LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-844538

Megahertz pulse trains enable multi-hit serial femtosecond crystallography experiments at X-ray free electron lasers

S. Holmes, H. J. Kirkwood, R. Bean, K. Giewekemeyer, A. V. Martin, M. Hadian-Jazi, M. O. Wiedorn, D. Oberthur, L. Adriano, M. A. Coleman, M. Frank, M. Shelby, H. Chapman, A. Mancuso, A. Barty, C. Darmanin, B. Abbey

January 30, 2023

Nature Communications

Disclaimer

Title : Megahertz pulse trains enable multi-hit serial femtosecond crystallography experiments at X-ray free electron lasers

Authors

S. Holmes[1], H. J. Kirkwood[2], R. Bean[2], K. Giewekemeyer[2], A.V. Martin[3], M. Hadian-Jazi[1,2], M. O. Wiedorn[4], D. Oberthur[4], L. Adriano[5], N. Al-Qudami[2], S. Bajt[5], I. Barák[6], S. Bari[5], J. Bielecki[2], S. Brockhauser[2], M. A. Coleman[7], F. Cruz-Mazo[8], C. Danilevski[2], K. Dörner[2], A. M. Gañán-Calvo[8], R. Graceffa[2], M. Heymann[9], H. Fangohr[2], M. Frank[7], S. Hauf[2], A. Kaukher[2], Y. Kim[2], B. Kobe[10], J. Knoška[4,18], T. Laurus[5], R. Letrun[2], L. Maia[2], M. Messerschmidt[20], M. Metz[4], T. Michelat[2], G. Mills[2], S. Molodtsov[2], D. C. F. Monteiro[11,12], A. Münnich[2], G. Previtali[2], A. Round[2], T. Sato[2,5], R. Schubert[2], J. Schulz[2], M. Shelby[7], C. Seuring,[4,11] J. A. Sellberg[13], M. Sikorski[2], A. Silenzi[2], S. Stern[2], J. Sztuk-Dambietz[2], J. Szuba[2], M. Trebbin[14,15], P. Vagovic[2], T. Ve[16], B. Weinhausen[2], K. Wrona[2], P. L. Xavier[4,19], C. Xu[2], O. Yefanov[4], K. A. Nugent[1,17], H. Chapman[4,11,18], A. Mancuso[2,1], A. Barty[4], C. Darmanin[1*], B. Abbey[1*].

Affiliations

[1]Australian Research Council (ARC) Centre of Excellence in Advanced Molecular Imaging, Department of Chemistry and Physics, La Trobe Institute for Molecular Sciences, La Trobe University, Bundoora, VIC 3086, Australia.

[2]European XFEL, Holzkoppel 4, 22869 Schenefeld, Germany.

[3]School of Science, RMIT University, Melbourne, VIC 3000, Australia.

[4]Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany.

[5]DESY, Notkestrasse 85, 22607 Hamburg, Germany.

[6]Institute of Molecular Biology, SAS, Dubravska cesta 21, 845 51 Bratislava, Slovakia.

[7]Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA.

[8]Dept. de Ingeniería Aeroespacial y Mecánica de Fluidos, ETSI, Universidad de Sevilla, 41092 Sevilla, Spain.

[9]Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Am Klopferspitz 18, D-82152 Martinsried.

[10]School of Chemistry and Molecular Biosciences, Institute for Molecular Bioscience and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, QLD 4072, Australia.
[11]The Hamburg Centre for Ultrafast Imaging, Universität Hamburg, Luruper Chaussee 149, Hamburg 22761, Germany.
[12]Hauptman-Woodward Medical Research Institute, 700 Ellicott St., Buffalo, NY 14203, USA,
[13]Biomedical and X-ray Physics, Department of Applied Physics, AlbaNova University Center. KTH Royal Institute of Technology, SE-106 91 Stockholm, Sweden.
[14]Department of Chemistry, State University of New York at Buffalo, 760 Natural Sciences Complex, Buffalo, New York 14260, USA.
[15]Hauptman-Woodward Medical Research Institute,700 Ellicott Street, Buffalo, New York 14203, USA.
[16]Institute for Glycomics, Griffith University, Southport, QLD 4222, Australia.
[17]Laser Physics Centre, Research School of Physics Australian National University, ACT, 2601 Australia. [18]Department of Physics, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany.
[19]Max-Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 175, 22761 Hamburg, Germany.
[20]School of Molecular Science, Arizona State University, Tempe, AZ, 85281, USA.

*Correspondence to: c.darmanin@latrobe.edu.au and b.abbey@latrobe.edu.au

The pulse structure of the European X-ray Free Electron Laser (EuXFEL) enables Serial Femtosecond Crystallography (SFX) data to be collected at megahertz (MHz) repetition rates. The future capabilities of MHz repetition rate are still to be explored. The combination of the high-sensitivity, ultra-fast, AGIPD detector and the sub-microsecond inter-pulse spacing of the EuXFEL, allows multiple diffraction images to be collected from the same crystal as it traverses the X-ray beam in a liquid jet. Here we exploit the MHz pulse structure of the EuXFEL to measure diffraction patterns (up to 2.1 Å resolution) from the same protein crystal at two time-points separated by less than a microsecond. We term this approach 'Multi-Hit SFX'. This is the first indication using the MHz regime that a single crystal can be illuminated twice, and a structure can be generated from the second hit data with no signs of radiation damage visible in the Fourier difference maps. This

technique takes advantage of the larger, faster to set-up beam sizes possible at the instrument. These results pave the way towards a new technique for developing *in-situ* single crystal, sub-microsecond, timing experiments at MHz repetition rate XFELs.

Introduction

Serial femtosecond crystallography (SFX) has changed the way in which we can probe macromolecular structures and investigate their dynamics. SFX is especially well suited for structure determination with small crystals and time-resolved systems exhibiting irreversible processes, which cannot be measured using conventional synchrotron or lab-based X-ray sources[1–5]. One hurdle to implementing SFX is obtaining a large enough data set for high-resolution 3D structure determination, which historically has meant high sample consumption rates. At traditional XFEL facilities, repetition rates are low and obtaining a large data set takes time[6]. With the development of new sources like the EuXFEL[7], capable of MHz repetition rates, both data collection times and sample consumption may be significantly reduced. Another avenue of research open to MHz XFEL facilities is the ability to use the pulse structure of the source to perform novel timing experiments[7-9]. Typically, X-ray pump/X-ray probe experiments at XFELs enable the study of dynamics such as bond breaking on sub-picosecond time-scales[10-13]. The spacing between consecutive pulses at the EuXFEL, is normally sub-microsecond. This creates an opportunity to perform a new class of time-resolved experiments in which molecular dynamics occurring on sub-microsecond time-scales can be studied on the same crystal. This timing regime is ideally suited to the study of fast (sub-millisecond) dynamics in proteins including transition states (i.e. active to inactive form)[14–18], as well as helix motions[19] and side-chain rotations[20]. The unique pulse train structure of the EuXFEL offers the opportunity to collect femtosecond diffraction data with variable sub-microsecond time delays. An example problem that could benefit from multi-hit SFX is understanding the mechanism of the Bacillus subtilis response regulator, SpoOF, which is involved in sporulation. SpoOF can induce a shift in protein conformation on picosecond to millisecond time-scales[16]. Previous NMR studies have shown the secondary structure dynamics of the protein involves a complex series of movements occurring over a wide range of time-scales. These include rotation of bonds in methyl groups (thought to occur within nanoseconds) to side chain flipping of buried residues (which occurs on the scale of seconds); critical protein-protein contacts between SpoOF and its binding partners take place on millisecond time-scales[16]. Examples which maybe important for this application include studying protein-protein complexes or binding of co-factors such as ATP or NADP which induce conformational changes within proteins and result in the

transition between active to inactive forms of the enzyme which occur within millisecond time frames. Multi-hit SFX performed on sub-microsecond time-scales would be able to capture the intermediate helical/sheet conformational changes in the secondary structures of proteins prior to the important side chain flipping. The advantage of this new technique is to evolve the existing time-resolved SFX capabilities[10-13] enabling capture of intermediate states 'live' based on the timing between the pulses rather than activation/mixing times in the jet. This work also questions whether a high number of photons and a good focus from XFEL is absolutely required if we have good crystal quality.

The most common method for delivering samples to the XFEL beam is via a liquid jet, formed using a flow focusing nozzle (commonly referred to as GDVN)[21-23], in which the sample is continuously replenished with respect to the X-ray beam. Typically, no single crystal flowing in a liquid jet can contribute to more than one diffraction pattern at an XFEL due to the crystal being destroyed by each single beam interaction or the jet velocity being too fast with respect to the X-ray pulse repetition rate[24–25]. The normally ultra-intense femtosecond duration pulses available at XFELs, typically allow high-resolution diffraction data to be collected from suitably crystalline samples during a single 'hit', prior to any significant nuclear motion occurring. The first experiments at EuXFEL addressed the issue of jet explosion and its effect on the crystals. As a result, multiple jet speeds (42, 78 and 102 m/s) were tested[1]. However, we discovered using the EuXFEL it is possible, with the appropriate beam intensity and flow rate, that a single crystal may be hit multiple times as it passes through the beam without exhibiting degradation of measured diffraction intensities. Hence one infers there is also no extensive crystal damage. This "multi-hit diffraction" scenario is clearly distinct from "multi-crystal diffraction", in which two or more crystals arrive at the X-ray beam at the same time, and which can be identified in the diffraction pattern and by using multi-crystal indexing options[26–28]. Wiedorn *et. al* [1] also investigated the hit rates as a function of train pulse number for each jet speed, and confirmed that there is no bias on which train pulse number hits a crystal.

