[3] Christopher Milne,[3] (ID) Anton Barty,[4] and Jasper J. van Thor[1,b)] (ID)

AFFILIATIONS

[1]Department of Life Sciences, Faculty of Natural Sciences, Imperial College London, London SW7 2AZ, United Kingdom
[2]Central Laser Facility, Rutherford Appleton Laboratory, Didcot OX11 0QX, United Kingdom
[3]European XFEL, Holzkoppel 4, 22869 Schenefeld, Germany
[4]Center for Data and Computing in Natural Sciences (CDCS), Notkestrasse 10, D-22607 Hamburg, Germany
[5]Section for Biomedical Imaging, University Medical Center Hamburg-Eppendorf, D-20246 Hamburg, Germany
[6]Institute for Biomedical Imaging, Hamburg University of Technology, D-21073 Hamburg, Germany

Note: Paper published as part of the special topic on Celebrating the work and achievements of Keith Moffat.
a)**Present address:** Cambridge Institute for Medical Research, University of Cambridge, The Keith Peters Building, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0XY, United Kingdom
b)**Author to whom correspondence should be addressed:** j.vanthor@imperial.ac.uk

ABSTRACT

X-ray Free Electron Lasers (XFELs) allow the collection of high-quality serial femtosecond crystallography data. The next generation of megahertz superconducting FELs promises to drastically reduce data collection times, enabling the capture of more structures with higher signal-to-noise ratios and facilitating more complex experiments. Currently, gas dynamic virtual nozzles (GDVNs) stand as the sole delivery method capable of best utilizing the repetition rate of megahertz sources for crystallography. However, their substantial sample consumption renders their use impractical for many protein targets in serial crystallography experiments. Here, we present a novel application of a droplet-on-demand injection method, which allowed operation at 47 kHz at the European XFEL (EuXFEL) by tailoring a multi-droplet injection scheme for each macro-pulse. We demonstrate a collection rate of 150 000 indexed patterns per hour. We show that the performance and effective data collection rate are comparable to GDVN, with a sample consumption reduction of two orders of magnitude. We present lysozyme crystallographic data using the Large Pixel Detector at the femtosecond x-ray experiment endstation. Significant improvement of the crystallographic statistics was made by correcting for a systematic drift of the photon energy in the EuXFEL macro-pulse train, which was characterized from indexing the individual frames in the pulse train. This is the highest resolution protein structure collected and reported at the EuXFEL at 1.38 Å resolution.

INTRODUCTION

X-ray Free Electron Lasers (XFELs) have proven to be a powerful tool for studying chemical and biological mechanisms through crystallography.[1–7] XFELs deliver femtosecond pulses with peak brilliance over a billion times more intense than that of their synchrotron counterparts.[8] The short duration of the pulses allows tracking of structural and coherent dynamics on a sub-picosecond level,[1,9] for many photo-reactions where the initial product-determining reaction step occurs.

However, due to the high peak brilliance, these X-rays destroy the sample upon exposure. The consequent necessity to replace the crystal after each shot means that XFEL crystallography is typically conducted using a serial approach, where crystals are rapidly introduced and replaced for each x-ray pulse.[10] High-resolution structure determination through serial femtosecond crystallography (SFX) requires large datasets ∼10 000–100 000 frames. This translates to significant sample consumption, which can further increase by up to two orders of

magnitude for sample delivery techniques with low hit rate, a factor that is prohibitive for targets that cannot be easily produced in such quantities.

The advent of superconducting XFELs, such as the European XFEL (EuXFEL) and Linac Coherent Light Source (LCLS) II-HE, which operates at megahertz (MHz) repetition rate, promises faster data collection and better signal-to-noise for serial crystallography.[11] EuXFEL is currently the highest repetition rate XFEL, capable of delivering up to 27 000 pulses per second in a "burst mode," with a train duration of 600 μs at a repetition rate of 10 Hz.[12] LCLS-II is expected to start accepting users in the near future with a true megahertz repetition rate. Once operational, it is anticipated that the x-ray source will not be the limiting factor in SFX measurements. Instead, the focus will shift to improving detector technologies, optical excitation methods, and sample delivery systems to fully leverage the capabilities of the latest generation of x-ray sources.[13,14]

Crystal delivery methods for SFX experiments can broadly be split into two categories. First, fixed-target methods place crystals on solid supports and move them through the beam. These boast high hit rates but possess their own limitations, such as device loading time and potential sample evaporation issues.[15–19] Although fixed-targets can achieve high hit rates, they cannot keep up with MHz repetition rate of the latest FELs due to physical limits on translating a fixed target and the finite physical size. Second, liquid phase injectors, such as grease injectors, offer low sample consumption and high hit rates.[20,21] However, grease injectors are not suitable for MHz XFELs due to their slower delivery speeds and inability to refresh the sample volume. An alternative liquid target is the gas dynamic virtual nozzle (GDVN), which is the only method at present capable of delivering samples at >100 Hz repetition rates.[22,23] GDVNs deliver a high velocity stream of crystal slurry to the XFEL beam focused by helium gas. Variations like double flow-focusing nozzles (DFFNs)[24] and co-flow systems offer improved stability and reduced waste. The low hit rate and rapid jet velocities required to replace the crystals on a shot-to-shot basis, on the other hand, lead to an extremely low index rate of all frames (<5%). These issues are exacerbated at the EuXFEL, where, with continuous GDVN injection, the majority of the sample arrives between the trains and is consequently wasted. The very high sample consumption renders MHz SFX unfeasible for many biologically relevant targets that are not available in the quantities required, typically grams. Sample delivery at FELs has been previously presented and discussed in detail.[25–28] Droplet-on-demand (DoD) injection methods have proved very successful at delivering high data rates with small sample consumption at ~100 Hz FELs.[1,29–34] While the repetition rate of this method is physically capped by the acoustic wave generation and dampening speed of the injector, this technology can be theoretically be advanced to MHz operation.[35,36] The feasibility of this method at the latest high repetition rate FELs is yet to be demonstrated past 100 Hz.

Here, we report kHz droplet-on-demand serial crystallography, yielding rates of indexed crystals per second comparable with published MHz GDVN experiments,[27,37–40] with a two order of magnitude reduction in sample consumption. The data were collected at the Femtosecond X-ray Experiments (FXE) beamline of the EuXFEL, marking the first report of serial protein crystallographic measurements at this beamline. The use of the Large Pixel Detector (LPD) in protein crystallography for the first time resulted in achieving higher resolution limits than any previous protein crystallographic experiments at the EuXFEL. The implementation of supervised learning algorithms for masking, along with efficient data handling techniques, streamlined the process of managing protein SFX data with this detector. By developing and optimizing these processing methods, we were able to maximize data quality from the LPD, leading to the highest resolution protein structures obtained from the EuXFEL. Our work demonstrates the feasibility of collecting high-resolution protein crystallography data at MHz FELs with sample volumes not uncommon in biochemistry laboratories, lowering the barrier to entry for the emerging field of SFX.

METHODS

Droplet-on-demand injection design

Droplet-on-demand (DoD) injection techniques utilize a piezo actuator to create an acoustic wave in the sample medium. This negative pressure wave travels down the nozzle, pushing the sample's meniscus out of the orifice. The deformation of the meniscus then breaks off, forming a droplet due to surface tension forces. The actuation of the piezoelectric material, controlled by the supplied electrical waveform, allows for tuning of the acoustic wave and the resultant droplet characteristics. Additionally, the properties of the sample medium, such as crystal density and viscosity, are crucial factors influencing droplet formation.[41,42] Solutions with viscosities between 0.5 and 40 cP can be jetted at these higher rates. We have jetted solutions of PEG 6000 (25%) and sodium malonate (3.2M), common carriers for biological crystals. The DoD injector was an adapted inkjet printing system manufactured by MicroFab Technologies, consisting of the JetDrive controller (CT-M3-02), pressure controller (CT-PT-21-1), printhead assembly (PH-47-AB), and 80 μm microdispensing device (MJ-AB-15-80 DLC). A schematic diagram of the setup can be found in supplementary information Fig. A.

The major practical concern of EuXFEL is the macro-pulse (train) structure of the X-rays.[43] The precise timing control of the DoD method significantly reduces sample usage by allowing the injection to be paused during the >99% of total time that the X-ray is idle. However, for a usable data rate in SFX experiments (exceeding 10 Hz), it is crucial to achieve diffraction from multiple pulses within a single train. To meet this requirement, droplets must be generated at frequencies ranging from kHz to MHz and contain sufficient crystal density.

The primary challenge in this process arises from fluid dynamics. Specifically, the fundamental limitation to the injection rate stems from the damping effect of the acoustic wave in the sample medium and the speed of droplet ejection.[36] Ensuring the dampening of the previous acoustic wave is critical for stable jetting, as it guarantees that each subsequent droplet is unaffected by its predecessor. To maximize the intra-pulse injection rate, our method uses a combination of two distinct approaches. First, high voltage long electrical waveforms allow for an ejection of an elongated droplet, here named "Knackwurst" (a Hamburg sausage). The Knackwurst can then be hit multiple times over the train by matching the pulse rate to the droplet velocity. The longer, elongated shape of these droplets affords a larger cross section than a smaller spherical droplet. Variations in the crystal density, precipitation, and debris resulting from droplet explosion over the data collection period can induce positional changes in a droplet within the pulse sequence. The increased volume of the Knackwurst droplets

provides a larger margin of error, contributing to a sustained high hit rate. Second, four of these Knackwurst droplets are ejected over a train, each triggered by a distinct electrical waveform (Fig. 1). The XFEL had an intra pulse repetition rate of 47 kHz, meaning that data were collected at an effective repetition rate of 160 Hz. We note that future work could extend this repetition rate, as the MicroFab injector specifications allow for producing droplets at 45 kHz, which could be subjected to multiple hits.

