SCIENTIFIC REPORTS

Double-flow focused liquid injector for efficient serial femtosecond crystallography

Dominik Oberthuer[1], Juraj Knoška[1,2], Max O. Wiedorn[1,2], Kenneth R. Beyerlein[1], David A. Bushnell[3], Elena G. Kovaleva[4], Michael Heymann[1,†], Lars Gumprecht[1], Richard A. Kirian[5], Anton Barty[1], Valerio Mariani[1], Aleksandra Tolstikova[1,2], Luigi Adriano[6], Salah Awel[1,7], Miriam Barthelmess[1], Katerina Dörner[1,‡], P. Lourdu Xavier[1,2,8], Oleksandr Yefanov[1], Daniel R. James[5,$], Garrett Nelson[5], Dingjie Wang[5], George Calvey[9], Yujie Chen[9], Andrea Schmidt[10], Michael Szczepek[10], Stefan Frielingsdorf[11], Oliver Lenz[11], Edward Snell[12], Philip J. Robinson[4], Božidar Šarler[13,14], Grega Belšak[14], Marjan Maček[14], Fabian Wilde[15], Andrew Aquila[16], Sébastien Boutet[16], Mengning Liang[16], Mark S. Hunter[16], Patrick Scheerer[10], John D. Lipscomb[17], Uwe Weierstall[5], Roger D. Kornberg[3], John C. H. Spence[5], Lois Pollack[9], Henry N. Chapman[1,2,17] & Saša Bajt[6]

Serial femtosecond crystallography requires reliable and efficient delivery of fresh crystals across the beam of an X-ray free-electron laser over the course of an experiment. We introduce a double-flow focusing nozzle to meet this challenge, with significantly reduced sample consumption, while improving jet stability over previous generations of nozzles. We demonstrate its use to determine the first room-temperature structure of RNA polymerase II at high resolution, revealing new structural details. Moreover, the double flow-focusing nozzles were successfully tested with three other protein samples and the first room temperature structure of an extradiol ring-cleaving dioxygenase was solved by utilizing the improved operation and characteristics of these devices.

Using pulses from X-ray free-electron lasers (XFEL) with sufficiently short duration to outrun radiation damage, the structures of biological macromolecules can be obtained from nanocrystals at room temperature[1,2]. After interaction with the pulse, the sample is ultimately vaporized, requiring the recording of many diffraction snapshots, each from a fresh crystal, in order to acquire a complete dataset. The resulting method is referred to as serial femtosecond crystallography (SFX)[2]. New structures[3,4] have been obtained this way, recently even by employing

[1]Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany. [2]Department of Physics, University of Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany. [3]Department of Structural Biology, Stanford University, School of Medicine, Stanford, California 94305, USA. [4]SSRL, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. [5]Department of Physics, Arizona State University, Tempe, AZ, USA. [6]Photon Science, Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany. [7]Centre for Ultrafast Imaging, Luruper Chaussee 149, 22761 Hamburg, Germany. [8]IMPRS-UFAST, Max-Planck Institute for the Structure and Dynamics of Matter, 22675 Hamburg, Germany. [9]School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853, USA. [10]Institut für Medizinische Physik und Biophysik, Group Protein X-ray Crystallography and Signal Transduction, Charité - Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin, Germany. [11]Institut für Chemie, Sekr. PC14, Technische Universität Berlin, Straße des 17, Juni 135, 10623 Berlin, Germany. [12]Hauptman-Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, New York 14203, USA. [13]Laboratory for Multiphase Processes, University of Nova Gorica, Vipavska 13, SI-5000 Nova Gorica, Slovenia. [14]Laboratory for Simulation of Materials and Processes, Institute of Metals and Technology, Lepi pot 11, SI-1000 Ljubljana, Slovenia. [15]Helmholtz-Zentrum Geesthacht, Max-Planck-Straße 1, 21502 Geesthacht, Germany. [16]LCLS, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. [17]Department of Biochemistry, Molecular Biology & Biophysics, University of Minnesota, Minneapolis, Minnesota 55455, USA. †Present address: Max-Planck-Institute for Biochemistry, Munich, Germany. ‡Present address: European XFEL GmbH, Hamburg, Germany. $Present address: Paul Scherrer Institut, Villingen, Switzerland. Correspondence and requests for materials should be addressed to S.B. (email: sasa.bajt@desy.de)
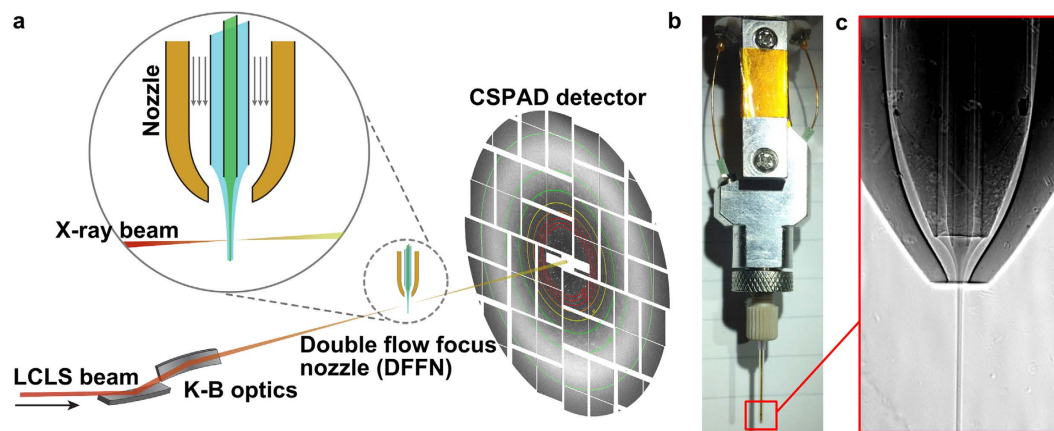
Figure 1. (**a**) Diagram of a SFX experiment at LCLS using a double flow-focusing nozzle (DFFN). Crystals of RNA polymerase II were injected in their crystallization buffer (inner green stream) into the sheath jet formed by ethanol (middle blue stream) through a 40 μm ID capillary. The sheath liquid is accelerated by Helium gas (outer blank channels). A detailed view of the DFFN used in this experiment is shown in (**b**) and in (**c**) a radiograph of a working DFFN, in which the inner jet (4 M KI) can be clearly seen within the sheath jet (ethanol).

experimental phasing[4]. The approach is well suited for time-resolved crystallography over timescales from femtoseconds to seconds or longer[5–7], allowing even the investigation of irreversible reactions since each crystal is used only once[7].