From our analysis, we can obtain high-resolution diffraction data from the first and second hits of a single crystal in a MHz SFX experiment, hence the term 'Multi-Hit SFX'. This is in stark contrast to conventional liquid-jet SFX experiments in which each single crystal gives rise to just one diffraction pattern prior to exiting the beam or being destroyed. Though it should be emphasised that the caveat to performing multi-hit SFX is that the dose from the initial hit must be below the

threshold for significant radiation damage of the crystal to occur. Whilst this likely precludes the use of nanocrystals, due to the weak diffraction from these samples, it does allow for multi-hit, femtosecond diffraction experiments to be performed on microcrystals. Microcrystals are often used for time-resolved SFX using either an optical pump or rapid mix-and-inject to initiate structural changes within the sample[11].

Multi-hit SFX can be thought of as a bridge between SFX and traditional crystallography. SFX; where thousands of crystals are hit once and destroyed, and traditional crystallography; where one crystal is continuously rotated and illuminated multiple times before significant radiation damage occurs. Previously, this phenomenon has only been detected using highly-viscous jet streams. This is because, unlike liquid jets, highly-viscous jets can flow at speeds slow enough that even at lower (e.g. non-MHz) repetition rate XFEL facilities multiple hits can be observed[29–31]. A key benefit, in addition to reducing sample consumption, is that multi-hit SFX at a MHz XFEL source can be used to probe multiple time-points sub-microseconds apart and given good crystal quality, the need for high photon counts and a good focus from the XFEL is no longer an essential requirement. The important questions we address in this paper are:

- How often are multi-hits (in this case double hits) observed under the conditions probed during the initial experiment at the SPB/SFX instrument[1,32]?
- Why do multi-hits occur and what is the influence of the experimental geometry?
- How do multi-hits impact the data quality and resolution?
- Under what experimental conditions will more than one hit of the same sample occur and how can this information be used to tailor future MHz XFEL experiments?

Based on the analysis presented by Wiedorn et al. (2018)[1] we have constructed a model that demonstrates how crystals are hit either once or twice by the beam and demonstrate that under the conditions used here the second hits do not negatively impact the quality of the data. This implies that the data from multiple hits of the same crystal can be used for structure determination. This result allow users of XFEL sources to select experimental parameters that can increase the probability of multiple hits, or at least try not to minimise them, potentially increasing the amount of usable data per crystal in an SFX experiment and paving the way to novel forms of microsecond, time-resolved measurements.

Results

EuXFEL experimental set-up and injector jet speeds

SFX was carried out on the SPB/SFX instrument[32] at the EuXFEL using compound refractive lenses (CRL) to focus the beam as described in Wiedorn et al. (2018)[1]. Lysozyme (HEWL) crystals were delivered to a 1.125 MHz XFEL beam using a flow focused jet (also referred to as GDVN),[21-,23] at three different target jet speeds for the experiment: 42 m/s, 78 m/s and 102 m/s[1]. The two slowest jet speeds were checked experimentally[1], whilst the fastest jet speed (102 m/s) was theoretically determined (see supplementary data). A summary of the liquid and gas flow rates and their equivalent jet speeds from Wiedorn *et. al.* (2018)[1] are shown in Table 1, alongside the theoretically calculated 102 m/s jet speed.

Identifying consecutive hits for a single lysozyme crystal

To assess this multi-hit phenomenon, lysozyme data from the first SPB/SFX commissioning round of the EuXFEL experiment was analysed. The X-ray data statistics corresponding to the three different jet speeds tested during this experiment are shown in Table 1. The number of images collected is compared to the number of images with hits detected by Cheetah[33] and indexed by *CrystFEL*[34,35], as well as the percentage of those patterns indexed where crystals were hit twice as they passed through the beam. A crystal was determined a 'multi-hit' if it satisfied the following criteria; i) the two hits were from consecutive pulses and ii) the patterns showed very similar crystallographic orientations (the angle between each pair of basis vectors were less than 5 degrees and the lengths of the basis vectors were in agreement to within 10%, determined using *CrystFEL*'s W*hirligig* program[34,35]).

The data shows that, as expected, when the jet speed increases the percentage of double hits decreases. For the slowest (42 m/s), intermediate (78 m/s), and fastest (102 m/s) jet speeds the percentage of double hits were 6.4%, 0.9% and 0.3%, respectively.

Crystal rotation and Bragg peak analysis

To confirm that our method for determining crystals hit twice was reliable, we calculated the change in crystal rotation for all the diffraction data collected from consecutive hits in the pulse trains. Rotation of the crystal as it 'tumbles' in the liquid jet between consecutive X-ray pulses results in small changes in the position of the Bragg peak. A significant increase in the number of crystals that had <5 degree change in orientation between consecutive images was observed for the 42 m/s

jet speed (Fig. 1 and supplementary Figs. S1 and S2). Fig 1 illustrates the change in crystal rotation between consecutive pulses as the crystal passes through the X-ray beam. A significant increase in the number of crystals with a <5 degree change in rotation can be seen, which is indicative of the fact that many of the crystals in this region are hit twice by the XFEL beam before they either exit the interaction region or are destroyed. By contrast, the number of crystals with change in rotation of >5 degrees between consecutive hits does not vary significantly. This suggests that for the 42 m/s jet speed the crystal does not rotate more than 5 degrees between the first and second hit for double hit crystals. These results verify the predictions of the *whirligig* program and show that a significant fraction of the crystals were hit twice as they traversed the X-ray interaction region. The *whirligig* program was therefore used to determine the number of multiple hits that occurred for the two faster jet speeds. Due to the lower hit rates for these two jets speeds (78 and 102 m/s) the number of diffractions patterns collected which contained Bragg peaks, is less than compared to the 42 m/s data set (Table 1).

Modelling the multi-hit scenario

Now to the question of how if, as expected, the crystal is destroyed in the beam it is possible to get more than one diffraction measurement from a single crystal. Irrespective of crystal concentration, multiple hits on a single crystal are primarily determined by the jet speed, beam focus size, and pulse repetition rate. To illustrate how multiple hits occur, a graphical representation of the crystal path through the beam based on the experimental data is presented in Fig. 2. For the slowest jet speed (42 m/s) the crystal is initially hit within the tail (lowest intensity) region of the beam followed by a second hit within the FWHM of the beam. As exposure to the second hit is within the most intense part of the X-ray beam, it leads to the destruction of the crystal and this is consistent with the observation that no third hit of the crystal was observed. As the jet speed increases to 78 m/s, multiple crystal hits become possible only within the tail regions of the X-ray beam, where the crystal is clipped by both ends of the beam (Fig. 2). Therefore, the intensities of both these hits are less than the second hit for the 42 m/s jet speed. At the fastest jet speed (102 m/s), multiple crystal hits can only occur in the extreme tail regions of the X-ray beam (Fig. 2). Therefore, the slowest jet speed (42 m/s) resulted in the greatest number of double-hits.

X-ray beam profile

The X-ray beam size and profile determines the percentage of double hits observed during the experiment and therefore is fundamental to achieving double-hits. To independently verify that our model for the double-hit condition is correct, we calculated the beam size and profile using optical

measurements accounting for the point spread function of the YAG crystal and the optical resolution of the system (see Methods). The beam profile was consistent with a Lorentzian model as observed previously at the SPB/SFX beam line. The Lorentzian full-width-at-half-maximum (FWHM) ($2\gamma = 50\%$) of each optical image was experimentally determined with a 7.5% image threshold applied to eliminate background scatter. The full width (FW) ($7.04\gamma = 82.38\%$) was then calculated and the variation in beam width was determined as shown in Fig. 3. The mean, minimum, and maximum FWHMs were calculated to be 18.7 μm, 11.9 μm and, 23.3 μm, respectively. Whilst the mean, minimum, and maximum, FWs were calculated to be 65.8 μm, 41.7 μm and 82.1 μm, respectively. The experimentally determined beam size was consistent with the optical measured X-ray beam size, which shows double hits are possible.

Lysozyme crystal movement through the X-ray beam

The crystal trajectory through the X-ray beam can be calculated based on the experimental and theoretical jet speeds provided in Table 1. The minimum distance a lysozyme crystal of length 8 μm (with respect to the flow direction) could clear between two X-ray pulses whilst having at least 1 μm of crystal within the beam window and with an inter-pulse spacing of 886 ns was calculated. For the 42 m/s jet, a crystal will clear a minimum distance of 35.4 μm, whereas for the 78 m/s jet, a crystal will clear a minimum distance of 56.7 μm and, at the fastest jet speed 102 m/s, the crystal will clear a minimum distance of 79.9 μm. By minimum distance, we refer to the distance calculated using the minimum jet speed including a 5% margin of error. This error is consistent with the observation that hits are observed with the 102 m/s jet speed (Table 1). Fig. 2 shows a schematic of how far the crystal has travelled at each jet speed. Based on this information it confirms that for each of the three jet speeds, the X-ray beam diameter was sufficiently large that double hits occur, which supports our experimental observations.