The four Knackwursts were generated as follows. A Transistor-Transistor Logic (TTL) trigger from the XFEL was routed through a
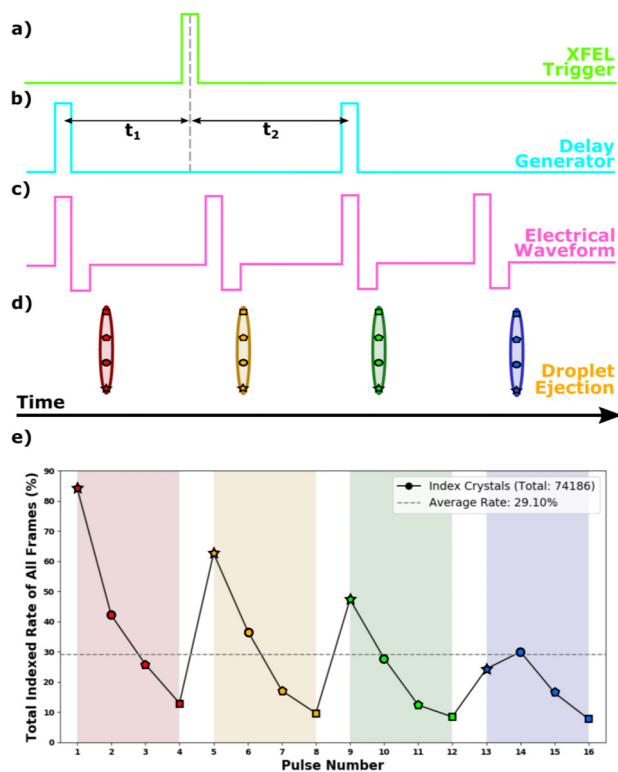

FIG. 1. Knackwurst schematic depicting the process of triggering and generating waveforms for droplet formation. (a) The sequence begins with a trigger from the Free-Electron Laser (FEL). (b) This trigger is followed by two additional triggers generated by a delay generator. The first of these delayed triggers occurs before the original FEL trigger. (c) These two delayed triggers are then fed into an electrical waveform generator, which is electronically controlled by MicroFab software. The electrical waveform generator produces a specific voltage pattern: initially, a large positive voltage of $+80$ V is applied, followed by a short negative voltage of -40 V. This voltage pattern is key in forming a droplet with a Knackwurst shape. Each Transistor-Transistor Logic (TTL) signal from the delay generator is responsible for creating two electrical waveforms, resulting in a total of four waveforms. These four electrical waveforms are all independently controlled and are crucial in the operation of the piezo-actuator. (d) The piezo-actuator, in response to these waveforms, ejects four droplets. Each of these droplets is then hit four times by the XFEL pulses. (e) The resulting hit rate for each hit within the Knackwursts. A decline in the total indexing rate is evident as the hit rate decreases, attributed to increased instability with each additional hit. A slight reduction from Knackwurst-to-Knackwurst is observed due to incomplete dampening of the previous acoustic wave affecting subsequent Knackwurst ejection. The average hit rate over the collection period was 29%. All FEL pulses hit a Knackwurst none missed between the injection.

delay generator, which produced two independent TTL triggers. The first trigger was timed to precede the XFEL TTL, initiating the first Knackwurst injection before the first XFEL pulse—effectively syncing it with the preceding FEL pulse. The minor shot-to-shot drift is negligible for these injection timescales. These two independent triggers from the delay generator were used to trigger the JetDrive III signal generator. For each trigger, two separate voltage signals were sent to the piezo-actuator printer assembly (Fig. 1). Consequently, the printer assembly received four electrical waveforms, resulting in the output of four consecutive Knackwursts. A rapid-rise-and-fall waveform, which quickly ascends (5 μs) to a high positive voltage (70–90 V), briefly (5 μs) maintains this level and then swiftly descends (5–10 μs) to a negative value (-40 V) for a short period (10–20 μs), before rapidly returning (5 μs) to 0 V. This waveform generates the elongated droplets (knackwursts) at a high repetition rate.

A crystallization buffer was used to partially fill the sample reservoir and all tubing. Subsequently, the crystalline sample (crystal density 18%) was reverse-loaded through the nozzle by applying a -0.5 psi backing pressure, serving as a final filtration step. Typical 100 μl of sample was loaded, under stable conditions enough for $\sim$2 h. Care was taken to avoid introducing any air bubbles into the system. Anecdotally, it was found that maintaining a small air bubble just inside the 80 μm exit hole of the jet did aid in jetting less viscous buffers. This is assumed to be due to a dampening effect that disrupts the acoustic wave traveling in reverse direction after droplet ejection. A slight backing pressure was applied throughout jetting at -0.2 psi, to maintain the meniscus flush without the orifice of the nozzle. Jetting was maintained during the hutch closing procedure to prevent clogging. However, to minimize sample wastage, a separate waveform scheme operating at 10 Hz was employed. The sample concentration was adjusted to approximately 3–6×10^7 Crystals/ml, varying with the sample, to ensure consistent hits without clogging the nozzle. Live feedback on the hit rate was provided by EXtra-Foam, with a consistent average hit rate of 30%–50% being achievable after some optimization (supplementary information D).[44] The sample was replenished when the hit rate dropped below 15% and continued operation without disruption was typically possible for 1-2 h. Fine tuning of trigger delays and spatial position of the jet was done via the live feedback, and this was adjusted over the run accounting for any debris buildup on the jet. Live imaging of the jetting process, as shown in Fig. 2, was conducted using a FLIR Blackfly S (BFS-U3-04S2M-CS) camera focused on the jet. Additionally, a small LED on a variable delay was employed as a strobe to record the droplet injection at various time delays (Fig. 2). Initially, the alignment of the DoD was achieved by increasing the X-ray flux to visually observe the explosion. Subsequently, data collection was conducted at a lower flux to minimize any shockwave-induced crystal deformities, which have been reported at high repetition rates.[45]

The four distinct Knackwursts allow examination of potential shock-wave-induced alterations in crystal dimension and quality of diffraction. GDVN sample delivery at 1.1 MHz repetition rates have shown no such differences.[37] However, at the higher repetition rate of 4.5 MHz, there has been significant shock induced damage reported.[45] The GDVN and DoD injection conditions differ in the thickness of the jet as well as the velocity. A distance of $\sim$100 μm separates subsequent hits within a Knackwurst (total length $\sim$600 μm), which corresponds to a volume of 0.09 nl. We note that this is difficult to estimate with
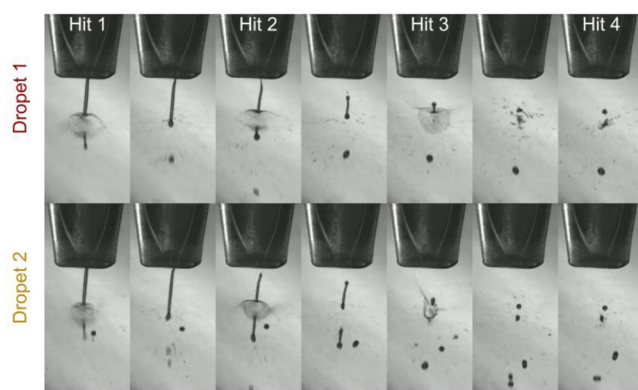
FIG. 2. Droplet-on-demand injection scheme. Top: Images from the first Knackwurst (elongated droplet) ejection sequence, wherein the Knackwurst undergoes four successive hits (labelled hit 1 to 4), each marked by a notable explosion. Each frame is shown at intervals of 10.5 μs. Bottom: A similar sequence for the second Knackwurst, in total four Knackwursts were ejected, each experiencing four hits. Consequently, 16 X-ray pulses were delivered over the macro-pulse at a frequency of 47 kHz. Note images at higher flux for droplet alignment and only display first two out of four Knackwursts.

varying dynamics of the droplet formation and injection due to surface tension and perturbation by the X-rays. With the DoD delivery method operating at 47 kHz, we observed no shot-to-shot changes within a single droplet, as discussed from the crystallographic results below. Figure 2 displays the images of the injection scheme with successive hits on the same droplet demonstrating a decreasing hit rate for shots within each Knackwurst, attributed to the random instability caused by the explosion from the previous hit(s). However, when averaged over the entire injection scheme, each pulse demonstrated an average indexing rate of 29% (of all frames), which is significantly higher compared to other sample delivery methods [GDVN $\sim$ 5% (Ref. 46)]. This substantial increase in the hit rate, in comparison with GDVN, enables a higher effective data rate, or number of indexed patterns per second, surpassing all but one of the previous lysozyme MHz SFX experiments conducted at the EuXFEL, shown in Table II.[37–40,47]

FXE data collection instrumentation

Data were collected at Femtosecond X-ray Experiments (FXE) endstation at the European XFEL.[48,49] Sample delivery was as described as above and in the main text. The sample environment enclosure was purged with helium before data collection, beyond this there was approximately $\sim$15 cm of air before the detector. The photon energy of the XFEL was $\sim$9.3 keV, and up to 2.6 mJ pulse energy, 16 pulses were provided within the train at a repetition rate of 47 kHz at 20% transmission. The x-ray beam spot size was $10 \times 11 \, \mu m$ at interaction region (supplementary Fig. P). The data collection was performed using the Large Pixel Detector (LPD),[50] which features a pixel size of $500 \times 500 \, \mu m$ and three gain stages, at a detector distance of 0.235 m. A number of additional data processing steps were required nonstandard to other detectors, which are discussed below.

Data analysis

EXtra-Xwiz[51] was used to process data with CrystFEL 10.2.[52–54] Peak finding was performed by "peakfinder8"[55] and integration carried

out by XGANDALF (extended gradient descent algorithm for lattice finding),[56] with multi crystal indexing turned on. Additionally, the nonstandard "integration = rings-cen" was used to recenter the indexing area, with ring integration radii of 1,2,3 due to the large pixels and detector size of the LPD. Several iterations of "geoptimiser" and "detector-shift" were employed to refine the detector geometry. Due to the LPD's large size and the large sample to detector distance, errors in detector geometry have a more pronounced effect on indexing when compared to other crystallographic detectors at different endstations.