Many SFX experiments to date use a gas-dynamic virtual nozzle (GDVN)[8] to deliver crystals into the X-ray focus. Such jets can operate in an air or vacuum environment to achieve low diffraction background. Current GDVNs require sample flow rates up to 30–40 μl/min to produce a stable jet of protein crystal suspensions at concentrations in excess of 10^{11} crystals/ml. However, many biological macromolecules of high importance cannot be produced in such large quantities. Using a viscous matrix such as lipidic cubic phase (LCP), sample consumption can be reduced to about 100 nl/min or less[9]. But, the slow flow rates and properties of the viscous jets are not always suitable for many time-resolved light activated measurements due to pre-illumination in a slow jet or for fast mixing experiments due to long diffusion times. The slow flow also impairs their use at high repetition rate X-ray sources such as the European XFEL or LCLS-II. Additionally, the 40–80 μm diameter of the LCP jet leads to higher levels of X-ray background scattering from the carrier medium as compared to typical GDVN jets of around 2–6 μm diameter. Low sample consumption can also be achieved with nanoflow electrospinning injection using a microfluidic electro-kinetic sample holder (MESH), which requires the sample to be mixed into a cryo-protective buffer[10]. An improvement of this method using concentric flow (coMESH)[11], overcomes this limitation although this injector design is optimized for slow sample flow, rendering mixing experiments at fast time scales difficult.

In the double-flow focusing nozzle (DFFN) that we present here, an inner liquid jet is focused by a coaxial faster outer liquid jet that is itself focused by gas[12,13] as in the traditional GDVN (Fig. 1). This is achieved by injecting the sample stream into the meniscus of the outer sheath liquid where it is being formed into a jet by the gas constriction[12]. In this region, the flow streams are laminar, and the two liquids (sample and sheath liquids) do not have time to fully mix as the meniscus is sped up by the gas flow and drawn out into a fine 2–5 μm jet with the concentrated protein solution at its core (Suppl. Fig. 1). As the jet is primarily formed by the sheath liquid, the sample consumption can be much less than for a GDVN.

Results and Discussion

Our DFFN system (Fig. 1) reduces sample consumption up to eight fold compared to a GDVN, can inject crystals in their native crystallization buffers, and is suitable for time-resolved mixing experiments and future high repetition rate X-ray sources such as LCLS-II. Furthermore, low surface tension and low viscosity sheath liquids such as ethanol or n-decane reliably jet at reduced total liquid flow rates and jet diameters, hence further reducing diffraction background (Fig. 2 and Suppl. Figs 2 and 3). Using our DFFN system at LCLS we obtained the first room-temperature structure of RNA polymerase II (RNA pol II, Fig. 3a and Suppl. Table 1). A full dataset, based on 8,854 indexed diffraction patterns (Suppl. Fig. 4), was collected during 160 min of beam time, injecting sub-micron sized crystals (approx. 800 nm width) at an average sample flow-rate of 5 μl/min, consuming 10.8 mg of protein. The SFX structure, refined to 3.8 Å resolution, revealed a number of additional residues previously not observed in the cryo-cooled 12 subunit RNA polymerase[14] structure (PDB-entry 1WCM), thus providing more complete protein structural information. The majority of these residues are in RPB2 subunit (amino acid residues 70–76, 335–345, 466–477, 717–722), with others in RPB1 (1174–1187, 1243–1254), the RPB4 N-terminus, RPB6 (69–71) and RPB8 (65–67).

Because of the low X-ray background and focusing of the sample, DFFN injection is also well suited for diffraction data collection of very small crystals. *Cydia pomonella* granulovirus (CpGV) crystals are the smallest crystals used so far for structure determination by SFX[15], with an average size of $400 \times 200 \times 200$ nm^3 (and about 10,000 unit cells per crystal). The crystal structure of CpGV was determined to 2.56 Å resolution (Fig. 3b) and the
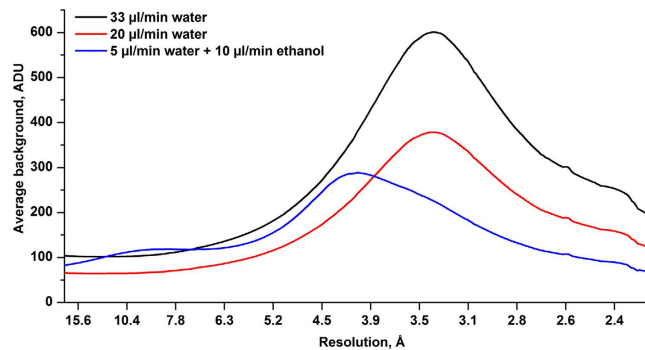
Figure 2. Plots of the diffraction background contribution (in detector ADUs) versus resolution length d (Å) for three jet flow rates (black = 33 μl/min water, red = 20 μl/min water, blue = 5 μl/min water plus 10 μl/min ethanol) obtained by averaging the recorded diffraction over annuli of constant scattering angles after first correcting for the linear polarization of the X-ray beam and scaling by pulse energy. DFF injection using an ethanol sheath yields significantly reduced background in comparison to typical water based jets, especially at resolution lengths of 3.5 to 2.2 Å.
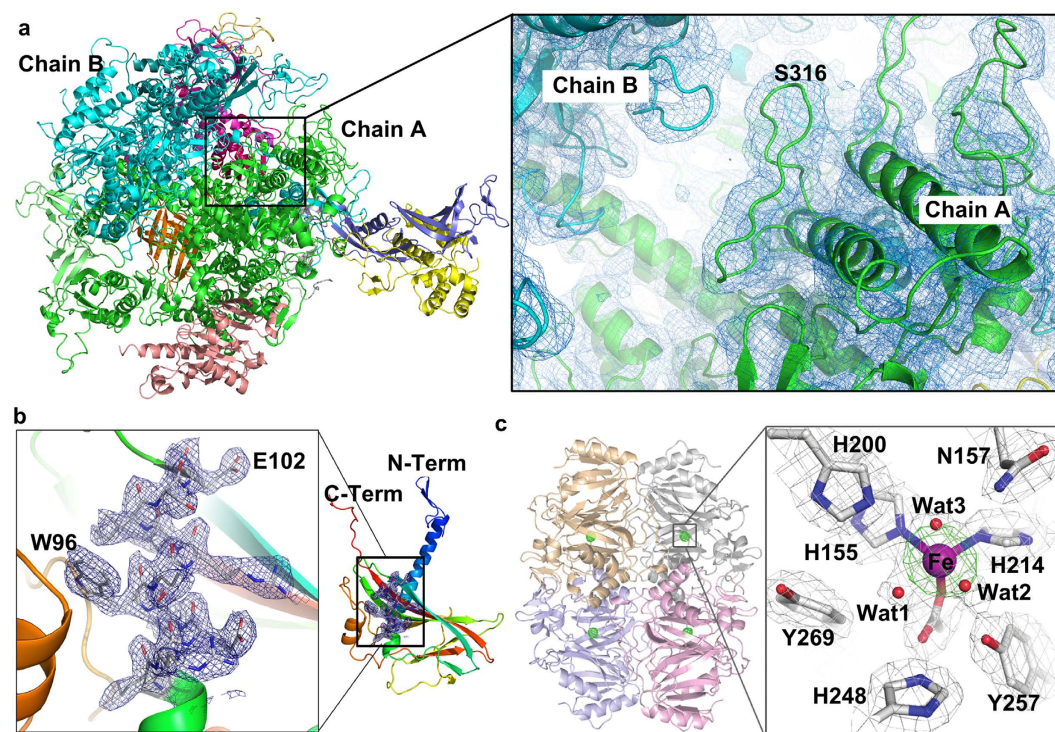