Data quality check

Quality control checking of the data is necessary to ensure we don't see any degradation in the second hit data with respect to the first hit data. To check this, we compared the Bragg intensities and resolution of the two data sets (first and second hit on the same crystal). The crystal diffraction data was used to generate three independent powder plots (first hits, second hits, and single hits). For the slowest jet speed used in this experiment (42 m/s), we observe that the integrated intensity of the first of the double hits was lower than the second hit, with the integrated intensity for the second hit more similar to that measured from crystals hit just once by the X-ray beam (Fig. 4a and 5a). This supports the interpretation that the second hit occurs within the FWHM of the X-ray

beam. As the jet speed increases the diffraction intensity profiles change generating a larger difference in the integrated intensity between the double hit crystal and the single hit crystal. At a jet speed of 78 m/s, the integrated intensity profiles for the first and second hit of the double hit crystal is decreased further and found to be more similar in intensity to each other rather than the single hit crystal intensity profile (Fig. 4b, 5a). This supports the fact that the crystal is hit within both the tail regions of the X-ray beam, either side of the FWHM. At the fastest jet speed (102 m/s) the difference in integrated intensities between the double hits and single hits is even more pronounced. The integrated intensity profiles of all double hit crystals are very similar and much less intense than the single hit crystals (Fig. 4c, 5a), which is again consistent with our model as shown in Fig. 2. Irrespective of the integrated intensity data, if we specifically select for only high-resolution reflections in the data (Fig. 5b), little difference ($\leq 5\%$) is detected in the data obtained above 2.5 Å resolution, when comparing crystal hits for the same jet speed. When comparing the data from the different jet speeds, an overall reduction in the percentage of high-resolution peaks can be seen for the 42 m/s jet speed, however, this can be explained by the larger quantity of data obtained for the slowest jet speed compared to the two higher jet speeds (Table 1). We would expect the higher jet speeds to have been comparable to the 42 m/s jet if the equivalent amount of data was collected. To further check if radiation damage has occurred within the crystal after the first hit, we compared two data sets, single hit crystal data (where crystals where only hit once) versus the second hit crystal data (where the same crystal was hit for the second time). Comparisons of the two structures showed very little differences. Importantly, closer examination of the disulfide bonds, which are highly prone to radiation damage, also showed no evidence of radiation damage in the Fourier difference map (Fig. 6A). A statistical comparison between the single hit and second hit data (in the double hit pair) show the data quality is similar (Fig. 6 and Table 2). The Wilson plots for both data sets were included in the supplementary file (Supplementary Fig. S4). CC* is a commonly used metric to assess data quality for structure determination. If we compare half the data set of the first hits to half of the data sets of the second hits, there are double the reflections in the second hit data set and the CC* are comparable (0.85 from a total of 6506 reflection pairs) further indicating there is no significant change in quality between the two data sets (Fig. 6B and Table 2).

Discussion

During this experiment, different jet speeds were tested to study the effect of jet explosion at the EuXFEL. The jet explosion begins at the centre of the interaction region and produces an initial

gap that expands in the vertical direction aligned with the liquid jet. From our observations, the gap initiates in most cases from the vaporization of a jet slice with a thickness about 1/5 to 1/4 of the jet diameter, and initially expands according to a time power law that has been investigated.[36] That power law yields initial expansion velocities of the order of 200 m/s for our experiments, which decay approximately as $(t/t_0)^{-3/7}$, where t_0 is approximately to 1 ns in our experiments. Thus, if the crystal is hit near the boundaries of the irradiated jet segment, the gap will not reach the crystal before the next pulse arrives. It is, therefore, important to establish the pulse power, beam size, jet diameter, and jet speed to avoid the interaction of the crystal with the expanding front of the opening gap formed by a previous pulse. Nevertheless, for the jet speeds tested in this experiment we did not see an effect from jet explosions. In these conditions it was indeed possible to collect high quality data on the same crystal in a subsequent pulse. This is further supported by the results of Wiedorn *et. al.*[1] which shows the jet recovery for each jet speed tested and the Yefanov *et. al*[9] results which compared structures from different pulses within the Megahertz train with no observable difference between the first and any subsequent pulse seen.

The results from this experiment represent the first demonstration of multi-hit SFX using a flow-focusing injector. The same experiment is not currently possible at lower repetition rate sources with GDVNs, as the unique pulse structure of the MHz XFEL source is required for these experiments. Furthermore, the new focusing system now installed at SPB/SFX allows tuning of the horizontal and vertical focus size independently and could be used to maximise double hits. For example, by using a vertical line focus with a similar width to the jet and a height, tuned to a length as selected for by the distance the crystal will travel for the given jet speed between X-ray pulses. Hence, multi-hit SFX could be a standard mode of operation for time-resolved experiments on microcrystals with this configuration. Our data shows that depending on the jet speed and beam profile the second hit in the double hit series can be more intense than the first with comparable integrated intensities to that of the single hit crystals. Our observations and analysis show that the initial interaction of the X-ray beam with the crystal must occur within the low intensity region of the beam profile, i.e. within the 'tail' regions in order for a second interaction to occur. The experimental value of the beam size using the YAG (accounting for the PSF and optical resolution) is consistent with the beam size required to generate double hits.

The time interval between pulses at the EuXFEL is extremely short when compared to other non-MHz XFEL sources, however sub-microsecond pulse spacing is still long on the time-scale of radiation damage[37,]. Crystal dose estimates within the tail regions, calculated based on the

measured beam and crystal size[38,39], give an average absorbed dose of 0.3 Gy for the first hit compared to 1.5 Gy within the FWHM of the beam. Typically, protein crystals can withstand a radiation dose of 30 MGy[40,41] under cryo-cooled conditions as it slows the effects of radiation damage, whilst at room temperature the limit decreases by approximately two orders of magnitude, i.e. ~ 0.2 MGy, as a result of the diffusion of free radicals[41-44]. In addition, room temperature radiation damage becomes dependent on dose rate[45]. It has been reported that this dose limit can be increased (30-150 MGy)[3,46,47] by the introduction of highly intense femtosecond pulses at XFEL facilities due to the pulse duration being short enough to significantly reduce the effects of radiation damage under room temperature conditions[37,48]. Radiation damage effects at XFELs is still an active area of research[49–51] and is dependent on both dose rate, and the specific choice of sample. However, in the present case, the dose estimates calculated for the double hit crystals measured here fall well below published thresholds for radiation damage and is supported by the fact that no significant degradation in the quality of diffraction was observed in the second hit structure. With this in mind, we have developed an explanation for the observed integrated intensity data considering the distance the crystal travels between X-ray pulses for the three different jet speeds.

Our interpretation is consistent with the observed changes in the measured integrated intensity profile and it explains why for the 42 m/s jet speed, the second hit has a similar intensity profile to the single hit crystal data (Fig. 4). This was initially contrary to expectations based on the assumption that the crystal would be damaged during the first hit reducing the measured intensity of the second hit. For the two higher jet speeds, the inter-pulse spacing and beam size are such that to observe double hits the crystal must only interact with the tails of the beam for both the first and second hit (Figures 2 & 3). In other words, for the faster jet speeds, it is not possible to observe a double hit where one of the pair of hits is the result of an interaction within the central portion of the beam. This is consistent with a weaker measured intensity for this data. However, it is important to note that even in the tails of the beam, there is sufficient intensity at the XFEL to obtain high-resolution diffraction from micron sized crystals, as has been observed in previous experiments[3].

For the 78 m/s and 102 m/s jet speeds we demonstrate the estimated area of the beam ‘seen’ by the crystal can be determined simply by considering the distance the crystal travels between pulses. Whereas for the slowest jet speed (42 m/s) the second hit must occur within the high flux interaction region to generate similar diffraction to the single hits and guarantee destruction of the crystal (Fig. 2). Similarly, for the faster jet speeds, the first hit occurs in the tail region of the beam where the incident flux is low and does not significantly damage the crystal which reduces the chance of

double hits occurring. Using the beam size, jet speed and an understanding of how these parameters relate to the region of crystal interaction, it is now possible to determine parameters, which either minimise or maximise the chances of multi-hit SFX. This is critical, particularly for current and future experiments at the EuXFEL, where we can investigate alternate time-resolved techniques using the unique time structure of the source. To optimise for multiple hits it is possible to tailor the jet speed and beam focus. More importantly, the spacing between the successive pulses in a EuXFEL pulse train can be modified to allow for multi-hit SFX to probe different time regimes. Multi-hit SFX will be a critical consideration for when the EuXFEL operates at planned 4.4 MHz repetition rates, where multiple hits are significantly more frequent (and potentially unavoidable) than at the 1.125 MHz that the current experiment was performed at. To predict how multi-hit SFX at 1.125 MHz compares to 4.4 MHz, we have provided a figure to estimate optimal, beam size and jet speed to achieve multiple hits. Fig. 7 demonstrates the criteria to obtain multi-hit regime for the current rate (1.1 MHz, Fig. 7a) and future rate (4.4 MHz, Fig. 7b) of the EuXFEL.

In addition to the integrated intensity, we also studied the resolution of the diffraction data from both single and double hits. A direct comparison shows that whilst the integrated intensity varies, there are no significant differences in the resolution of single hits compared to the first and second hits that make up the double hit pair and this is supported by the structural results that no radiation damage is present. This suggests that all the data collected during this experiment can be merged into a single dataset for structural analysis without detriment to the quality, allowing us to increase the amount of data without increasing crystal consumption. Furthermore, it also implies that under the experimental configuration used here, the shock-waves generated in the liquid jet, as a result of the first hit, do not appear to impact the crystal quality or packing. If the crystal quality was affected by the first X-ray pulse, we would expect the higher resolution diffraction peak(s) to disappear; indicative of crystal damage, however this was not observed which supports the fact that no radiation damage occurred. A comparison of the single hit and second hit molecular structures confirm this, showing no signs of radiation damage in the Fourier difference maps.