To identify "hot pixels" on the LPD, as previously reported by Aleksich *et al.*,[57] a series of masking procedures was performed. Briefly, before data collection, a crude mask was created based on peakfinder results from water and lysozyme crystals to eliminate peaks significantly above the average occurrence. Additionally, indexing results using the MOSFLM algorithm in CrystFEL[54] on lysozyme were utilized. Pixels causing significant deviations in indexing at the extremes of standard lysozyme unit cell distributions were removed. This initial masking facilitated a reasonably accurate live hit rate determination through EXtra-Foam.[44] Post-beamtime, for larger datasets, peakfinder and indexing results were revisited to remove pixels that showed peaks significantly above average. The mask is generated by an isolation forest algorithm. Specifically, we used the implementation from scikitlearn.[58] The features that were fed to the algorithm are the differences between the mean, standard deviation, skewness, and kurtosis to its respective median value. Further detector masking was then informed by the peakogram output from CrystFEL.[54] The peakogram distributions, represented as histograms of intensity and resolution over a run, were used to identify areas with non-continuous peak accumulation. The locations of these "hot areas" were then plotted on the detector to pinpoint additional masking zones, as detailed in the supplementary information Figs. E and F. Furthermore, the average intensity recorded for each pixel across all indexed Bragg peaks helped identify areas that were not correctly intensity-calibrated. We note that only 15 panels (out of 16) on the LPD detector were operational during our experiment. Ultimately, approximately 17% of the usable part of the LPD detector were masked.

RESULTS

Crystallographic results

Initially, parameters, such as detector distance, beam center, and photon energy, were optimized to yield high indexing rates and symmetrical, low standard deviation unit cell distributions.[53] Once processed, the data were separated according to pulse number within the train to check for variations in crystallographic statistics and structures. The unit cell distributions of each pulse were fitted with a Gaussian function using the relevant tool within CrystFEL's "cell_explorer." Figure 3(a) displays the centers of these Gaussian fits in red along with their standard deviations (the raw histograms and fits can be found in supplementary information Fig. I). A noticeable trend of decreasing unit cell dimensions, of approximately 0.15% along axes a, b, and c, is observed over the course of the train. A changing unit cell dimension could be induced by a variety of factors, such as a varying sample-to-detector distance, droplet heating, or a drift in photon energy over the train. Given the large detector distance used in this experiment (236 mm), a change in the unit cell of 0.10 Å would necessitate a shift in the sample-to-detector distance of about 400 μm over the duration of the train. It is unlikely that the interaction region systematically
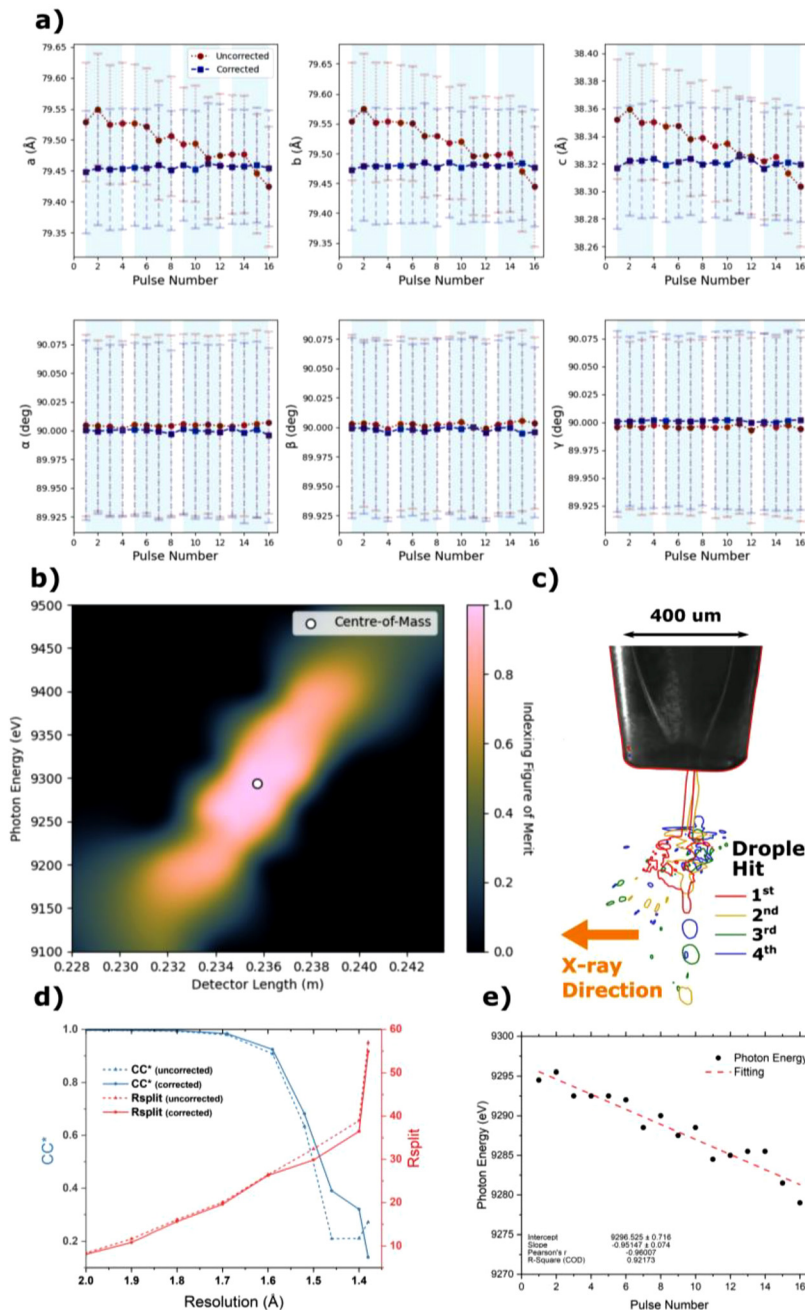
FIG. 3. Photon energy drift and subsequent correction. (a) Unit cell dimensions for pulses over injection scheme, and points are centers of fitted Gaussians with errors being standard deviation of the fit supplementary information Fig. I. Red circles are values from indexing all pulses with the same photon energy (9290 eV), and blue squares are the results after fitting a photon energy to each pulse to maintain constant unit cell dimension (corrected data). Light blue background on plots indicate groups of droplets. (b) Indexing figure of merit for various photon energies and detector distances, see the main text for details. (c) Overlayed images of sequential hits traced in their respective color and show no obvious drift in the sample to detector distance. (d) Crystallographic merging statistics, CC*, and Rsplit, for various resolution shells show an improvement after photon energy correction. (e) Photon energy used to correct unit cell distribution for each pulse, a drift of ∼15.5 eV was observed over the train.

moves by 400 μm in the direction of the XFEL beam, and also not supported by the jet imaging. Figure 3(c) traces the outlines of images of the explosion from every hit superimposed on top of each other. Since the camera was mounted perpendicular to the x-ray beam, any horizontal drift from hit-to-hit, and therefore, detector distance would be visible. However, no such trend is observed. Spatial or thermal trends from consecutive drops should reset with the introduction of each new Knackwurst and are therefore also unlikely to be the cause. Furthermore, heating would be expected to change the unit cell by

enlargement rather than reduction, while we observe a decrease in unit cell dimensions.

These observations strongly suggest a systematic drift in the photon energy of the train. Unfortunately, accurate photon energy data were not recorded during the beamtime, so only a comparative photon energy reference can be used to correct one pulse relative to another. To address this, we employed a strategy where the first pulse within the train, which had the highest number of indexed frames, was used to optimize both the detector distance and photon energy. Subsequent

pulses could then be corrected based on the parameters established during this optimization, assuming no change in the detector distance.

A figure of merit (FOM$_{indexing}$) was defined to assess the quality of specific detector distance and photon energy combinations. This metric relied on two primary factors: the number of indexed patterns (IR) and the Gaussian quality of the unit cell distribution, with a focus on standard deviation (σ) and skewness (κ). In total, 638 different combinations of detector distance and photon energy were explored for indexing, resulting in a unique FOM for each combination. To characterize the quality of indexing, these FOM values were interpolated to create a 2D matrix and normalized, with the maximum value set to 1 [Fig. 3(b)]. The center-of-mass was calculated from this matrix to determine the ideal detector distance and photon energy, yielding unit cell dimensions of a = 79.45 Å, b = 79.47 Å, c = 38.32 Å, and $\alpha = \beta = \gamma = 90°$. The non-interpolated individual indexing results are shown in supplementary information Fig. J. For each subsequent pulse, the photon energy was scanned to fit the unit cell dimensions to that of the first optimized pulse,

$$\text{FOM}_{indexing} = \text{IR} \times \left(1 - \sum_{i=a,b,c,\alpha\beta\gamma} \frac{\sigma_i \kappa_i}{6} \right).$$

The photon energy drift required to correct the unit cells over the train was 15.5 eV [Fig. 3(e)]. The corrected unit cell plot is shown in Fig. 3(a) with blue squares. The photon energy for each of the pulses is depicted in Fig. 3(e). It is noteworthy that the same amount of unit cell drift from the start to the end of the macropulse was observed at half the repetition rate. The drift remains within the Self-Amplified Spontaneous Emission (SASE) bandwidth of 25 eV.[59] A similar magnitude of energy drift is observed in the study by Kujala *et al.* (Figs. 12 and 13).[59] Although no explicit conclusion is provided, the energy drift appears to be correlated with the pulse intensity in the SASE process.[59] Shot-to-shot x-ray photon energy via the HIgh REsolution hard X-ray single-shot spectrometer (HIREX) spectrometer would negate this issue and is recommended for SFX experiments conducted at EuXFEL.[59] CrystFEL has the ability to apply a separate photon energy on a frame-by-frame basis. This correction is important for time-resolved SFX (TR-SFX) experiments where a very accurate estimate of the unit cell for the detection of small light induced differences. Figure 3(d) shows crystallographic merging statistics before and after the photon energy correction. A clear decrease in Rsplit and increase in CC* are observed at higher resolutions, indicative of increase in crystallographic data quality.[60,61]