Figure 3. (**a**) Overall room-temperature structure of RNA polymerase II (PDV entry 5U5Q), shown as a cartoon plot. The close-up view shows parts of the interface between chain A and chain B with overlayed $2F_{obs} - F_{calc}$ electron density map (contoured at 1.5 σ). (**b**) The asymmetric unit of CpGV (PDB entry 5 MND), with the blue $2F_{obs} - F_{calc}$ electron density map (contoured at 1 σ) is displayed over an alpha-helical part of the molecule ranging from L92 to E102 and shown as well in the close-up view. (**c**) Crystal structure of HPCD homotetramer (PDB entry 5TRX), and a close-up view of coordination sphere of the catalytic Fe atom in the resting state and key active site residues. The grey $2F_{obs} - F_{calc}$ electron density map is contoured at 1.0 σ. The green ligand omit $F_{obs} - F_{calc}$ difference map, contoured at 7.0 σ, was calculated with Fe atoms removed from the model. Atom color code: grey, carbon; blue, nitrogen; red, oxygen; purple, iron. Cartoons depict secondary structure elements for the 4 subunits.

statistics (Suppl. Table 1) show that this was limited by the detector acceptance angle rather than the nanocrystal size. Using CpGV, we further analyzed how the fraction of pulses, which show crystal diffraction (the "hit fraction"), depends on the sample flow rate and the type of sheath liquid, for a given crystal concentration. For CpGV, the flow-rate of 3–5 μl/min produced an inner jet width matched both to the size of the X-ray beam and that of the crystals (Suppl. Fig. 5). The average hit fractions of 8–10% obtained at these low sample flow rates indicates the continuous operation of a stable jet over the course of the experiment and the absence of interruptions that
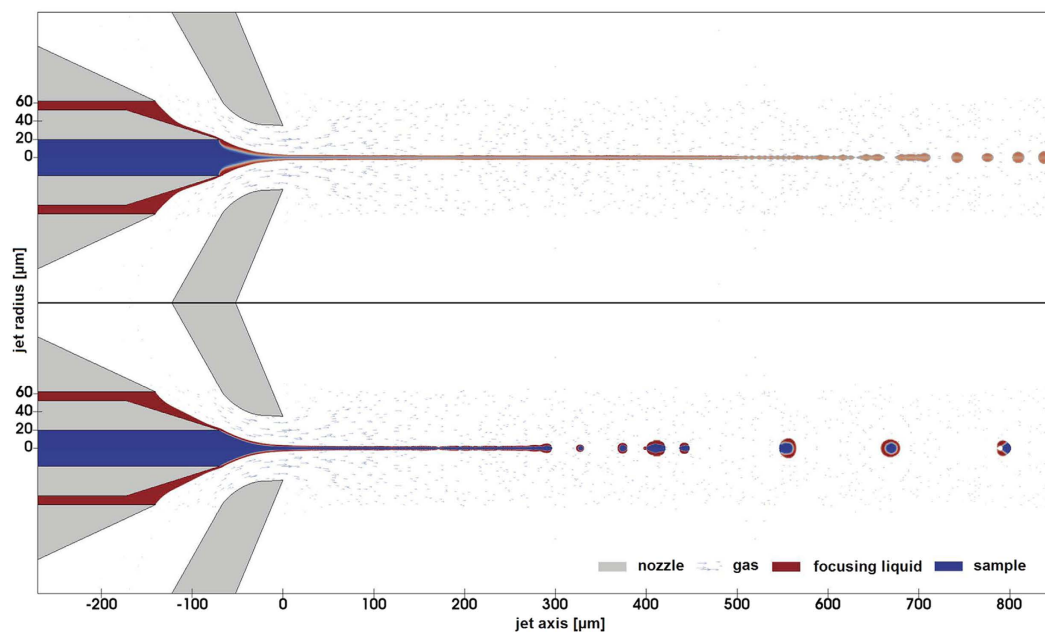
Figure 4. Comparison of a snapshot at the 1 ms time point of two simulations solved with the assumptions of mixing (top) and non-mixing (bottom) of the sample and focusing liquids. Simulations were run in such a way that the focusing liquid - alcohol (red color) jet is created first and is later joined by the liquid of 2.5 M NaCl water solution (blue color), simulating the experimental conditions. The simulation for miscible liquids was obtained with interMixingFoam solver, characterized by a diffusion of 10^{-9} m²/s between the miscible phases. The immiscible case shown at the bottom was calculated with multiphaseInterFoam solver. Other parameters were equal in both cases as described in Methods.

often accompany experiments that use GVDNs. Notably, as shown in Suppl. Fig. 5, the hit fraction is higher when ethanol is used as sheath-liquid as compared to water, for the same sample flow-rate. This can be attributed to the lower surface tension of ethanol which leads to a more stable outer jet at lower flow rates, which increases the probability of the center of the jet (where crystals flow) being in the X-ray focus when an X-ray pulse arrives. Even so, either the use of water or ethanol as sheath liquids leads to stable jets at much reduced sample flow-rates as compared to normal GDVN injection. Additionally, we could independently vary the outer and inner flow-rates during operation without impairing the jetting of the DFFN.