With reference to the key questions posed at the start of the paper, we find that in the experimental configuration used here, double-hits occur between 0.3 and 6.3% of the time depending on beam-size and jet speed, with the slowest jet speed being optimal for multi-hit SFX. This percentage is expected to increase at 4.4 MHz. The quality and frequency of double-hits is dominated by the experimental geometry, though for smaller beam sizes the effect of radiation damage will become

increasingly significant so larger beams may be more desirable for time-resolved multi-hit SFX experiments. Under the present experimental conditions, we observed no evidence of radiation damage, or damage from XFEL induced shock-waves when looking at the resolution and integrated intensity data or the molecular structures.

In summary, multi-hit SFX will be developed as a new mode of sub-microsecond time-resolved serial crystallography, taking advantage of the unique time structure of MHz XFEL sources and large beam focuses. This method will become increasingly important as additional MHz XFEL sources come on-line, and as facilities turn to faster and faster MHz repetition rates. The beam sizes used for this experiment can be easily realised within the standard set-up at SFX/SPB and requires no modification of beam-line or upstream optics. In addition, control over the initial sample state (structure) and subsequent dynamics of the injected crystals can be achieved via a rapid mix-and-inject approach using a chemical trigger. This provides an opportunity for measuring molecular dynamics at intermediate time-scales between those currently achievable with SFX (femtosecond regime) and serial millisecond crystallography (SMX) (millisecond regime). Through analysis of the individual Bragg peaks, by comparing plots from the first and second hits as well as successfully solving the structure based solely on second hit data, we demonstrate that for the system studied here, radiation damage effects are not observed. Most importantly, the second hit structure did not reveal any radiation damage surrounding the disulfide bonds. It should be emphasised however, that for proteins that are particularly radiation sensitive, the effects of radiation damage (especially specific damage) may not be easily separable from the molecular dynamics. Finally, we have outlined the criteria to observe multi-hits in terms of jet speed and beam size (Fig. 7) and confirmed, through discussions with beam-line staff, that these parameters may be readily achieved at the European XFEL.

Methods

Further details on the sample preparation and delivery, as well as the experimental set-up and initial data analysis can be found in ref [1]. The beam focusing at the SPB/SFX beam-line for this experiment was achieved using compound refractive lenses (CRLs). The crystal size varied between 6 x 6 x 6 µm and 8 x 8 x 8 µm as characterised via an optical microscope. A subset of the data generated and analysed from the first EuXFEL experiment (run numbers r0069-r0082, r0145-r0150, r0153) in the form of *CrystFEL* stream files were used to identify crystals which had been

hit consecutively, it is this subset that was used for the analysis presented here. The data used for this paper is available from CXIDB-ID80 (www.cxidb.org/id-80.html).

Identification of multi-hits in the data

The stream files with the already indexed crystal data were used to identify the crystals classified as multi-hits. The stream files were filtered for these multi-hits using the *whirligig* script from the *CrystFEL* crystallography suite[34,35], which defines crystals hit twice as those that had similar crystallographic orientations in consecutive frames. Based on our analysis of the relative change in crystal orientation between consecutive hits we confirmed that for multi-hit crystals the change in angle between each pair of basis vectors was less than 5 degrees. In addition, the lengths of the basis vectors had to be within 10% agreement. Filtering using the *whirligig* program based on these criteria alone does not account for the pulse/train format of the EuXFEL. Therefore, python code was implemented to filter out any adjacent frames that were not from the same pulse train. This data was further sorted into three separate categories: data from crystals hit just once, data from the first hit of crystals hit twice, and data from the second hit of crystals hit twice. These three stream files were then used as the input files for further analysis in Python 3, using multiple parameters (integrated intensity, a^*, b^*, c^*, h, k, l) within the stream files to calculate $1/d$, normalised integrated intensity plots.

Python 3 was also used to analyse the *CrystFEL* stream files, where the change in orientation between all consecutive images was calculated in an identical manner to the *whirligig* program. The output was plotted as a histogram, showing the change in angle (degrees) for consecutive images to confirm those consecutive hits that were selected as our double hit data sets were truly from the same crystal.

Analysis of data quality

Merging and scaling of the Bragg peaks were performed using Partialator in the *CrystFEL suite*[34,35]. Figures of merit were calculated using compare_hkl (Rsplit, $CC_{1/2}$, CC*) and check_hkl (SNR, multiplicity, completeness), also using the *CrystFEL* suite[34,35]. To generate a complete second hit data-set, experimental data from all jet speeds, were merged. The statistics for the single hits and second hits data sets were generated (Table 2) and Wilson plots for both data sets are shown in the supplementary (Supplementary, Fig. S4).

Structure refinement was performed in CCP4i2 using the MTZ output from CrystFEL. A solvent free version of lysozyme (pdb code 6FTR) was used as the initial starting model for Molecular Replacement in Phaser[52] and the R_{free} flags were generated (utilizing 10% of the data) followed by iterative cycles of Refmac5[53] refinement and rebuilding of the model in Coot[54]. MolProbity[55] and Xtriage (Phenix)[56] tool was used to validate the model.

To assess the quality of the data a Fourier difference (F_o-F_c) map was generated using Refmac5[53] (Fig. 6A).

X-ray Beam Profile Analysis

To find the optimum focus, a photon energy scan of the accelerator was carried out. The nominal energy for the SFX data collection was 9.232 keV. The beam size and profile were estimated based on 6773 individual YAG images collected using the in-line microscope positioned within in the chamber. The optical images were generated from single shots using a 15 µm thick Ce:YAG screen (refer to Supplementary Fig. S3). The point spread function (PSF) of the YAG was determined based on published estimates[57] (see Supplementary Fig. S3) and estimated to be approximately 2 µm. Using Edmund standards, the optical microscope resolution was determined to be 8 µm. A Lorentzian distribution was fitted to the optical images using Python code. The actual beam size was determined via a convolution taking into account the PSF for the YAG and optical microscope. A Lorentzian FWHM (FWHM = 2ɣ) of each image was determined using a 3 × 3 median filter and a 7.5% noise threshold. Furthermore, the Full Width (FW = 7.04ɣ) of 82.38% was determined based on the Lorentzian above the noise threshold; and for each image it was calculated and plotted along with the overall mean, minimum and maximum FW.

References

1. Wiedorn, M. O. *et al.* Megahertz serial crystallography. *Nat. Commun.* **9**, 4025 (2018).
2. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011).
3. Boutet, S. *et al.* High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* **337**, 362–364 (2012).
4. Gati, C. *et al.* Atomic structure of granulin determined from native nanocrystalline granulovirus using an X-ray free-electron laser. *Proc. Natl. Acad. Sci.* **114**, 2247–2252 (2017).
5. Suga, M. *et al.* Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99–103 (2015).
6. Liu, H. & Lee, W. The XFEL Protein Crystallography: Developments and Perspectives. *Int. J. Mol. Sci.* **20**, 3421 (2019)
7. Pandey, S. et al. Time-resolved serial femtosecond crystallography at the European XFEL. Nat. Methods 17, 73–78 (2020)
8. Gisriel, C. et al. Membrane protein megahertz crystallography at the European XFEL. Nat. Commun. 10, 1–11 (2019).
9. Yefanov, O. et al. Evaluation of serial crystallographic structure determination within megahertz pulse trains. Struct. Dyn. 6, 064702 (2019).
10. Aquila A, Hunter MS, Doak RB, Kirian RA, Fromme P, White TA, et al. Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Optics Express*. 20(3):2706–2716 (2012).
11. Tenboer J, Basu S, Zatsepin N, Pande K, Milathianaki D, Frank M, et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* 346(6214):1242–1246 (2014).
12. Allen M. Orville Entering an era of dynamic structural biology… *BMC Biology,* **16**, 55 (2018)

13. Schmidt, M. Time-Resolved Macromolecular Crystallography at Pulsed X-ray Sources. *Int. J. Mol. Sci.* ***20***, 1401 (2019) doi:10.3390/ijms20061401

14. Grünbein, M. L. *et al.* Megahertz data collection from protein microcrystals at an X-ray free-electron laser. *Nat. Commun.* **9**, 3487 (2018).

15. Volkman, B. F., Lipson, D., Wemmer, D. E. & Kern, D. Two-State Allosteric Behavior in a Single-Domain Signaling Protein. *Science* **291**, 2429–2433 (2001).

16. Moy, F. J. *et al.* Assignments, Secondary Structure, Global Fold, and Dynamics of Chemotaxis Y Protein Using Three- and Four-Dimensional Heteronuclear (13C,15N) NMR Spectroscopy. Biochemistry, **33**, 10731-10742 (1994).

17. Feher, V. A. & Cavanagh, J. Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. *Nature* **400**, 289–293 (1999).

18. Nicolaï, A., Delarue, P. & Senet, P. Decipher the Mechanisms of Protein Conformational Changes Induced by Nucleotide Binding through Free-Energy Landscape Analysis: ATP Binding to Hsp70. *PLOS Comput. Biol.* **9**, e1003379 (2013).