Crystallographic comparison of sequential hits and droplets

Data from each pulse were indexed with the corrected photon energy as described above. All the data were then concatenated, scaled, and merged with CrystFEL's partialator, to ensure consistent scaling across all data. The stream files were merged either with point group 4/mmm for the standard case or 422 for the anomalous difference map (merging statistics for both are shown in Table I). In total, 74,048 indexed frames were collected over a 27 min period. R$_{free}$ sets of 10% were then generated using CCP4i2's FreeRFlag.[63] PHENIX's function phenix.refine[64] was used for iterative rounds of model building based on a lysozyme model [Protein

Data Bank (PDB)[65] accession code 6H0K[47]]. The appropriate resolution cutoff for the crystallographic data was determined using Diederichs and Karplus "paired refinement."[62,66,67] This approach implements successive rounds of refinement including higher and higher resolution shells. A comparison of the R-factors is used to define when data at a certain resolution no longer contribute to a better model. The PAIREF package outputs a "strict" and "benevolent" cutoff, which, for the data presented here, was 1.5 and 1.38 Å, respectively.[62] The full output from the refinement procedure is shown in supplementary information Figs. M and N. Notably, the Fourier shell correlation (FSC) is still significantly above 50% at the limiting resolution of 1.38 Å (supplementary information Fig. N).[69] The resulting merging and refinement statistics for both cutoffs are given in Table I.

The high-quality crystallographic data result in the electron density maps shown in Fig. 4 and are displayed in PyMOL.[68] Figure 4(a$_i$) displays an anomalous Fourier difference map at 9.3 keV, revealing clear density on the sulfur atoms contoured at 3.5σ. Furthermore, the density is localized on the individual atoms within the disulfide bond between CYS-30 and CYS-115. Clear density on the sulfurs indicates good signal-to-noise ratio, as this experiment was conducted with 9.3 keV, far from sulfur's K-edge absorption at 2.4 keV, and hence, sulfur has a minimal cross section at this wavelength. In Fig. 4(a$_{ii}$), the 2Fo-Fc map of the data, cutoff to 1.38 Å, displays clear density for the aromatic groups (PDB: 8RUS). Figures 4(b) and 4(c) report isomorphous difference maps at $\pm 3\sigma$, comparing the first Knackwurst to subsequent Knackwurst and the first hit to within a Knackwurst to the subsequent hits. No interpretable features were found in any difference map, with the limited visible density attributed to noise.

Data rate and sample consumption

The high hit rate of the DoD method means that, despite not running at the MHz repetition rates of the GDVN injection methods, the actual data rate, which we define as the indexed frames per second, is still competitive. Table II is a comparison of published datasets taken at EuXFEL on lysozyme. The data rate achieved in this work is second to only one megahertz study.[37] We note, however, that the work presented here was a time-limited testing at the start of a beamtime, and it is not unreasonable to imagine further improvement of this rate with more time and optimization.

The DoD method minimizes the sample volume required for an SFX experiment, particularly for the EuXFEL macro-pulse structure because there is no sample waste. Table II presents the sample consumption rates of various lysozyme SFX experiments at EuXFEL, which shows that the DoD's consumption rate is two orders of magnitude lower than GDVN. Once optimized, the DoD method can reliably deliver samples over an extended period without clogging. Furthermore, sample changeovers are less frequent due to the lower consumption rate. In this work, a single load would last approximately 1-2 h, compared to the more common 30 min for GDVN. This duration could be extended further, limited only by crystal settling in the sample delivery tube. Table II additionally lists the amount of sample required for 1×10^6 indexed frames, which is typical of a very successful TR-SFX experiment. This reduces the amount of protein required for an SFX experiment from the grams to milligrams scale, making this structural measurement much more feasible for a vast range of proteins.

TABLE I. Lysozyme crystallographic and refinement statistics for merged runs 48,49, and 50 at EuXFEL beamtime p002808. Highest resolution shells for each strict and benevolent cutoff shown in parentheses at (a) 1.50–1.53 Å and (b) and (c) 1.38–1.40 Å. Resolution cutoff informed from paired refinement procedure PAIREF.[62] Data merged in P4$_3$2$_1$2 (PDB 8RUS) for most of data processing except for the anomalous difference maps shown in Fig. 4(a$_i$) merged P422, a non-centrosymmetric point group.

Data collection			
Wavelength (eV)			9295.5–9279
No. of indexed patterns			74 048

	P4$_3$2$_1$2		P422
Merged statistics	Strict (1.5 Å)	Benevolent (1.38 Å)	Anomalous maps
Resolution limit (A)	18.73–1.5	18.73–1.38	18.73–1.38
Total number of reflections	11 981 740 (322 575)	12 383 860 (107 809)	12 333 371
Number of unique reflections	20 240	25 840	48 251
Completeness	100 (100)[a]	100 (100)[b]	100 (99.98)[c]
Redundancy	591.0 (163.8)[a]	479.3 (42.7)[b]	255.6 (22.3)[c]
Signal to noise	8.34 (1.66)[a]	6.77 (0.90)[b]	5.07 (0.79)[c]
Wilson B-factor	30.43	25.21	25.09
Rsplit(%)	8.29 (29.30)[a]	8.36 (54.88)[b]	10.85 (75.67)[c]
CC*	0.997 (0.399)[a]	0.997 (0.134)[b]	0.996 (0.000)[c]
CC$_{1/2}$	0.990 (0.086)[a]	0.990 (0.010)[b]	0.985 (0.001)[c]

	P4$_3$2$_1$2		P422
Refinement statistics	Strict (1.5 Å)	Benevolent (1.38 Å)	Anomalous maps
Resolution range	18.73–1.5	18.73–1.38	18.73–1.38
Point group	P 43 21 2	P 43 21 2	P 422
Unit cell	79.47, 79.47, 38.32, 90, 90, 90	79.47, 79.47, 38.32, 90, 90, 90	79.47, 79.47, 38.32, 90, 90, 90
Reflection for refinement	18 172	23 142	43 241
Reflections for R-free	2 023	2 639	4 899
No. none-hydrogen atoms	2 165	2 165	2 165
R-work	0.1595	0.1679	0.1489
R-free	0.1927	0.2015	0.1969
RMS (bonds)	0.009	0.009	0.008
RMS (angles)	1.036	1.002	0.932
Ramachandran favored (%)	99.21	99.21	99.21
Ramachandran allowed (%)	0.79	0.79	0.79
Ramachandran outliers (%)	0	0	0
Rotamer outliers (%)	0	0	0
Clash-score	3.8	2.9	1.44
Average B-factors	38.6	35.6	38.0

We obtain for the first-time high-quality protein crystallographic data from the LPD that was not designed for crystallography. The high-resolution data are facilitated by the large detector area of the LPD and the short distance between the detector and the sample, which is adaptable with the movable LPD—unlike other endstations with fixed distances. The quality of the data benefits from the LPD's extensive dynamic range (1:100 000 photons), which tolerates a higher x-ray flux, and the partial helium environment that lowers the signal-to-noise ratio. We expect data quality was degraded with the scatter from the larger liquid volume of the droplets of the DoD. Further reductions in sample consumption and noise could be achieved by using a smaller orifice size. To improve future experiments, it would be beneficial to reduce the amount of air between the sample and detector, for example, using a helium bag.

CONCLUSIONS

This work demonstrates kHz droplet-on-demand injection at the EuXFEL, achieving data rates competitive with current state-of-the-art MHz GDVN sample delivery methods. This approach significantly
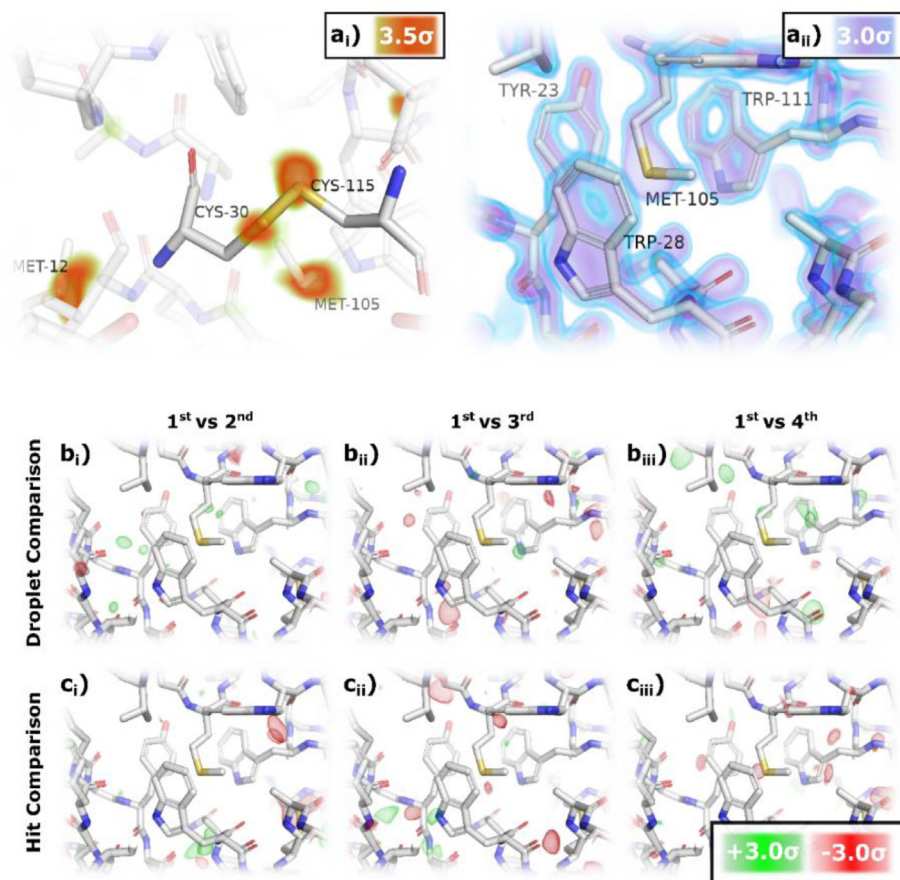
13 May 2024 15:17:15

FIG. 4. Electron density maps for Lysozyme. (a_i) Phased anomalous Fourier difference map shown at 3.5 σ, highlighting sulfur atoms of various residues. (a_{ii}) 2Fo-Fc map at 3σ (purple) to 3.5σ (blue) highlighting resolution of data with clear holes in density around aromatic groups (PDB: 8RUS). (b) and (c) Isomorphous difference maps at 3.0σ comparing all hits merged in the first Knackwurst to subsequent Knackwursts (b) and all Knackwursts merged for first hit to subsequent hits (c). No obvious difference density indicates no issues with hitting droplets multiple times kHz rates.

reduces the volume of consumed sample by two orders of magnitude over GDVN.