We explored a range of focusing liquids and found that the use of n-decane, which is immiscible with water or aqueous buffers, produced a very unstable hit fraction, whereas ethanol and water proved to be good focusing liquids for aqueous suspensions of protein crystals. To better understand the influence of different parameters on the DFFN jet, we performed computational fluid dynamics simulations. Using the nozzle geometry, two cases were studied, one including diffusion between the buffer and the focusing fluid (two miscible liquids) and the other simulating immiscible fluids. The results (Fig. 4) match the experimental observation that sheath liquids miscible with the buffer are better suited for DFFN injection.

DFFN's are also ideal for jetting samples in buffers that contain high concentrations of polymers (e.g. PEG) and/or salt. With GDVNs operated in vacuum, even low concentrations of PEGs or salts result in an increased risk of ice formation in the X-ray interaction region, which strongly diffracts and can damage the detector. Often only venting of the vacuum chamber removes the ice, reducing the data collection time. The ethanol sheath liquid prevented this ice formation in all tested PEG concentrations (5–15%, w/v). In all tests, the sheath liquid also prevented the formation of PEG and/or salt debris at the nozzle tip, which is very often observed with normal GDVN and can impair data collection in a similar manner to icing.

Previously, injection of the crystal suspension of membrane-bound [NiFe] hydrogenase from *Ralstonia eutropha* (*Re*MBH)[16,17] with a standard GDVN failed due to the high viscosity of the sample, while a modified GDVN with a 100 μm inner diameter clogged after less than four minutes at flow rates of 30 μl/min. Injection of *Re*MBH using DFFN allowed uninterrupted nozzle flow with a flow-rate of 8.5 μl/min over the course of a 90 minute experiment, enabling the collection of 3576 indexable images. Data collection statistics show that a complete dataset could be obtained to a resolution limit of at least 2.2 Å (Suppl. Table 1, Suppl. Fig. 6). Structural refinement is still on-going and will be published in a separate paper focused on the structure and function of *Re*MBH.

To further test sample delivery with DFFN, micro-crystals of an extradiol ring-cleaving dioxygenase from *Brevibacterium fuscum* (HPCD), at a sample density of only 10^8–10^9 crystals/ml, were used to acquire a complete dataset, based on 2,185 indexed patterns, during 49 min of data collection at an average sample flow-rate of 7 μl/min. The SFX structure of the HPCD enzyme in the resting state was refined to 2.38 Å resolution (Fig. 3c, Suppl. Table 1). Comparison of the SFX and cryo-MX (cryo-macromolecular crystallography) structures shows no significant differences in the global protein backbone, but contractions of both macromolecule and

unit cell are observed in the low-temperature structure (Suppl. Fig. 7). These changes are absorbed via slight increases in inter-molecular contact areas (Suppl. Fig. 8). Previous cryo-MX studies have shown that the HPCD homo-tetramer remains catalytically active in crystalline form, and that different reaction intermediates accumulate in the 4 nominally identical subunits within a single asymmetric unit[18]. Results of the current study demonstrate that unique crystal packing environments for the four subunits of the homo-tetramer are present at room temperature, supporting earlier conclusions that the observed differential subunit reactivity is due to differences in crystal packing interactions and dynamic restrictions. This finding eliminates cryo-cooling as the cause of differential intermediate stabilization[18,19].

To conclude, a new injection system for SFX experiments, referred to as a double-flow focus nozzle (DFFN), enables structural biology experiments at room temperature on samples of high biological importance as demonstrated by the first structures of RNA polymerase II and HPCD determined at non-cryogenic temperatures. The DFFN significantly reduces sample consumption, overcomes limitations regarding buffer composition and viscosity and substantially extends uninterrupted run-times, while maintaining the benefits of an overall fast flowing jet. Furthermore, it is possible to transform the DFFN into a mix-and-inject nozzle[13,20] for time-resolved structural enzymology experiments (Suppl. Fig. 9).

Methods

Double Flow-Focusing Nozzles. A fused silica capillary with a square-profile inner bore of $100\,\mu m$ width (Polymicro, USA) was used as the outer liquid line. Into this, a $40\,\mu m$ inner diameter (ID)/$105\,\mu m$ outer diameter (OD) round fused silica capillary (Polymicro, USA) was inserted after slightly etching the inner wall of the outer capillary with hydrofluoric acid to accommodate the $105\,\mu m$ diameter. All fused silica capillaries were sharpened by wet mechanical polishing[21] before assembly. Two Kapton rings (so called 'spacers') were placed around the outer liquid line (as reported before ref. 21) to ensure centering of this line within the outermost nozzle piece. The inner round capillary self-centers in the square bore of the outer capillary and does not need to be kept in place by spacers. This capillary ensemble was built into the DFFN-housing (see Fig. 1 and Suppl. Fig. 10). A glass capillary with $500\,\mu m$ ID was flame-polished at its end to form a gas orifice of $60\,\mu m$. This outer nozzle piece was then fitted over the outer liquid line and sealed with a fitting-ferrule-sleeve assembly (IDEX). The DFFN was designed so that it would connect the outermost glass capillary to the gas line and the outer liquid capillary to the focusing liquid line. The inner (sample) line extends from the nozzle tip to the upper part of the nozzle housing (~15 cm). The relative position of the tip of the inner liquid line to the tip of the outer liquid line was adjusted after assembly while running liquid in the outer line to optimize nozzle performance.

Sheath liquids. Focusing liquids were loaded into a stainless steel reservoir and delivered to the DFFN by applying gas pressure on the reservoir. For all experiments, N_2 was used to pressurize this reservoir. The flow-rate of sheath liquid was adjusted by changing the gas pressure using a Proportion Air GP1 controller. To avoid clogging of the DFFN, a stainless steel in-line HPLC filter (IDEX) with $0.5\,\mu m$ pore size was placed between the sheath-liquid reservoir and the nozzle. Water was purified with the Milli-Q Integral water system (EMD Millipore) available in the LCLS sample preparation laboratory. Pure ethanol (Sigma Aldrich, CHROMASOLV®, SIAL-34852) and n-decane (Sigma Aldrich, SIAL-457116) were used as purchased.