19. Xie, M. *et al.* Functional protein dynamics on uncharted time scales detected by nanoparticle-assisted NMR spin relaxation. *Sci. Adv.* **5**, eaax5560 (2019).

20. Hansen, A. L. & Kay, L. E. Quantifying millisecond time-scale exchange in proteins by CPMG relaxation dispersion NMR spectroscopy of side-chain carbonyl groups. *J. Biomol. NMR* **50**, 347–355 (2011).

21. Oberthuer, D. *et al.* Double-flow focused liquid injector for efficient serial femtosecond crystallography. *Sci. Rep.* **7**, 44628 (2017).

22. Gañán-Calvo, A. M. Generation of Steady Liquid Microthreads and Micron-Sized Monodisperse Sprays in Gas Streams. *Phys. Rev. Lett.* **80**, 285–288 (1998).

23. DePonte, D. P. *et al.* Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Phys. Appl. Phys.* **41**, 195505 (2008).

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

25. Stan, C. A. *et al.* Liquid explosions induced by X-ray laser pulses. *Nat. Phys.* 12, 966–971 (2016)

26. Beyerlein, K. R. *et al.* FELIX: an algorithm for indexing multiple crystallites in X-ray free-electron laser snapshot diffraction images. *J. Appl. Crystallogr.* **50**, 1075–1083 (2017).

27. Ginn, H. M. *et al.* TakeTwo: an indexing algorithm suited to still images with known crystal parameters. *Acta Crystallogr. D. Struct. Biol.* **72**, 956–965 (2016).

28. Gildea, R. J. *et al.* New methods for indexing multi-lattice diffraction data. *Acta Crystallogr. D Struct. Biol.* **70**, 2652–2666 (2014).
29. Liu, W., Ishchenko, A. & Cherezov, V. Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nat. Protoc.* **9**, 2123–2134 (2014).
30. Nogly, P. *et al.* Lipidic cubic phase serial millisecond crystallography using synchrotron radiation. *IUCrJ,* **2**, 168–176 (2015).
31. Weierstall Uwe. Liquid sample delivery techniques for serial femtosecond crystallography. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20130337 (2014).
32. Mancuso, A. P., et.al. The Single Particles, Clusters and Biomolecules and Serial Femtosecond crystallography instrument of the European XFEL: initial installation. *J. Synchrotron Radiat*. **26**, (2019).
33. Barty, A. *et al.* Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.* **47**, 1118–1131 (2014).
34. White, T. A. *et al.* CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* **45**, 335–341 (2012).
35. White, T. A. *et al.* Recent developments in CrystFEL. *J. Appl. Crystallogr.* **49**, 680–689 (2016).
36. Ganan-Calvo, A. M. The scaling of exploding liquid jets under intense X-ray pulses. *PRL* **123**, 064501-5 (2019).
37. Holton, J. M. A beginner's guide to radiation damage. *J. Synchrotron Radiat.* **16**, 133–142 (2009).
38. Coughlan, H. D. *et al.* Bragg coherent diffraction imaging and metrics for radiation damage in protein micro-crystallography. *J. Synchrotron Radiat.* **24**, 83–94 (2017).
39. Marman, H., Darmanin, C. & Abbey, B. The Influence of Photoelectron Escape in Radiation Damage Simulations of Protein Micro-Crystallography. *Crystals* **8**, 267 (2018).
40. Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proc. Natl. Acad. Sci.* **103**, 4912–4917 (2006).
41. Owen, R. L. *et al.* Outrunning free radicals in room-temperature macromolecular crystallography. *Acta Crystallogr. D.* **68**, 810–818 (2012).
42. Gotthard, G., Aumonier, S., De Sanctis, D., Leonard, G., von Stettena, D., Royant, A. Specific radiation damage is a lesser concern at room temperature. *IUCrJ*, **6**, 665–680, (2019). https://doi.org/10.1107/S205225251900616X

43. de la Mora, E., Coquelle, N., Bury, C.S., Rosenthal, M., Holton, J.M., Carmichael, I, Garman, E.F., Burghammer, M., Colletier, J.F., Weik, M. Radiation damage and dose limits in serial synchrotron crystallography at cryo- and room temperatures *PNAS*, **117**, 4142-4151, (2020) https://doi.org/10.1073/pnas.1821522117
44. Ebrahim, A., Moreno-Chicano, T., Appleby, M.V., Chaplin, A.K., Beale, J.H., Sherrell, D.A., Duyvesteyn, H.M.E., Owada, S., Tono, K., Sugimoto, H., Strange, R.W., Worrall, J.A.R., Axford, D., Owenb, R.L., Hougha, M.A. Dose-resolved serial synchrotron and XFEL structures of radiation-sensitive metalloproteins. *IUCrJ*, 6, 543–551, (2019). https://doi.org/10.1107/S2052252519003956
45. Kern, J. *et al.* Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature. *Science* **340**, 491–495 (2013).
46. Liu, W. *et al.* Serial Femtosecond Crystallography of G Protein–Coupled Receptors. *Science* **342**, 1521–1524 (2013).
47. Lomb, L. *et al.* Radiation damage in protein serial femtosecond crystallography using an x-ray free-electron laser. *Phys. Rev. B Condens. Matter Mater. Phys.* **84**, 214111 (2011).
48. Nass, K. Radiation damage in protein crystallography at X-ray free-electron lasers. *Acta Crystallogr. Acta D.* **75**, 211–218 (2019).
49. Abbey, B. *et al.* X-ray laser–induced electron dynamics observed by femtosecond diffraction from nanocrystals of Buckminsterfullerene. *Sci. Adv.* **2**, e1601186 (2016).
50. Amin, M., Badawi, A. & Obayya, S. S. Radiation Damage in XFEL: Case study from the oxygen-evolving complex of Photosystem II. *Sci. Rep.* **6**, 36492 (2016).
51. Standfuss, J. Membrane protein dynamics studied by X-ray lasers–or why only time will tell. *Current Opin. Struc. Biol***.** **57**, 63-71, (2019) https://doi.org/10.1016/j.sbi.2019.02.001
52. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. **40**, 658–674 (2007)
53. G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, and A. A. Vagin, *REFMAC*5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr*. 67, 355 (2011).
54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr*. **66**, 486–501 (2010)
55. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
56. Afonine, P., Grosse-Kunstleve, R.W., Chen,V.B., Headd, J.J. Moriarty, N.W. Richardson, J.S. Richardson, D.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D. Phenix.model_vs_data: a

high-level tool for the calculation of crystallographic model and data statistics, *J. App. Cryst.* **43**, 669-676 (2010) https://doi.org/10.1107/S0021889810015608

57. Koch, A., Raven, C., Spanne, P., Snigirev, A. X-ray imaging with submicrometer resolution employing transparent luminescent screens. *J. Opt. Soci. America A.* **115**, 1940-1951, (1998)

Acknowledgments

General: We acknowledge European XFEL in Schenefeld, Germany, for provision of X-ray free electron laser beam-time at Scientific Instrument SPB/SFX and use of the XBI biological sample preparation laboratory, enabled by the XBI User Consortium and would like to thank the instrument group and facility staff for their assistance and everyone who attended the first experiment.

Funding: We would also like to acknowledge the funding support for B.A., C.D, K.A.N and S.H from Australian Research Council Center of Excellence in Advanced Molecular Imaging (CE140100011), www.imagingcoe.org and the Australian Nuclear Science and Technology Organisation (ANSTO); the International Synchrotron Access Program (ISAP) managed by the Australian Synchrotron, part of ANSTO, and funded by the Australian Government. M.L.S., M.A.C. and M.F. performed, in part, under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. M.L.S., M.A.C. and M.F. were supported by the NIH grant 1R01GM117342-01. I.B acknowledges support by Ministry of Education, Science, Research and Sport of the Slovak Republic and by grant from the Slovak Research and Development Agency under contract APVV-14-0181. M.L.S., M.A.C. and M.F. performed, in part, under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. M.L.S., M.A.C. and M.F. were supported by the NIH grant 1R01GM117342-01. T.V is a recipient of ARC DECRA (DE170100783). B.K work is funded by National Health and Medical Research Council (NHMRC) grants 1107804 and 1160570 to BK and TV; BK is Australian Research Council (ARC) Laureate Fellow (FL180100109). F.C-M and A.M.G-C acknowledge Ministerio de Economía y Competitividad (Spain), Plan Estatal 2013-2016 Retos, project DPI2016-78887-C3-1-R. S.B.