A systematic drift in the average photon energy over the train was observed, and correcting this chirp improved the crystallographic data. We develop methods to process SFX data from the LPD detector that result in excellent crystallographic statistics and high-quality structure. We report the highest resolution protein structure collected at EuXFEL (1.38 Å), attributed to using a smaller detector distance with a large dynamic range of the LPD, used for the first protein crystallography experiment at the FXE endstation.

It is anticipated that SFX experiments performed at this facility will benefit from shot-to-shot photon energy diagnostics. Comparison

TABLE II. Comparison of published lysozyme data collected at EuXFEL. Highlighting the data rate, quality, and low sample consumption of a DoD. All parameters calculated from most conservative values listed in various publications.

Delivery method	Intra train repetition rate (MHz)	Effective repetition rate (kHz)	Data rate (Indexed frames per second)	Resolution (Å)	Sample consumption (mg/min)	Sample over beamtime[a]
GDVN[37]	1.2	1.2	99	1.6	1.8	0.303
GDVN[38]	1.125	0.15	5.1	1.7	1.8	5.882
GDVN[47]	1.125	0.5	11	1.9	1.75	2.727
GDVN[39]	1.125	0.14	3.1	2.1	1.64	8.817
GDVN[40]	1.1	0.3	10.4	2	· · ·	· · ·
DoD	0.047	0.16	45.7	1.5 (1.38)	0.0175	0.006

[a]Assuming 1 000 000 indexed frames and no sample wastage (grams).

of multiple droplet hits showed no significant changes over the train, making the DoD a suitable injection method at these rates. With advancements in inkjet technology injectors, it is predicted that the limitations of current DoD setups will decrease, approaching megahertz droplet injection rates.[35] The reduced sample consumption, while maintaining high crystallographic quality data, will make SFX beamtimes more feasible for a wider range of biologically and chemically relevant samples.

AUTHOR DECLARATIONS
Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Samuel Perrett: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – original draft (lead). **Alisia Fadini:** Data curation (equal); Formal analysis (equal); Investigation (equal). **Christopher D. M. Hutchison:** Data curation (equal); Investigation (equal); Methodology (equal). **Sayantan Bhattacharya:** Data curation (equal); Investigation (equal). **Cade Morrison:** Investigation (equal). **Oleksii Turkot:** Formal analysis (equal); Methodology (equal); Software (equal). **Mads Bregenholt Jakobsen:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal). **Michael Größler:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal). **José Licón-Saláiz:** Data curation (equal). **Florian Griese:** Formal analysis (equal); Methodology (equal); Software (equal). **Samuel Flewett:** Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal). **Joana Valério:** Data curation (equal); Methodology (equal); Resources (equal). **Joachim Schulz:** Methodology (equal); Resources (equal). **Mykola Biednov:** Data curation (equal); Formal analysis (equal); Methodology (equal); Resources (equal). **Yifeng Jiang:** Data curation (equal); Formal analysis (equal); Methodology (equal); Resources (equal). **Han Huijong:** Data curation (equal); Investigation (equal); Methodology (equal). **Hazem Yousef:** Data curation (equal); Formal analysis (equal); Methodology (equal); Software (equal); Validation (equal). **Dmitry Khakhulin:** Investigation (equal); Methodology (equal); Resources (equal). **Christopher Milne:** Data curation (equal); Methodology (equal); Resources (equal). **Antony Barty:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Software (equal). **Jasper J. van Thor:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]C. D. M. Hutchison, J. M. Baxter, A. Fitzpatrick, G. Dorlhiac, A. Fadini, S. Perrett, K. Maghlaoui, S. B. Lefèvre, V. Cordon-Preciado, J. L. Ferreira, V. U. Chukhutsina, D. Garratt, J. Barnard, G. Galinis, F. Glencross, R. M. Morgan, S. Stockton, B. Taylor, L. Yuan, M. G. Romei, C. Y. Lin, J. P. Marangos, M. Schmidt, V. Chatrchyan, T. Buckup, D. Morozov, J. Park, S. Park, I. Eom, M. Kim, D. Jang, H. Choi, H. J. Hyun, G. Park, E. Nango, R. Tanaka, S. Owada, K. Tono, D. P. DePonte, S. Carbajo, M. Seaberg, A. Aquila, S. Boutet, A. Barty, S. Iwata, S. G. Boxer, G. Groenhof, and J. J. van Thor, "Optical control of ultrafast structural dynamics in a fluorescent protein," Nat. Chem. **15**, 1607–1615 (2023).

[2]T. Gruhl, T. Weinert, M. J. Rodrigues, C. J. Milne, G. Ortolani, K. Nass, E. Nango, S. Sen, P. J. M. Johnson, C. Cirelli, A. Furrer, S. Mous, P. Skopintsev, D. James, F. Dworkowski, P. Båth, D. Kekilli, D. Ozerov, R. Tanaka, H. Glover, C. Bacellar, S. Brünle, C. M. Casadei, A. D. Diethelm, D. Gashi, G. Gotthard, R. Guixà-González, Y. Joti, V. Kabanova, G. Knopp, E. Lesca, P. Ma, I. Martiel, J. Mühle, S. Owada, F. Pamula, D. Sarabi, O. Tejero, C. J. Tsai, N. Varma, A. Wach, S. Boutet, K. Tono, P. Nogly, X. Deupi, S. Iwata, R. Neutze, J. Standfuss, G. Schertler, and V. Panneels, "Ultrafast structural changes direct the first molecular events of vision," Nature **615**, 939–944 (2023).

[3]E. A. Schriber, D. W. Paley, R. Bolotovsky, D. J. Rosenberg, R. G. Sierra, A. Aquila, D. Mendez, F. Poitevin, J. P. Blaschke, A. Bhowmick, R. P. Kelly, M. Hunter, B. Hayes, D. C. Popple, M. Yeung, C. Pareja-Rivera, S. Lisova, K. Tono, M. Sugahara, S. Owada, T. Kuykendall, K. Yao, P. J. Schuck, D. Solis-Ibarra, N. K. Sauter, A. S. Brewster, and J. N. Hohman, "Chemical crystallography by serial femtosecond x-ray diffraction," Nature **601**, 360–365 (2022).

[4]H. I. J. Kang, Y. Lee, S. Lee, H. Ki, J. Kim, J. Gu, Y. Cha, J. Heo, K. W. Lee, S. O. Kim, J. P. J. Kang, Y. Lee, S. Lee, H. Ki, J. Kim, J. Gu, Y. Cha, J. Heo, and K. W. L. Seong, "Dynamic three-dimensional structures of a metal–organic framework captured with femtosecond serial crystallography," chemRxiv, 2023.

[5]N. Christou, V. Apostolopoulou, D. V. M. Melo, M. Ruppert, A. Pateras, A. R. Mashhour, O. M. Yefanov, and M. Galchenkova, "Time-resolved crystallography captures light-driven DNA repair," Science **1020**, 1015–1020 (2023).

[6]B. Maity, M. Shoji, F. Luo, T. Nakane, S. Abe, S. Owada, J. Kang, K. Tono, R. Tanaka, T. T. Pham, M. Kojima, J. Tanaka, J. Tian, H. Noya, Y. Nakasuji, and A. Asanuma, "Real-time observation of a metal complex-driven reaction intermediate using a porous protein crystal and femtosecond crystallography serial," chemRxiv, 2023.

[7]A. M. Orville, "Recent results in time resolved serial femtosecond crystallography at XFELs," Curr. Opin. Struct. Biol. **65**, 193–208 (2020).

[8]P. H. Bucksbaum and N. Berrah, "Brighter and faster: The promise and challenge of the x-ray free-electron laser," Phys. Today **68**(7), 26–32 (2015).

[9]K. Pande, C. D. M. Hutchison, G. Groenhof, A. Aquila, J. S. Robinson, J. Tenboer, S. Basu, S. Boutet, D. P. DePonte, M. Liang, T. A. White, N. A. Zatsepin, O. Yefanov, D. Morozov, D. Oberthuer, C. Gati, G. Subramanian, D. James, Y. Zhao, J. Koralek, J. Brayshaw, C. Kupitz, C. Conrad, S. Roy-Chowdhury, J. D. Coe, M. Metz, P. L. Xavier, T. D. Grant, J. E. Koglin, G. Ketawala, R. Fromme, V. Srajer, R. Henning, J. C. Spence, A. Ourmazd, P. Schwander, U. Weierstall, M. Frank, P. Fromme, A. Barty, H. N. Chapman, K. Moffat, J. J. Van Thor, and M. Schmidt, "Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein," Science **352**(6286), 725–729 (2016).

[10]T. R. M. Barends, B. Stauch, V. Cherezov, and I. Schlichting, "Serial femtosecond crystallography," Nat. Rev. Methods Primers **2**, 60 (2022).

[11]L. C. Johansson, B. Stauch, A. Ishchenko, and V. Cherezov, "A bright future for serial femtosecond crystallography with XFELs," Trends Biochem. Sci. **42**, 749–762 (2017).