Numerical simulations. To computationally model the GDVN operation we need to consider its geometrical and surface design, the gas and liquid flow rates (Q_g and Q_l) and pressures (P_g and P_l), as well as material properties of the flow: densities ρ_g and ρ_l, viscosities μ_g and μ_l, and surface tension between the liquid and the gas γ_{lg}. The DFFN operation is even more complex, requiring more parameters, since the gas and the sample liquid flow rates and pressures are accompanied by the flow focusing fluid flow rate Q_f and pressure P_f with accompanying material properties ρ_f, and μ_f. The surface tensions between the sample and the focusing liquids γ_{lf}, and the focusing liquid and the gas γ_{fg} have to be considered as well, along with the diffusion between the buffer and the flow focusing liquid D_{lf}. The adjustment of the DFFN is thus a highly complex task that depends, in addition to the geometrical design, on 16 parameters. The flow rates and pressures of the sample liquid, focusing fluid and gas are thus difficult to optimize in a measurement without prior extensive experimentation. A complementary methodology that helps this optimization of the operating parameters is computational fluid dynamics (CFD). In the case of DFFN this is particularly difficult because compressible gas, two miscible or immiscible fluids, and moving interphase boundaries between all three must be simulated. Compounding the difficulty are the large aspect ratio of the jet length to width and the flow morphology. The flow behavior can be reasonably well represented in axial symmetry only for dripping and jetting modes, whereas the whipping mode[21] has a highly three-dimensional character.

The OpenFOAM open source CFD solvers "multiphaseInterFoam" and "interMixingFoam" were used respectively for the non-mixing and mixing cases of the sample and focusing liquids. The code for the mixing case had to be modified in order to properly describe the interface between the focusing fluid and the gas. The geometry was prepared by FreeCAD. To reduce calculation time the non-axial symmetry of the DFFN (round inner capillary inserted into square outer capillary) was treated as axially symmetric (consisting of a round capillary inserted into round outer capillary) with the same cross-sectional area of the channel. The simulations were visualized using ParaView, with additional code to extract the length, thickness, and velocity of the sample and fluid focusing jets, frequency and size of drops, drop separation length and other relevant parameters. A typical calculation requires discretization into several hundred thousand computational cells with a width that decreases towards the jet axis to a finest width of $0.3\,\mu m$. To assess the DFFN jet behavior it is sufficient to simulate the behavior within the first 3 ms. The respective computational time is several days on a modern server with 20 cores. Figure 4 displays the results of calculations of mixing of the sample and focusing liquids and non-mixing, carried out with the following parameters: $Q_g = 10\,mg/min$, $Q_f = 10\,\mu L/min$, $Q_l = 5\,\mu L/min$, $\mu_g = 1.9*10^{-5}\,kg/ms$, $\mu_f = 1.12*10^{-3}\,kg/ms$,

$\mu_l = 17.49 * 10^{-3}$ kg/ms, $\gamma_{lg} = 0.0728$ N/m, $\gamma_{fg} = 0.0223$ N/m, $\gamma_{lf} = 0.0505$ N/m, $\rho_g = 1.13$ kg/m^3, $\rho_f = 789$ kg/m^3 and $\rho_l = 882.4$ kg/m^3. In the case of mixing the diffusion between the miscible phases was set at 10^{-9} m^2/s. Stable jetting is obtained in both cases, but it is seen that the use of miscible liquids produces a longer and thinner jet, which is preferred for SFX experiments, as compared with immiscible liquids.

Operation of the DFFN at CXI.   For experiments at the CXI beamline we designed and built an adapter to mount the DFFN on the rod assembly currently used to hold and position nozzles in the vacuum chamber (see Suppl. Fig. 10). Using this nozzle positioning system, the tip of the DFFN was positioned approximately 100 μm above the X-ray interaction point. To start the jet, first the flow of sheath gas (He) was started with a pre-tested stable value of around 10 mg/min, requiring about 2.07×10^6 Pa. Then, the focusing liquid reservoir was pressurized with N_2 to about 1.38×10^6 Pa to get a stable jet with a flow rate between 10–15 μl/min. Once a stable jet was produced, water was injected into the meniscus through the inner capillary using a high-pressure liquid chromatography (HPLC) pump. A stainless steel in-line filter (IDEX) with a pore size ranging from 5 to 20 μm (depending on the sample crystal size) was used. An initial flow-rate of 10 μl/min was selected for the inner flow. Once the jet was aligned with the X-ray interaction region the sample injection switch box at CXI was used to pressurize a sample reservoir plunger using the HPLC pump. Initially the flow-rate of sample was set to be 5 μl/min, and increased in 1 μL/min steps to 10 μL/min if no diffraction hits were observed. Once a hit fraction of about 5% was achieved with 5 μl/min, the flowrate could be slightly reduced to maximize the hit fraction to sample consumption ratio, if desired. To switch between samples, the inner line (sample line) was purged with water until no more hits could be detected by the online analysis (OnDA)[22], followed by a switch to the next sample. To turn off a DFFN without risk of icing or back-flow of sample into the nozzle, the inner channel was first purged with water and then depressurized, then the focusing liquid was depressurized and the DFFN removed from the CXI vacuum chamber with sheath gas flowing. All lines were dried with gas before removing the DFFN from the nozzle rod.

X-ray radiography of the DFFN.   Phase-contrast radiographic X-ray images of the DFFN were acquired at the imaging beamline P05 (PETRA III/DESY and Helmholtz-Zentrum Geesthacht)[23], experimental hutch (EH2) at 5 keV X-ray energy. Images were acquired using a 3.7 s exposure. In order to enhance image quality a background image, taken using the same exposure but with the injector moved out of the field of view, was subtracted from an average of three subsequent images. For the images shown in Fig. 1c and Suppl. Fig. 1, we used a 4 M aqueous solution of potassium iodide as a sample liquid to provide contrast between the inner and outer streams. Injection was carried out with 5 μl/min sample flow, 20 μl/min outer flow rate (ethanol) and 8.27×10^5 Pa focusing gas (Helium) pressure, which matches typical injection conditions for SFX experiments.