620 acknowledged the Helmholtz Initiative and Networking Fund through the Young Investigators
621 Program and the Deutsche Forschungsgemeinschaft SFB755/B03.
622
623 **Author contributions:** The experiment was conceived by H.C., A.B. and A.P.M. as an open
624 experiment for the EuXFEL user community. A.B. coordinated and led the collaboration. B.A. and
625 C.D conceived the idea of studying the multi-hit phenomena and edited the manuscript. S.H.
626 developed the model for double hit scenario and wrote the manuscript. H.K, A.V.M, M.H-J and
627 K.G. contributed to the beam profile data collection and analysis. M.O.W, A.B., A.P.M., K.A.N.
628 were involved in discussions. R.B. was the instrument scientist responsible for this experiment.
629 A.P.M. led development of the SPB/SFX instrument. M.O.W and S.B. led the sample injection
630 team. C.D., D.O., R.S., I.B., M.A.C., M.F., M.S., T.V. B.K., K.D., R.L., G.P., S.H., M.M., G.P.,
631 P.L.X., H.F. and C.S. contributed to sample preparation. S.B, A.B., D.C.FM., M.T., M.O.W., J.S.,
632 J.B., R.G., H.F., and L.A. contributed to sample injector delivery. A.B led the data processing team
633 and H.C., M.S., O.Y., A.M., K.A., A.T., and P.V., contributed to data processing. A.P.M., R.B.,
634 J.S., S.S., K.G., Y.K., M.M., G.M., A.R., P.V., B.W., and J.S., developed and operated the
635 SPB/SFX instrument at EuXFEL. F.C-M., A.M.G-C., S.M., J.A.S., B.A. M.A.C., M.F., and M.S.,
636 participated in the XFEL 2012 experiment. T.L., K.W., L.M., G.P., N.A-Q., S.B., T.M., C. Dan.,
637 A.S., C.X., S.Ha., developed XFEL DAQ and controls. J.S-D., T.L., A.K. and A.M., developed and
638 responsible for AGIPD. The manuscript was written by S.H., C.D. and B.A. with input from all
639 authors.
640
641 **Competing interests:** The authors declare no competing interest.
642
643
644
645
646
647
648
649
650
651
652
653

Figures and Tables

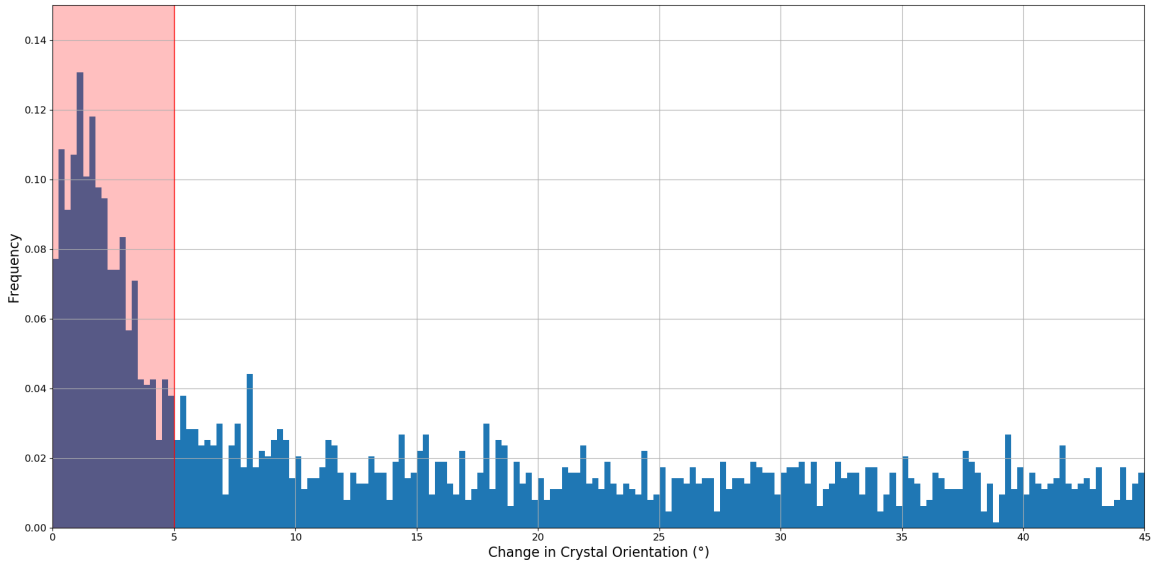


Fig. 1. Frequency of diffraction patterns as a function of the relative change in crystal rotation. This is characterized by reciprocal space vector $\vec{a}$, between consecutive diffraction measurements (separated by 886 ns) within the X-ray pulse train. The liquid jet flow rate was 42 m/s. An increase in the frequency of images with a change in orientation of less than 5 degrees can be observed due to the occurrence of double hits.

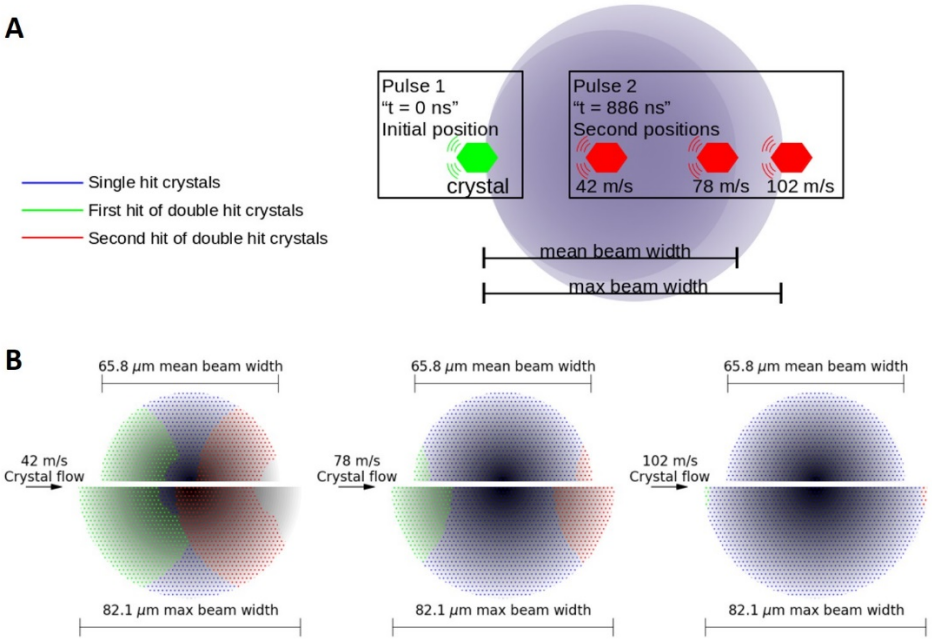


Fig. 2. Distance travelled by the crystal. **A**. Schematic (not to scale) illustrating the minimum distances travelled by an 8µm crystal for the different jet speeds overlaid with the average beam FW (dark purple shaded region) and the max beam FW (light purple shaded region). All three jet speeds are represented, the green crystal depicts the initial position, and the red crystal illustrates how far the crystal travels after the first hit for 42 m/s, 78 m/s, and 102 m/s jet speeds. **B.** Schematic representation of the crystal path through the X-ray beam for each of the three jet speeds for the mean beam FW (upper half) and maximum beam FW (lower half). The beam profile (shaded grey), is overlaid with the regions that the crystal travels through for the single hits (blue) as well as the first (green) and second (red) hits of the double hit crystal. Note for 42 m/s and 78 m/s, regions where no hits occur are possible.

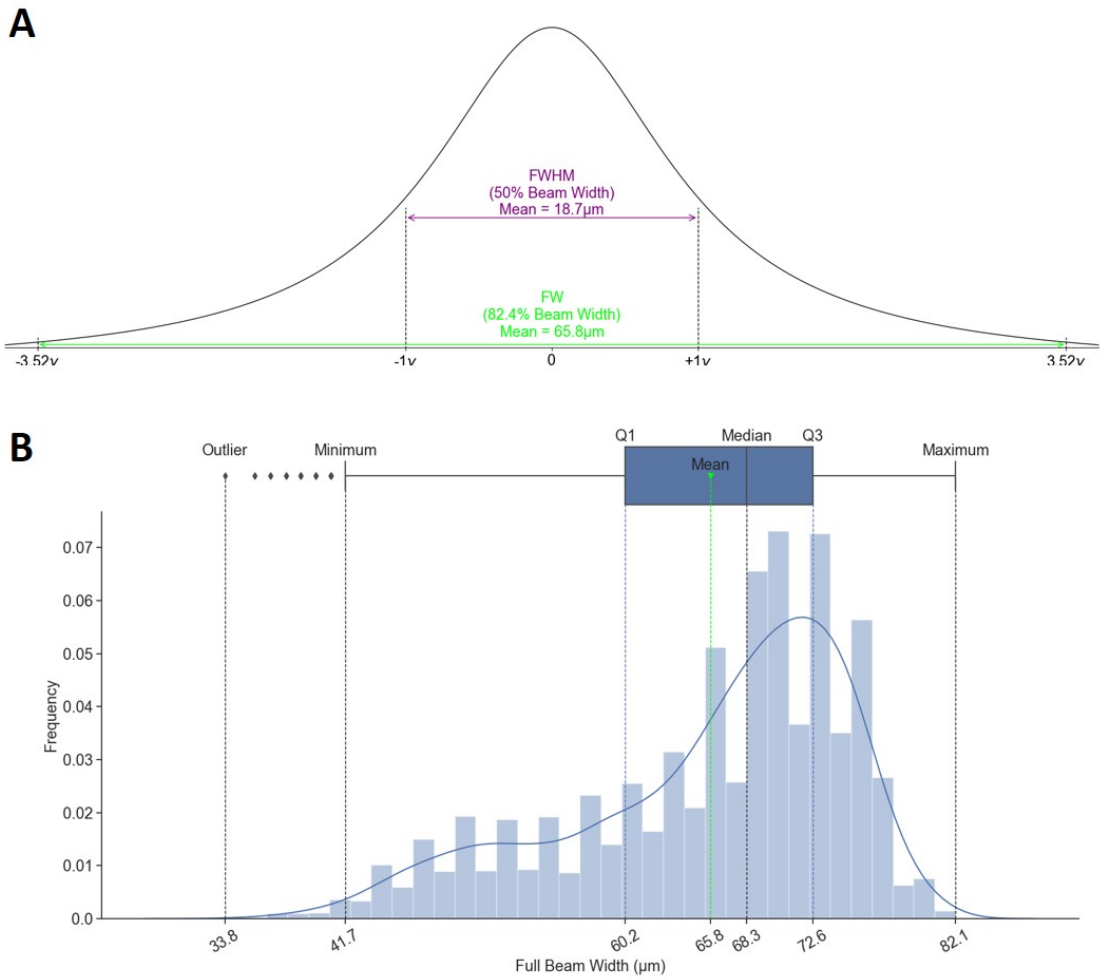


Fig. 3. Characteristic beam profile. **A.** A histogram showing the X-ray Beam profile. The beam profile was modelled using a Lorentzian distribution with a Full Width Half Maximum (FWHM) = 2$\boldsymbol{\gamma}$ (50% of beam) and Full Width (FW) = 7.04$\boldsymbol{\gamma}$ (82.4% of beam). **B.** The Lorentzian distribution used to determine the FWHM and FW is shown for all 6773 YAG images obtained. The furthermost outlier, minimum, Q1 (25th percentile), mean, median, Q3 (75th percentile), and maximum have been indicated. To obtain the FWHM and FW for each YAG image, a 7.5% noise threshold was applied to the image combined with a 3 × 3 median filter to account for the noise.