[12]J. C. P. Koliyadu, R. Letrun, H. J. Kirkwood, J. Liu, M. Jiang, M. Emons, R. Bean, V. Bellucci, J. Bielecki, S. Birnsteinova, R. De Wijn, T. Dietze, E. Juncheng, J. Grünert, D. Kane, C. Kim, Y. Kim, M. Lederer, B. Manning, G. Mills, L. L. Morillo, N. Reimers, D. Rompotis, A. Round, M. Sikorski, C. M. S.

Takem, P. Vagovič, S. Venkatesan, J. Wang, U. Wegner, A. P. Mancuso, T. Sato, and M. Yabashi, "Pump-probe capabilities at the SPB/SFX instrument of the European XFEL," J. Synchrotron Radiat. **29**, 1273–1283 (2022).

[13]J. J. Van Thor, "Advances and opportunities in ultrafast X-ray crystallography and ultrafast structural optical crystallography of nuclear and electronic protein dynamics," Struct. Dyn. **6**(5), 050901 (2019).

[14]S. Perrett, V. Chatrchyan, T. Buckup, and J. J. van Thor, "Application of density matrix Wigner transforms for ultrafast macromolecular and chemical x-ray crystallography," J. Chem. Phys. **160**(10), 100901 (2024).

[15]D. A. Sherrell, A. J. Foster, L. Hudson, B. Nutter, J. O'Hea, S. Nelson, O. Paré-Labrosse, S. Oghbaey, R. J. D. Miller, and R. L. Owen, "A modular and compact portable mini-endstation for high-precision, high-speed fixed target serial crystallography at FEL and synchrotron sources," J. Synchrotron Radiat. **22**, 1372–1378 (2015).

[16]P. Rabe, J. H. Beale, A. Butryn, P. Aller, P. A. Lang, D. N. Axford, S. B. Carr, T. M. Leissing, M. A. McDonough, B. Davy, A. Ebrahim, J. Orlans, S. L. S. Storm, A. M. Orville, C. J. Schofielda, and R. L. Owen, "Anaerobic fixed-target serial crystallography," IUCrJ **7**, 901–912 (2020).

[17]S. Y. Park, H. Choi, C. Eo, Y. Cho, and K. H. Nam, "Fixed-target serial synchrotron crystallography using nylon mesh and enclosed film-based sample holder," Crystals **10**, 803 (2020).

[18]M. Carrillo, T. J. Mason, A. Karpik, I. Martiel, M. W. Kepa, K. E. McAuley, J. H. Beale, and C. Padeste, "Micro-structured polymer fixed targets for serial crystallography at synchrotrons and XFELs," IUCrJ **10**(6), 678–693 (2023).

[19]R. L. Owen, D. De Sanctis, A. R. Pearson, and J. H. Beale, "A standard descriptor for fixed-target serial crystallography," Acta Crystallogr., Sect. D **79**, 668–672 (2023).

[20]U. Weierstall, D. James, C. Wang, T. A. White, D. Wang, W. Liu, J. C. H. Spence, R. Bruce Doak, G. Nelson, P. Fromme, R. Fromme, I. Grotjohann, C. Kupitz, N. A. Zatsepin, H. Liu, S. Basu, D. Wacker, G. Won Han, V. Katritch, S. Boutet, M. Messerschmidt, G. J. Williams, J. E. Koglin, M. M. Seibert, M. Klinker, C. Gati, R. L. Shoeman, A. Barty, H. N. Chapman, R. A. Kirian, K. R. Beyerlein, R. C. Stevens, D. Li, S. T. A. Shah, N. Howe, M. Caffrey, and V. Cherezov, "Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography," Nat. Commun. **5**, 3309 (2014).

[21]M. Vakili, H. Han, C. Schmidt, A. Wrona, M. Kloos, I. De Diego, K. Dörner, T. Geng, C. Kim, F. H. M. Koua, D. V. M. Melo, M. Rappas, A. Round, E. Round, M. Sikorski, J. Valerio, T. Zhou, K. Lorenzen, J. Schulz, and J. Hajdu, "Mix-and-extrude: High-viscosity sample injection towards time-resolved protein crystallography," J. Appl. Crystallogr. **56**, 1038–1045 (2023).

[22]M. O. Wiedorn, S. Awel, A. J. Morgan, K. Ayyer, Y. Gevorkov, H. Fleckenstein, N. Roth, L. Adriano, R. Bean, K. R. Beyerlein, J. Chen, J. Coe, F. Cruz-Mazo, T. Ekeberg, R. Graceffa, M. Heymann, D. A. Horke, J. Knoška, V. Mariani, R. Nazari, D. Oberthür, A. K. Samanta, R. G. Sierra, C. A. Stan, O. Yefanov, D. Rompotis, J. Correa, B. Erk, R. Treusch, J. Schulz, B. G. Hogue, A. M. Gañán-Calvo, P. Fromme, J. Küpper, A. V. Rode, S. Bajt, R. A. Kirian, and H. N. Chapman, "Rapid sample delivery for megahertz serial crystallography at x-ray FELs," IUCrJ **5**, 574–584 (2018).

[23]J. Knoška, L. Adriano, S. Awel, K. R. Beyerlein, O. Yefanov, D. Oberthuer, G. E. P. Murillo, N. Roth, I. Sarrou, P. Villanueva-Perez, M. O. Wiedorn, F. Wilde, S. Bajt, H. N. Chapman, and M. Heymann, "Ultracompact 3D microfluidics for time-resolved structural biology," Nat. Commun. **11**, 657 (2020).

[24]D. Oberthuer, J. Knoška, M. O. Wiedorn, K. R. Beyerlein, D. A. Bushnell, E. G. Kovaleva, M. Heymann, L. Gumprecht, R. A. Kirian, A. Barty, V. Mariani, A. Tolstikova, L. Adriano, S. Awel, M. Barthelmess, K. Dörner, P. L. Xavier, O. Yefanov, D. R. James, G. Nelson, D. Wang, G. Calvey, Y. Chen, A. Schmidt, M. Szczepek, S. Frielingsdorf, O. Lenz, E. Snell, P. J. Robinson, B. Šarler, G. Belšak, M. Maček, F. Wilde, A. Aquila, S. Boutet, M. Liang, M. S. Hunter, P. Scheerer, J. D. Lipscomb, U. Weierstall, R. D. Kornberg, J. C. H. Spence, L. Pollack, H. N. Chapman, and S. Bajt, "Double-flow focused liquid injector for efficient serial femtosecond crystallography," Sci. Rep. **7**, 44628 (2017).

[25]R. K. Y. Cheng, "Towards an optimal sample delivery method for serial crystallography at XFEL," Crystals **10**(3), 215 (2020).

[26]R. de Wijn, D. V. M. Melo, F. H. M. Koua, and A. P. Mancuso, "Potential of time-resolved serial femtosecond crystallography using high repetition rate XFEL sources," Appl. Sci. **12**, 2551 (2022).

[27]M. L. Grünbein and G. N. Kovacs, "Sample delivery for serial crystallography at free-electron lasers and synchrotrons," Acta Crystallogr., Sect. D **75**, 178–191 (2019).

[28]R. G. Sierra, U. Weierstall, D. Oberthuer, M. Sugahara, E. Nango, S. Iwata, and A. Meents, *X-Ray Free Electron Lasers: A Revolution in Structural Biology* (Springer, Cham, 2018).

[29]A. Echelmeier, J. C. Villarreal, M. Messerschmidt, D. Kim, J. D. Coe, D. Thifault, S. Botha, A. Egatz-Gomez, S. Gandhi, G. Brehm, C. E. Conrad, D. T. Hansen, C. Madsen, S. Bajt, J. D. Meza-Aguilar, D. Oberthür, M. O. Wiedorn, H. Fleckenstein, D. Mendez, J. Knoška, J. M. Martin-Garcia, H. Hu, S. Lisova, A. Allahgholi, Y. Gevorkov, K. Ayyer, S. Aplin, H. M. Ginn, H. Graafsma, A. J. Morgan, D. Greiffenberg, A. Klujev, T. Laurus, J. Poehlsen, U. Trunk, D. Mezza, B. Schmidt, M. Kuhn, R. Fromme, J. Sztuk-Dambietz, N. Raab, S. Hauf, A. Silenzi, T. Michelat, C. Xu, C. Danilevski, A. Parenti, L. Mekinda, B. Weinhausen, G. Mills, P. Vagovic, Y. Kim, H. Kirkwood, R. Bean, J. Bielecki, S. Stern, K. Giewekemeyer, A. R. Round, J. Schulz, K. Dörner, T. D. Grant, V. Mariani, A. Barty, A. P. Mancuso, U. Weierstall, J. C. H. Spence, H. N. Chapman, N. Zatsepin, P. Fromme, R. A. Kirian, and A. Ros, "Segmented flow generator for serial crystallography at the European x-ray free electron laser," Nat. Commun. **11**, 4511 (2020).

[30]F. D. Fuller, S. Gul, R. Chatterjee, E. S. Burgie, I. D. Young, H. Lebrette, V. Srinivas, A. S. Brewster, T. Michels-Clark, J. A. Clinger, B. Andi, M. Ibrahim, E. Pastor, C. De Lichtenberg, R. Hussein, C. J. Pollock, M. Zhang, C. A. Stan, T. Kroll, T. Fransson, C. Weninger, M. Kubin, P. Aller, L. Lassalle, P. Bräuer, M. D. Miller, M. Amin, S. Koroidov, C. G. Roessler, M. Allaire, R. G. Sierra, P. T. Docker, J. M. Glownia, S. Nelson, J. E. Koglin, D. Zhu, M. Chollet, S. Song, H. Lemke, M. Liang, D. Sokaras, R. Alonso-Mori, A. Zouni, J. Messinger, U. Bergmann, A. K. Boal, J. M. Bollinger, C. Krebs, M. Högbom, G. N. Phillips, R. D. Vierstra, N. K. Sauter, A. M. Orville, J. Kern, V. K. Yachandra, and J. Yano, "Drop-on-demand sample delivery for studying biocatalysts in action at x-ray free-electron lasers," Nat. Methods **14**, 443–449 (2017).