Preparation and crystallization of RNA polymerase II.   12 subunit RNA polymerase II (RNA Pol II) was purified from the yeast *Saccharomyces cerevisiae* as described previously[24]. To make crystals suitable for SFX experiments previously published conditions were adapted to growth of crystals <1 μm^3 (ref. 25). Pol II was concentrated to 24 mg/ml in 5 mM Tris-Cl pH 7.5, 60 mM ammonium sulfate, 100 uM ZnCl$_2$, 10 mM DTT using a Via Spin 500 100 K MWCO spin concentrator. 24 mg/ml pol II was mixed 1:1 with the crystallization buffer (21–22% PEG 400, 5% PEG 5%, 100 mM ammonium phosphate pH 6.3, 100 mM NaCl, 1 mM ZnCl2 and 10 mM DTT) and placed in a 200 μl dialysis button (Hampton Research cat# HR3-330). The dialysis button was sealed using 7000 MWCO SnakeSkin Dialysis tubing (Thermo-Fisher cat#68700), placed in a 50 ml Falcon tube with 20 ml of crystallization buffer, and then in a 16.5 °C incubator. Crystals were harvested after 1 week in the 16.5 °C incubator and stored at 4 °C. Crystal size and concentration were measured using a NanoSight (Malvern). The sample used for the experiments had a size distribution between 500—1000 nm with a mean size of ~800 nm and a concentration of 10^{11} particles/ml. Prior to injection the Pol II crystals were filtered through a 10 μm PEEK filter.

Preparation of CpGV.   *Cydia pomonella* granulovirus (CpGV)[26] is used as an insecticide against the codling moth (*Cydia pomonella*) and is commercially available in various formulations (e.g. Madex Max, Madex Plus, both Andermatt Biocontrol or CYD-X, Certisusa). Here we used Madex HP (Certisusa). Madex HP contains about 10^{13} virus occlusion bodies (OB) per liter and the OBs were purified from the aqueous suspension by applying iterative washing and centrifugation cycles. The pellet was then re-suspended in ultra-pure water at pH7. After 3 h of incubation at room temperature the supernatant, containing the almost pure OBs, was removed from the pellet and subjected to filtration steps through a sequence of stainless steel filters with decreasing pore sizes (20 μm, 10 μm, 5 μm, 2 μm, 0.5 μm; all IDEX). To increase the concentration of CpGV to the desired 10^{11} particles/ml for injection at LCLS, the suspension was subjected to centrifugation at 21,000 g, the supernatant was removed, and the pellet re-suspended. Size distribution and particle concentration was estimated using a Nanosight LM14 instrument (Malvern).

Preparation and crystallization of membrane-bound [NiFe]-hydrogenase from *Ralstonia eutropha* (*Re*MBH).   Native *Re*MBH was expressed and purified in the as-isolated (air-oxidised) state as described elsewhere[16,17,27]. *Re*MBH nano- and microcrystals were grown at 4°–6 °C and pH 5.5–6.5 in the presence of 20–30% PEG3350 and Bis-(2-hydroxy-ethyl)-amino-tris(hydroxylmethyl)-methane following the micro-crystallization setup-protocol as described for PYP[28]. *Re*MBH solution (approx. 10 mg/ml) was mixed with the crystallization buffer in a 1:1 ratio. The mixture was vigorously stirred for 5 minutes at 4 °C and then incubated for 2–3 days to allow crystal growth to a well-defined size of 5–10 μm and a concentration of 10^{11} small crystals/ml. Prior to injection, crystals were subsequently filtered through 20 μm stainless steel filters.

Preparation and crystallization of HPCD dioxygenase. Recombinant homoprotocatechuate 2,3-dioxygenase from *B. fuscum* (HPCD) was expressed in *E. coli* and purified as described previously[29,30]. For injector-based studies, crystallization conditions and procedures that were optimized to produce diffraction quality macro-crystals of 2,3-HPCD enzymes[18,31] were used as the starting point to optimize and scale up production of micro-crystal slurries using batch crystallization. Protein solutions (5–7 mg/ml) were gently mixed in a 1:1 ratio with crystallization solution consisting of 12–14% PEG6000, 0.1–0.14 M calcium chloride, and 0.1 M MES pH 5.8, ensuring that no visible precipitate forms at these concentrations of components (protein, polymer and salt). After a few minutes of equilibration, a 20 μl aliquot of the 1:1 stock mixture (protein and crystallization solution) was seeded with roughly crushed macro-crystals in an Eppendorf tube, and the growth of many needle-like crystals was observed after a few hours of incubation at 20 °C. The growing seed stocks were diluted with a fresh mixture of protein and crystallization solution every 2–4 hours to maintain maximal growth of needle-like crystals and to reach a total of 2–3 ml sample volume. Following a final 12 hour period of growth to achieve saturation, slurries of needle-like crystals were concentrated by centrifugation at 2,000 rpm for ~5 minutes. Partial fragmentation was achieved by adding layers of 0.1 mm and 0.5 mm glass beads on top of the soft crystalline pellet, and additional centrifugation for 2 minutes. The samples were re-suspended by gentle pipetting, separated from the glass beads and passed through a 20 μm stainless steel filter to remove particulates and larger crystalline material prior to loading into the sample reservoir for diffraction studies. Sample density was estimated to be ~10^8–10^9 crystals/ml by bright field microscopy with average crystal needle dimensions of $(1–2 \mu m) \times (1–2 \mu m) \times (10–30 \mu m)$.

Data collection. SFX experiments with the DFFN were carried out at the Linac Coherent Light Source[32] (LCLS) experimental station CXI[33] at the SLAC National Accelerator Laboratory (Menlo Park, CA, USA) during beam times LH96 (Dioxygenase, Pol II and *Cp*GV) and LG68 (*Re*MBH). The LCLS X-ray beam was focused to 90×150 nm^2 FWHM (V × H) with Kirkpatrick-Baez mirrors. A photon energy of 8.0 keV, pulse length of 50 fs and a repetition rate of 120 Hz was used throughout the experiments. The Cornell-SLAC pixel-array detector[34] (CSPAD) was used for data collection. To improve the quality of data analysis, the geometry of the CSPAD was refined using 'geoptimiser'[35] during analysis of the CpGV data sets. The improved geometry was then used for the other data sets collected. The online monitor OnDA[22] was used during the experiments to assess hit fraction and data quality 'on the fly'. This rapid feedback was necessary to adjust the data collection strategy during the experiment for optimization of both hit fraction and sample consumption.

Data analysis and processing. Individual 'hits' were identified from the complete set of collected diffraction patterns and converted to HDF5 format using the software Cheetah[36]. Indexing, integration and data reduction was carried out using the program 'indexamajig' in CrystFEL[37] (Version 0.6.1). The resulting stream-files were subjected to post-refinement (scaling) and merged using 'partialator'. In the case of CpGV the resulting stream-files were detwinned[38] using 'ambigator' as implemented in CrystFEL to resolve the indexing ambiguity, before post-refinement and merging. MTZ-files for crystallographic data-processing were generated from CrystFEL-hkl-files using f2mtz (CCP4[39]). Figures of merit were calculated using 'compare_hkl' (R_{split}, $CC_{1/2}$, CC*) and 'check_hkl' (SNR, multiplicity, completeness), both from CrystFEL.