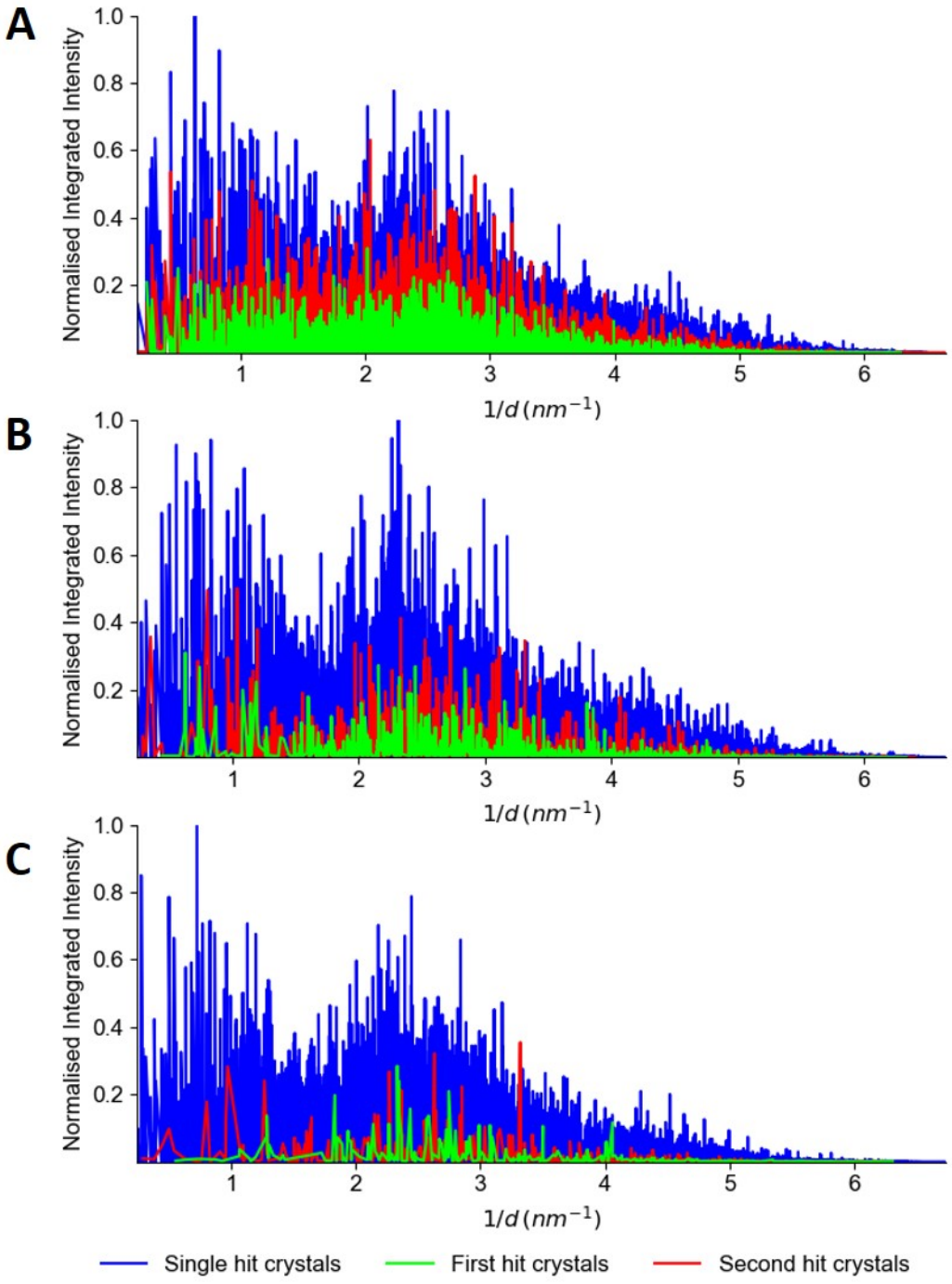


Fig. 4. Normalized integrated intensity plots. Integrated intensities were extracted from the data, normalized and plotted against 1/*d* (where *d* is the lattice spacing) for the **a.** 42 m/s jet data, **b**. 78 m/s jet data, and **c**. 102 m/s jet data. Blue represents data for single hit crystals only; green represents the first hit of the double hit crystal; red represents the second hit of the double hit crystal. A threshold of I/sig(I) > 2 was applied to the analysis.

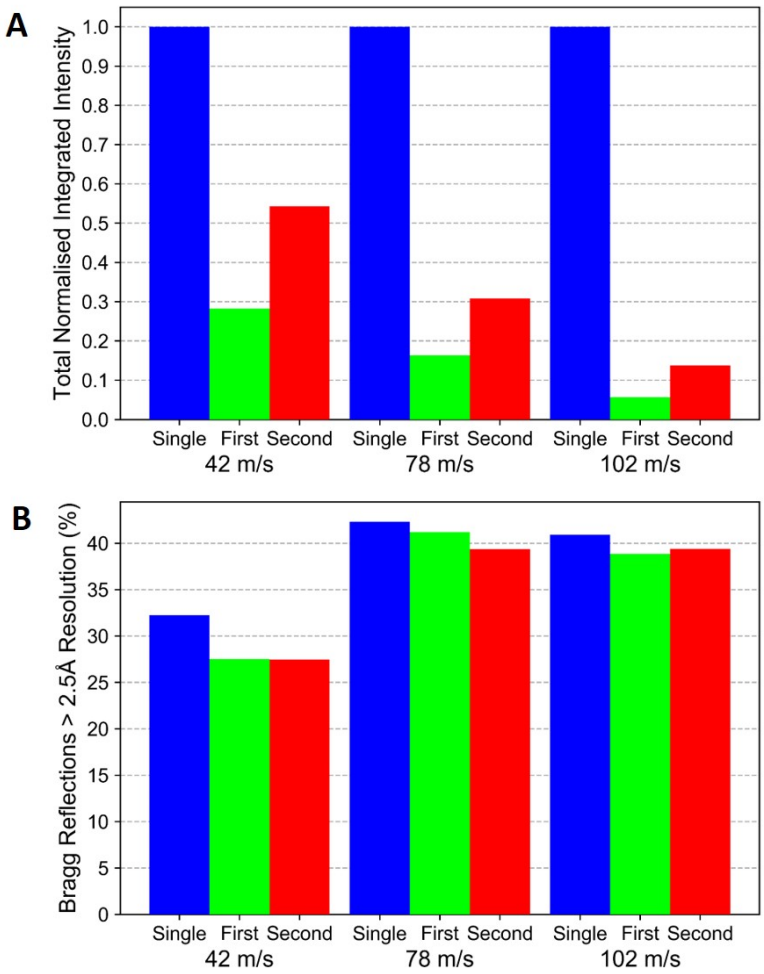
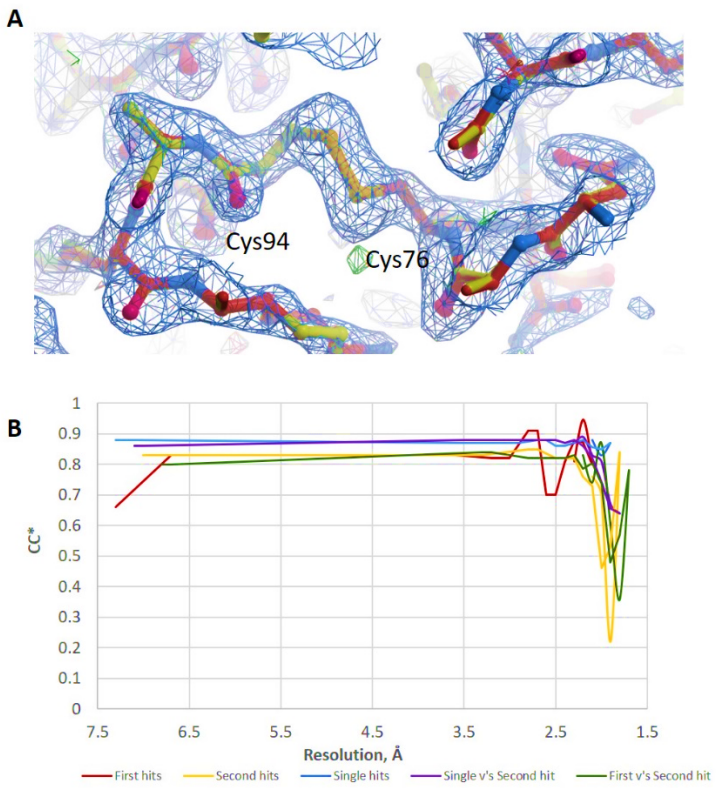