[31]N. A. Miller, L. B. Michocki, R. Alonso-Mori, A. Britz, A. Deb, D. P. Deponte, J. M. Glownia, A. K. Kaneshiro, C. Kieninger, J. Koralek, J. H. Meadows, T. B. Van Driel, B. Kräutler, K. J. Kubarych, J. E. Penner-Hahn, and R. J. Sension, "Antivitamins B12 in a microdrop: The excited-state structure of a precious sample using transient polarized x-ray absorption near-edge structure," J. Phys. Chem. Lett. **10**, 5484–5489 (2019).

[32]C. G. Roessler, R. Agarwal, M. Allaire, R. Alonso-Mori, B. Andi, J. F. R. Bachega, M. Bommer, A. S. Brewster, M. C. Browne, R. Chatterjee, E. Cho, A. E. Cohen, M. Cowan, S. Datwani, V. L. Davidson, J. Defever, B. Eaton, R. Ellson, Y. Feng, L. P. Ghislain, J. M. Glownia, G. Han, J. Hattne, J. Hellmich, A. Héroux, M. Ibrahim, J. Kern, A. Kuczewski, H. T. Lemke, P. Liu, L. Majlof, W. M. McClintock, S. Myers, S. Nelsen, J. Olechno, A. M. Orville, N. K. Sauter, A. S. Soares, S. M. Soltis, H. Song, R. G. Stearns, R. Tran, Y. Tsai, M. Uervirojnangkoorn, C. M. Wilmot, V. Yachandra, J. Yano, E. T. Yukl, D. Zhu, and A. Zouni, "Acoustic injectors for drop-on-demand serial femtosecond crystallography," Structure **24**, 631–640 (2016).

[33]F. Mafuné, K. Miyajima, K. Tono, Y. Takeda, J. Y. Kohno, N. Miyauchi, J. Kobayashi, Y. Joti, E. Nango, S. Iwata, and M. Yabashi, "Microcrystal delivery by pulsed liquid droplet for serial femtosecond crystallography," Acta Crystallogr., Sect. D **72**, 520–523 (2016).

[34]A. Fadini, C. D. M. Hutchison, D. Morozov, J. Chang, K. Maghlaoui, S. Perrett, F. Luo, J. C. X. Kho, M. G. Romei, R. M. L. Morgan, C. M. Orr, V. Cordon-Preciado, T. Fujiwara, N. Nuemket, T. Tosha, R. Tanaka, S. Owada, K. Tono, S. Iwata, S. G. Boxer, G. Groenhof, E. Nango, and J. J. van Thor, "Serial femtosecond crystallography reveals that photoactivation in a fluorescent protein proceeds via the hula twist mechanism," J. Am. Chem. Soc. **145**, 15796–15808 (2023).

[35]J. C. Miers and W. Zhou, "Droplet formation at megahertz frequency," AIChE J. **63**, 2367–2377 (2017).

[36]J. C. Miers and W. Zhou, in Proceedings of 26th Annual International Solid Freeform Fabrication Symposium - An Additive Manufacturing Conference SFF 2015 (2020).

[37]O. Yefanov, D. Oberthür, R. Bean, M. O. Wiedorn, J. Knoska, G. Pena, S. Awel, L. Gumprecht, M. Domaracky, I. Sarrou, P. L. Xavier, M. Metz, S. Bajt, V. Mariani, Y. Gevorkov, T. A. White, A. Tolstikova, P. Villanueva-Perez, C. Seuring, S. Aplin, A. D. Estillore, J. Küpper, A. Klyuev, M. Kuhn, T. Laurus, H.

Graafsma, D. C. F. Monteiro, M. Trebbin, F. R. N. C. Maia, F. Cruz-Mazo, A. M. Gañan-Calvo, M. Heymann, C. Darmanin, B. Abbey, M. Schmidt, P. Fromme, K. Giewekemeyer, M. Sikorski, R. Graceffa, P. Vagovic, T. Kluyver, M. Bergemann, H. Fangohr, J. Sztuk-Dambietz, S. Hauf, N. Raab, V. Bondar, A. P. Mancuso, H. Chapman, and A. Barty, "Evaluation of serial crystallographic structure determination within megahertz pulse trains," Struct. Dyn. **6**(6), 064702 (2019).

38 M. O. Wiedorn, D. Oberthür, R. Bean, R. Schubert, N. Werner, B. Abbey, M. Aepfelbacher, L. Adriano, A. Allahgholi, N. Al-Qudami, J. Andreasson, S. Aplin, S. Awel, K. Ayyer, S. Bajt, I. Barák, S. Bari, J. Bielecki, S. Botha, D. Boukhelef, W. Brehm, S. Brockhauser, I. Cheviakov, M. A. Coleman, F. Cruz-Mazo, C. Danilevski, C. Darmanin, R. B. Doak, M. Domaracky, K. Dörner, Y. Du, H. Fangohr, H. Fleckenstein, M. Frank, P. Fromme, A. M. Gañán-Calvo, Y. Gevorkov, K. Giewekemeyer, H. M. Ginn, H. Graafsma, R. Graceffa, D. Greiffenberg, L. Gumprecht, P. Göttlicher, J. Hajdu, S. Hauf, M. Heymann, S. Holmes, D. A. Horke, M. S. Hunter, S. Imlau, A. Kaukher, Y. Kim, A. Klyuev, J. Knoška, B. Kobe, M. Kuhn, C. Kupitz, J. Küpper, J. M. Lahey-Rudolph, T. Laurus, K. L. Cong, R. Letrun, P. L. Xavier, L. Maia, F. R. N. C. Maia, V. Mariani, M. Messerschmidt, M. Metz, D. Mezza, T. Michelat, G. Mills, D. C. F. Monteiro, A. Morgan, K. Mühlig, A. Munke, A. Münnich, J. Nette, K. A. Nugent, T. Nuguid, A. M. Orville, S. Pandey, G. Pena, P. Villanueva-Perez, J. Poehlsen, G. Previtali, L. Redecke, W. M. Riekehr, H. Rohde, A. Round, T. Safenreiter, I. Sarrou, T. Sato, M. Schmidt, B. Schmitt, R. Schönherr, J. Schulz, J. A. Sellberg, M. M. Seibert, C. Seuring, M. L. Shelby, R. L. Shoeman, M. Sikorski, A. Silenzi, C. A. Stan, X. Shi, S. Stern, J. Sztuk-Dambietz, J. Szuba, A. Tolstikova, M. Trebbin, U. Trunk, P. Vagovic, T. Ve, B. Weinhausen, T. A. White, K. Wrona, C. Xu, O. Yefanov, N. Zatsepin, J. Zhang, M. Perbandt, A. P. Mancuso, C. Betzel, H. Chapman, and A. Barty, "Megahertz serial crystallography," Nat. Commun. **9**, 4025 (2018).

39 S. Holmes, H. J. Kirkwood, R. Bean, K. Giewekemeyer, A. V. Martin, M. Hadian-Jazi, M. O. Wiedorn, D. Oberthür, H. Marman, L. Adriano, N. Al-Qudami, S. Bajt, I. Barák, S. Bari, J. Bielecki, S. Brockhauser, M. A. Coleman, F. Cruz-Mazo, C. Danilevski, K. Dörner, A. M. Gañán-Calvo, R. Graceffa, H. Fanghor, M. Heymann, M. Frank, A. Kaukher, Y. Kim, B. Kobe, J. Knoška, T. Laurus, R. Letrun, L. Maia, M. Messerschmidt, M. Metz, T. Michelat, G. Mills, S. Molodtsov, D. C. F. Monteiro, A. J. Morgan, A. Münnich, G. E. P. Murillo, G. Previtali, A. Round, T. Sato, R. Schubert, J. Schulz, M. Shelby, C. Seuring, J. A. Sellberg, M. Sikorski, A. Silenzi, S. Stern, J. Sztuk-Dambietz, J. Szuba, M. Trebbin, P. Vagovic, T. Ve, B. Weinhausen, K. Wrona, P. L. Xavier, C. Xu, O. Yefanov, K. A. Nugent, H. N. Chapman, A. P. Mancuso, A. Barty, B. Abbey, and C. Darmanin, "Megahertz pulse trains enable multi-hit serial femtosecond crystallography experiments at x-ray free electron lasers," Nat. Commun. **13**, 4708 (2022).

40 H. J. Kirkwood, R. de Wijn, G. Mills, R. Letrun, M. Kloos, M. Vakili, M. Karnevskiy, K. Ahmed, R. J. Bean, J. Bielecki, F. Dall'Antonia, Y. Kim, C. Kim, J. Koliyadu, A. Round, T. Sato, M. Sikorski, P. Vagovič, J. Sztuk-Dambietz, and A. P. Mancuso, "A multi-million image serial femtosecond crystallography dataset collected at the European XFEL," Sci. Data **9**, 161 (2022).

41 H. Wijshoff, "The dynamics of the piezo inkjet printhead operation," Phys. Rep. **491**, 77–177 (2010).

42 J. Eggers and E. Villermaux, "Physics of liquid jets," Rep. Prog. Phys. **71**, 036601 (2008).

43 T. Tschentscher, C. Bressler, J. Grünert, A. Madsen, A. Mancuso, M. Meyer, A. Scherz, H. Sinn, and U. Zastrau, "Photon beam transport and scientific instruments at the European XFEL," Appl. Sci. **7**, 592 (2017).

44 J. Fangohr, H. Göries, D. Kamil, E. Kirienko, Y. Kluyver, T. Michelat, T. Rosca, and R. Wrigley, "European XFEL open source data analysis tools: Extra-data, extra-GEOM, extra-foam," 2023, see https://github.com/European-XFEL/EXtra-foam.