Hit fraction and background analysis. In order to analyze how the hit fraction depends on the sample flow-rate and the type of sheath liquid the CpGV diffraction data was collected at several flow-rates with both water and ethanol as sheath liquid for at least 5 minutes at each point, chosen as a balance between obtaining sufficient data statistics for the hit fraction estimation at a single operating condition and a characterization of the nozzle performance over a range of conditions, given the limited beam time. Changing flow rates on the HPLC can require up to a minute to equilibrate while changes of the gas pressure (driving force for jet velocity) are almost instantaneous. The hit fraction was monitored for 5 minutes after equilibrium was reached. While online monitoring at the beamline was carried out using OnDA, the hit fraction analysis was calculated using the output of Cheetah. The sample flow-rate was measured with a flow meter and then averaged over the run. The sheath liquid flow rate was determined by the set value of the HPLC. Plots of the hit fraction are shown in Suppl. Fig. 5.

For the background analysis three runs were selected: one with the optimal DFFN injection conditions with ethanol as a sheath liquid and two with the flow-rates typical for the water-based jets. 20000 non-hit images from each run were averaged after detector corrections. Then, after applying the correct detector geometry, refined with the 'partialator' program in CrystFEL and masking bad pixels and shadowed regions on the detector the radially averaged plots were obtained, shown in Fig. 2.

Structural refinement. *RNA polymerase II.* Refinement of the room temperature 12-Subunit RNA Polymerase II from *Saccharomyces cerevisiae* was performed using PHENIX[40] with data spanning scattering angles corresponding to resolution lengths of 40–3.8 Å. Due to small differences in the unit cell dimensions between the room temperature data and previous cryo-cooled structure[25,41] the initial starting model (PDB entry 1WCM) was placed in the unit cell using rigid body refinement followed by simulated annealing, XYZ coordinates and individual *B*-factor refinements with optimized X-Ray/ADP (atomic displacement parameter) weights and secondary structure restraints. The resulting 2Fo-Fc electron density map showed additional density that was built using COOT[42]. Final rounds of refinement were performed with the inclusion of TLS parameters (for translation, libration (small movements) and screw-rotation of a group of atoms). TLS groups were chosen based on analysis using the TLSMD web server[43]. For RPB1 and RPB2, 20 TLS groups were defined; for RPB3-11, 10 TLS groups were defined; and for RPB12, 7 TLS groups were defined. The final refined structure had error metrics $R_{work} = 0.22$, $R_{free} = 0.27$, root-mean-square deviation (RMSD) for angles equal to 0.62°, RMSD for bonds equal to 0.003 Å and a final clash-score of 13.28. The Ramachandran statistics for energetically allowed regions for

backbone dihedral angles ψ against φ of amino acid residues were 84.6% favored, 11.8% allowed and 3.6% outliers. Notably, the geometry statistics of the final room temperature 12-subunit RNA Polymerase structure were better than for the previous cryo-cooled structure.

CpGV. Rigid-body refinement of the native CpGV structure[15] (PDB-ID: 5G0Z) against the new data using phenix.refine[44] did not give a solution. Molecular replacement phasing using PhaserMR[45] in Phenix[40] found a single solution (LLG = 3571, TFZ = 46.9). The combination of automatic maximum-likelihood refinement with phenix.refine (xyz, real space and isotropic ADP refinement) and manual model building in Coot resulted in a CpGV structure refined to 2.56 Å resolution. The final refined structure had R_{work} = 0.154, R_{free} = 0.207, RMSD (angles) equal to 0.54°, RMSD (bonds) equal to 0.003 Å and a final Clash-score of 2.50. The Ramachandran statistics were 97.9% favored, 2.1% allowed, and no outliers.

HPCD dioxygenase. The coordinates of the full length 2,3-HPCD (PDB 3OJT) were used as an initial model in rigid body refinement followed by cycles of restrained refinement with Refmac5[46] as part of the CCP4 program suite[39] and model building using Coot[42]. TLS was used in the final round of restrained refinement, with a single subunit defined as a TLS group. Link restraints to the iron were removed from the refinement to avoid bias in the refined metal-ligand distances. NCS restraints were not used during refinement, and the 4 subunits of the single enzyme molecule present in the asymmetric unit were refined independently. The final refined parameters for the SFX HPCD structure (PDB 5TRX) resulted in R_{work} = 0.18, R_{free} = 0.23, RMSD (angles) equal to 1.358°, RMSD (bonds) equal to 0.011 Å and a Clash-score of 2.

Figures. Figures showing diffraction images were generated using either Cheetah or hdfsee as implemented in CrystFEL. PyMol was used to generate figures showing protein structure models and electron density maps.