Fig. 5. Comparison of the data between the single and double hit crystals. A. The area under the curves in Fig. 4 were obtained and are seen here in order to more clearly show the differences. For the slowest jet speed, 42 m/s, the second hit had the most similar results compared to the single hit data as both crystals are hit within the FWHM region of the X-ray beam. For the faster jet speeds (78 and 102 m/s) the first and second hit data are similar in intensity as the crystal is hit within the tail region of the x-ray beam and not within the FWHM as occurs for the single hit crystal. **B.** This shows the percentage of the reflections with a better than 2.5 Å resolution for each hit type and jet speed. Note: The resolution for the fast jets is perceived to be higher in this figure but this is not statistically significant as the amount of data for each jets speed varies greatly, as denoted in Table 1. The data-sets were also collected on different days which resulted in varying beam intensities. A threshold of I/sig(I) > 2 was applied to the analysis.



718

719 **Fig. 6. Lysozyme structural analysis**. **A.** Electron density map of the Lysozyme double hit
720 structure showing no radiation damage is present surrounding the disulfide bond between Cys94
721 and Cys76, for the second hit data. The $2F_o$-F_c map at 1 sigma is shown in blue and overlaid with
722 the Fourier difference (F_o-F_c) map at 3 sigma (green). **B.** CC* for data separated into single hits
723 (blue), first hits (red) and second hits (yellow) as well a correlation of merged data of the second
724 hit relative to the first hit (green) and single hit (purple) data sets indicating they are very similar.

725

726

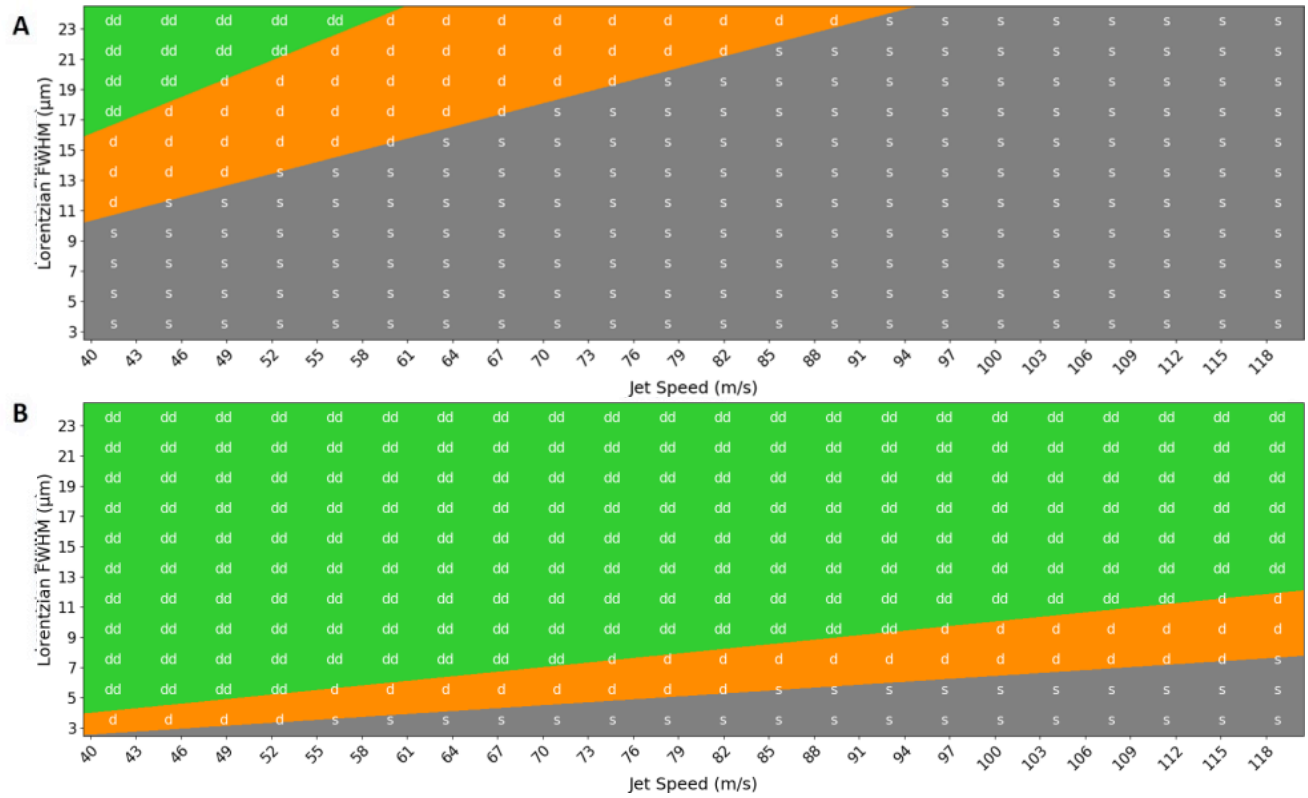


Fig. 7. Parameters for optimizing the collection of double-hit data at the EuXFEL. **A.** EuXFEL repetition rate of 1.1 MHz during this experiment and **B**. A Future capability of EuXFEL where a 4.4 MHz repetition rate is possible. The green shaded area (dd) indicates parameter combinations that will result in double-hits that allow the second hit to occur within the FWHM of the beam, the orange shaded area (d) indicates parameter combinations that will result in double-hits that allow a second hit to occur within in the tail region of the beam, and the grey shaded area (s) indicates parameter combinations that will result in only single hits. This was modeled independent of crystal size (i.e. crystal center-to-crystal center hits).

Table 1. Summary of jet speeds, experimental conditions and statistics for lysozyme crystals. Based on Wiedorn et al. (2018)[1], this table shows the flow rates of the gas and liquid used as well the experimentally and theoretically determined jet speeds for the current analysis. The data statistics were calculated utilising the *CrystFEL* software suite[34,35]. As can be seen below, as the jet speed increases the number of double hits decrease.

Target jet speed	**50 m/s**	**75 m/s**	**110 m/s***
Liquid Flow (µl/min)	15	13	13
Gas Flow (mg/min)	23	50	85
Experimental jet speed (m/s)	42 ± 2.1	78 ± 3.9	-
Theoretical jet speed (m/s)	-	-	102 ± 5.1
Total no. Images	440,000	60,000	240,000
No. hits	10,726	1,638	3,733
	(2.4%)	(2.7%)	(1.6%)
No. Indexed Frames	9,970	1,509	3,474
	(93%)	(92.1%)	(93.1%)
No. double hits	595	14	10
	(6.3% crystals)	(0.9% crystals)	(0.3% crystals)
No. single hits	8,780	1,481	3,454
	(93.7% crystals)	(99.1% crystals)	(99.7% crystals)

* A '100 m/s' jet speed was determined experimentally for a 13 µl/min liquid and 80 mg/min gas flow rate giving a speed of 105 m/s, which is similar to the theoretically calculated 102 m/s jet speed.

Table 2: SFX and refinement statistics for lysozyme single hit crystal data and lysozyme second hit crystal data.

Parameter	Single hit	Second hit
Photon energy (mean value)	9232 eV	9232 eV
X-ray focus		
Pulse energy at sample (assuming 50% beamline transmission)	290 µJ	290 µJ
Pulse length	50 fs	50 fs
Space group	$P4_32_12$	$P4_32_12$
Unit cell		
*a, b, c (*Å)	79.6, 79.6, 37.73	79.6, 79.6, 37.73
α, β, γ (°)	90, 90, 90	90, 90, 90
No. of hits/indexed lattices	10,106 / 10,725	962 / 1,073
No. of unique reflections	7,418 (535)	7,263 (494)
$CC_{1/2}$	0.906 (0.796)	0.615 (0.413)
CC*	0.975 (0.942)	0.873 (0.764)
Data Refinement		
Resolution range (Å)	21.66-2.10 (2.155-2.10)	21.66-2.10 (2.155-2.10)
Completeness (%)	99.84% (100%)	97.75% (92.34%)
Redundancy	47.24 (28.34)	6.66 (3.98)
$I/\sigma(I)$	5.1 (4.2)	2.9 (3.1)
R_{work}	0.152 (0.114)	0.249 (0.244)
R_{free}	0.216 (0.183)	0.299 (0.355)
Rmsd bonds (Å) /Rmsd angles (°)	0.0007 / 1.453	0.004 /1.254
Ramachandran (%)		
favored	98.43%	96.85%
allowed	1.57%	3.15%
outliers	0.00%	0.00%
Wilson B-factor (A^2)		
Average	19.58	20.4
Protein	20.61	21.44
Ligands	37.73	37.24
Ions	22.99	40.32
Waters	29.96	28.01
PDB code		
CXIDB data deposition		

Note: Statistics for the highest-resolution shell are shown in parentheses

772
773
774
775
776
777
778
779
780
781
782
783
784