45 M. L. Grünbein, A. Gorel, L. Foucar, S. Carbajo, W. Colocho, S. Gilevich, E. Hartmann, M. Hilpert, M. Hunter, M. Kloos, J. E. Koglin, T. J. Lane, J. Lewandowski, A. Lutman, K. Nass, G. N. Kovacs, C. M. Roome, J. Sheppard, R. L. Shoeman, M. Stricker, T. van Driel, S. Vetter, R. B. Doak, S. Boutet, A. Aquila, F. J. Decker, T. R. M. Barends, C. A. Stan, and I. Schlichting, "Effect of X-ray free-electron laser-induced shockwaves on haemoglobin microcrystals delivered in a liquid jet," Nat. Commun. **12**(1), 1672 (2021).

46 S. Pandey, R. Bean, T. Sato, I. Poudyal, J. Bielecki, J. Cruz Villarreal, O. Yefanov, V. Mariani, T. A. White, C. Kupitz, M. Hunter, M. H. Abdellatif, S. Bajt, V. Bondar, A. Echelmeier, D. Doppler, M. Emons, M. Frank, R. Fromme, Y. Gevorkov, G. Giovanetti, M. Jiang, D. Kim, Y. Kim, H. Kirkwood, A. Klimovskaia, J. Knoska, F. H. M. Koua, R. Letrun, S. Lisova, L. Maia, V. Mazalova, D. Meza, T. Michelat, A. Ourmazd, G. Palmer, M. Ramilli, R. Schubert, P. Schwander, A. Silenzi, J. Sztuk-Dambietz, A. Tolstikova, H. N. Chapman, A. Ros, A. Barty, P. Fromme, A. P. Mancuso, and M. Schmidt, "Time-resolved serial femtosecond crystallography at the European XFEL," Nat. Methods **17**(1), 73–78 (2020).

47 M. L. Grünbein, J. Bielecki, A. Gorel, M. Stricker, R. Bean, M. Cammarata, K. Dörner, L. Fröhlich, E. Hartmann, S. Hauf, M. Hilpert, Y. Kim, M. Kloos, R. Letrun, M. Messerschmidt, G. Mills, G. N. Kovacs, M. Ramilli, C. M. Roome, T. Sato, M. Scholz, M. Sliwa, J. Sztuk-Dambietz, M. Weik, B. Weinhausen, N. Al-Qudami, D. Boukhelef, S. Brockhauser, W. Ehsan, M. Emons, S. Esenov, H. Fangohr, A. Kaukher, T. Kluyver, M. Lederer, L. Maia, M. Manetti, T. Michelat, A. Münnich, F. Pallas, G. Palmer, G. Previtali, N. Raab, A. Silenzi, J. Szuba, S. Venkatesan, K. Wrona, J. Zhu, R. B. Doak, R. L. Shoeman, L. Foucar, J. P. Colletier, A. P. Mancuso, T. R. M. Barends, C. A. Stan, and I. Schlichting, "Megahertz data collection from protein microcrystals at an x-ray free-electron laser," Nat. Commun. **9**, 3487 (2018).

48 D. Khakhulin, F. Otte, M. Biednov, C. Bömer, T. K. Choi, M. Diez, A. Galler, Y. Jiang, K. Kubicek, F. A. Lima, A. Rodriguez-Fernandez, P. Zalden, W. Gawelda, and C. Bressler, "Ultrafast x-ray photochemistry at European XFEL: Capabilities of the femtosecond x-ray experiments (FXE) instrument," Appl. Sci. **10**(3), 995 (2020).

49 A. Galler, W. Gawelda, M. Biednov, C. Bomer, A. Britz, S. Brockhauser, T. K. Choi, M. Diez, P. Frankenberger, M. French, D. Görries, M. Hart, S. Hauf, D. Khakhulin, M. Knoll, T. Korsch, K. Kubicek, M. Kuster, P. Lang, F. A. Lima, F. Otte, S. Schulz, P. Zalden, and C. Bressler, "Scientific instrument femtosecond x-ray experiments (FXE): Instrumentation and baseline experimental capabilities," J. Synchrotron Radiat. **26**, 1432–1447 (2019).

50 R. M. Wheater, M. D. Hart, M. C. Veale, M. D. Wilson, D. Doblas-Jiménez, M. Turcato, C. Milne, H. Yousef, and D. Khakhulin, "Development of data correction for the 1M large pixel detector at the EuXFEL," J. Instrum. **17**, P04013 (2022).

51 O. Turkot, F. Dall'Antonia, R. J. Bean, E. Juncheng, H. Fangohr, D. E. Ferreira de Lima, S. Kantamneni, H. J. Kirkwood, F. H. M. Koua, A. P. Mancuso, D. V. M. Melo, A. Round, M. Schuh, E. Sobolev, R. de Wijn, J. J. Wrigley, and L. Gelisio, "EXtra-Xwiz: A tool to streamline serial femtosecond crystallography workflows at European XFEL," Crystals **13**(11), 1533 (2023).

52 T. A. White, V. Mariani, W. Brehm, O. Yefanov, A. Barty, K. R. Beyerlein, F. Chervinskii, L. Galli, C. Gati, T. Nakane, A. Tolstikova, K. Yamashita, C. H. Yoon, K. Diederichs, and H. N. Chapman, "Recent developments in CrystFEL," J. Appl. Crystallogr. **49**, 680–689 (2016).

53 T. A. White, "Processing serial crystallography data with crystFEL: A step-by-step guide," Acta Crystallogr., Sect. D **75**, 219–233 (2019).

54 T. A. White, R. A. Kirian, A. V. Martin, A. Aquila, K. Nass, A. Barty, and H. N. Chapman, "CrystFEL: A software suite for snapshot serial crystallography," J. Appl. Crystallogr. **45**, 335–341 (2012).

55 A. Barty, R. A. Kirian, F. R. N. C. Maia, M. Hantke, C. H. Yoon, T. A. White, and H. Chapman, "Cheetah: Software for high-throughput reduction and analysis of serial femtosecond x-ray diffraction data," J. Appl. Crystallogr. **47**(3), 1118–1131 (2014).

56 Y. Gevorkov, O. Yefanov, A. Barty, T. A. White, V. Mariani, W. Brehm, A. Tolstikova, R. R. Grigat, and H. N. Chapman, "XGANDALF - Extended gradient descent algorithm for lattice finding," Acta Crystallogr., Sect. A **75**, 694–704 (2019).

57 M. Aleksich, D. W. Paley, E. A. Schriber, W. Linthicum, V. Oklejas, D. W. Mittan-Moreau, R. P. Kelly, P. A. Kotei, A. Ghodsi, R. G. Sierra, A. Aquila, F. Poitevin, J. P. Blaschke, M. Vakili, C. J. Milne, F. Dall'Antonia, D. Khakhulin, F. Ardana-Lamas, F. Lima, J. Valerio, H. Han, T. Gallo, O. Turkot, I. J. B. Macias, T. Kluyver, P. Schmidt, L. Gelisio, A. R. Round, Y. Jiang, D. Vinci, Y. Uemura, M. Kloos, M. Hunter, A. P. Mancuso, B. D. Huey, L. R. Parent, N. K. Sauter, A. S. Brewster, and J. N. Hohman, "XFEL microcrystallography of self-assembling silver n-alkanethiolates," J. Am. Chem. Soc. **145**, 17042–17055 (2023).

[58]F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res. **12**, 2825–2830 (2011).

[59]N. Kujala, W. Freund, J. Liu, A. Koch, T. Falk, M. Planas, F. Dietrich, J. Laksman, T. Maltezopoulos, J. Risch, F. Dall'antonia, and J. Grünert, "Hard x-ray single-shot spectrometer at the European x-ray free-electron laser," Rev. Sci. Instrum. **91**, 103101 (2020).

[60]P. A. Karplus and K. Diederichs, "Linking crystallographic model and data quality," Science **336**, 1030–1033 (2012).

[61]K. Diederichs, "Pitfalls of crystallographic reasoning, and linking crystallographic model and data quality," Acta Crystallogr., Sect. A **71**, S22 (2015).

[62]M. Maly, K. Diederichs, J. Dohnalek, and P. Kolenko, "Paired refinement under the control of PAIREF," IUCrJ **7**, 681–692 (2020).

[63]L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. Von Delft, D. Waterman, K. Wilson, M. Winn, and M. Wojdyr, "CCP4i2: The new graphical user interface to the CCP 4 program suite," Acta Crystallogr., Sect. D **74**, 68–84 (2018).

[64]P. D. Adams, P. V. Afonine, G. Bunkóczi, 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, "PHENIX: A comprehensive Python-based system for macromolecular structure solution," Acta Crystallogr., Sect. D **66**, 213–221 (2010).

[65]H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The protein data bank," Nucleic Acids Res. **28**, 235–242 (2000).

[66]K. Diederichs and P. A. Karplus, "Better models by discarding data?," Acta Crystallogr., Sect. D **69**, 1215–1222 (2013).

[67]M. Bublitz, K. Nass, N. D. Drachmann, A. J. Markvardsen, M. J. Gutmann, T. R. M. Barends, D. Mattle, R. L. Shoeman, R. B. Doak, S. Boutet, M. Messerschmidt, M. M. Seibert, G. J. Williams, L. Foucar, L. Reinhard, O. Sitsel, J. L. Gregersen, J. D. Clausen, T. Boesen, K. Gotfryd, K. T. Wang, C. Olesen, J. V. Møller, P. Nissen, and I. Schlichting, "Structural studies of P-type ATPase-ligand complexes using an x-ray free-electron laser," IUCrJ **2**, 409–420 (2015).

[68]W. L. Delano, "Pymol: An open-source molecular graphics tool," CCP4 Newsl. Protein Crystallogr. **40**(1), 82–92 (2002).

[69]See the supplementary material for details on crystallization, sample injection instrumentation and characterization, detector masking, integrated peak intensity analysis, cell distribution statistics, Fourier Shell Correlation analysis, VERDICT limiting resolution analysis, and beam focus characterization.

13 May 2024 15:17:15