References

1. Boutet, S. *et al.* High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* **337,** 362–364, doi: 10.1126/science.1217737 (2012).
2. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470,** 73–77, doi: 10.1038/nature09750 (2011).
3. Redecke, L. *et al.* Natively Inhibited Trypanosoma brucei Cathepsin B Structure Determined by Using an X-ray Laser. *Science* **339,** 227–230, doi: 10.1126/science.1229663 (2013).
4. Colletier, J. P. *et al.* De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. *Nature*, doi: 10.1038/nature19825 (2016).
5. Barends, T. R. *et al.* Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350,** 445–450, doi: 10.1126/science.aac5492 (2015).
6. Pande, K. *et al.* Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science* **352,** 725–729, doi: 10.1126/science.aad5081 (2016).
7. Stagno, J. R. *et al.* Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature*, doi: 10.1038/nature20599 (2016).
8. DePonte, D. P. *et al.* Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics* **41,** 195505, doi: 10.1088/0022-3727/41/19/195505 (2008).
9. Weierstall, U. *et al.* Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature communications* **5,** 3309, doi: 10.1038/ncomms4309 (2014).
10. Sierra, R. G. *et al.* Nanoflow electrospinning serial femtosecond crystallography. *Acta crystallographica. Section D, Biological crystallography* **68,** 1584–1587, doi: 10.1107/S0907444912038152 (2012).
11. Sierra, R. G. *et al.* Concentric-flow electrokinetic injector enables serial crystallography of ribosome and photosystem II. *Nature methods* **13,** 59–62, doi: 10.1038/nmeth.3667 (2016).
12. Gañán-Calvo, A. M., González-Prieto, R., Riesco-Chueca, P., Herrada, M. A. & Flores-Mosquera, M. Focusing capillary jets close to the continuum limit. *Nature Physics* **3,** 737–742, doi: 10.1038/nphys710 (2007).
13. Wang, D., Weierstall, U., Pollack, L. & Spence, J. Double-focusing mixing jet for XFEL study of chemical kinetics. *Journal of synchrotron radiation* **21,** 1364–1366, doi: 10.1107/S160057751401858X (2014).
14. Armache, K. J., Mitterweger, S., Meinhart, A. & Cramer, P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *The Journal of biological chemistry* **280,** 7131–7134, doi: 10.1074/jbc.M413038200 (2005).
15. Gati, C. *et al.* Atomic structure of granulin determined from native nanocrystalline granulovirus using an X-ray free-electron laser. *Proceedings of the National Academy of Sciences of the United States of America*, doi: 10.1073/pnas.1609243114 (2017).
16. Fritsch, J. *et al.* The crystal structure of an oxygen-tolerant hydrogenase uncovers a novel iron-sulphur centre. *Nature* **479,** 249–252, doi: 10.1038/nature10505 (2011).
17. Kalms, J. *et al.* Krypton Derivatization of an O2 -Tolerant Membrane-Bound [NiFe] Hydrogenase Reveals a Hydrophobic Tunnel Network for Gas Transport. *Angewandte Chemie* **55,** 5586–5590, doi: 10.1002/anie.201508976 (2016).
18. Kovaleva, E. G. & Lipscomb, J. D. Crystal structures of Fe2+ dioxygenase superoxo, alkylperoxo, and bound product intermediates. *Science* **316,** 453–457, doi: 10.1126/science.1134697 (2007).
19. Kovaleva, E. G. & Lipscomb, J. D. Intermediate in the O-O bond cleavage reaction of an extradiol dioxygenase. *Biochemistry* **47,** 11168–11170, doi: 10.1021/bi801459q (2008).
20. Calvey, G. D., Katz, A. M., Schaffer, C. B. & Pollack, L. Mixing injector enables time-resolved crystallography with high hit rate at X-ray free electron lasers. *Structural Dynamics* **3,** 054301, doi: 10.1063/1.4961971 (2016).
21. Beyerlein, K. R. *et al.* Ceramic micro-injection molded nozzles for serial femtosecond crystallography sample delivery. *The Review of scientific instruments* **86,** 125104, doi: 10.1063/1.4936843 (2015).
22. Mariani, V. *et al.* OnDA: online data analysis and feedback for serial X-ray imaging. *J Appl Crystallogr* 49, 1073–1080, doi: 10.1107/S1600576716007469 (2016).
23. Wilde, F. *et al.* Micro-CT at the imaging beamline P05 at PETRA III. **1741,** 030035, doi: 10.1063/1.4952858 (2016).
24. Robinson, P. J. *et al.* Molecular architecture of the yeast Mediator complex. *eLife* **4,** doi: 10.7554/eLife.08719 (2015).
25. Bushnell, D. A. & Kornberg, R. D. Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 6969–6973, doi: 10.1073/pnas.1130601100 (2003).
26. Jehle, J. A. *et al.* Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. *Virology* **346,** 180–193, doi: 10.1016/j.virol.2005.10.032 (2006).

27. Frielingsdorf, S. *et al.* Reversible [4Fe-3S] cluster morphing in an O(2)-tolerant [NiFe] hydrogenase. *Nature chemical biology* **10,** 378–385, doi: 10.1038/nchembio.1500 (2014).

28. Tenboer, J. *et al.* Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* **346,** 1242–1246, doi: 10.1126/science.1259357 (2014).

29. Groce, S. L. & Lipscomb, J. D. Aromatic ring cleavage by homoprotocatechuate 2,3-dioxygenase: role of His200 in the kinetics of interconversion of reaction cycle intermediates. *Biochemistry* **44,** 7175–7188, doi: 10.1021/bi050180v (2005).

30. Mbughuni, M. M., Meier, K. K., Munck, E. & Lipscomb, J. D. Substrate-mediated oxygen activation by homoprotocatechuate 2,3-dioxygenase: intermediates formed by a tyrosine 257 variant. *Biochemistry* **51,** 8743–8754, doi: 10.1021/bi301114x (2012).

31. Kovaleva, E. G. & Lipscomb, J. D. Structural basis for the role of tyrosine 257 of homoprotocatechuate 2,3-dioxygenase in substrate and oxygen activation. *Biochemistry* **51,** 8755–8763, doi: 10.1021/bi301115c (2012).

32. Bostedt, C. *et al.* Linac Coherent Light Source: The first five years. *Rev Mod Phys* **88,** doi: 10.1103/RevModPhys.88.015007 (2016).

33. Liang, M. *et al.* The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *Journal of synchrotron radiation* **22,** doi: 10.1107/s160057751500449x (2015).

34. Blaj, G. *et al.* X-ray detectors at the Linac Coherent Light Source. *Journal of synchrotron radiation* **22,** 577–583, doi: 10.1107/S1600577515005317 (2015).

35. Yefanov, O. *et al.* Accurate determination of segmented X-ray detector geometry. *Opt Express* **23,** 28459–28470, doi: 10.1364/OE.23.028459 (2015).

36. 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, doi: 10.1107/S1600576714007626 (2014).

37. White, T. A. *et al.* Recent developments in CrystFEL. *J Appl Crystallogr* **49,** 680–689, doi: 10.1107/S1600576716004751 (2016).

38. Brehm, W. & Diederichs, K. Breaking the indexing ambiguity in serial crystallography. *Acta crystallographica. Section D, Biological crystallography* **70,** 101–109, doi: 10.1107/S1399004713025431 (2014).

39. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D* **67,** 235–242, doi: 10.1107/S0907444910045749 (2011).

40. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66,** 213–221, doi: 10.1107/S0907444909052925 (2010).

41. Armache, K. J., Kettenberger, H. & Cramer, P. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc Natl Acad Sci USA* **100,** 6964–6968, doi: 10.1073/pnas.1030608100 (2003).

42. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **66,** 486–501, doi: 10.1107/S0907444910007493 (2010).

43. Painter, J. & Merritt, E. A. TLSMD web server for the generation of multi-group TLS models. *Journal of Applied Crystallography* **39,** 109–111, doi: 10.1107/S0021889805038987 (2006).

44. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **68,** 352–367, doi: 10.1107/S0907444912001308 (2012).

45. McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography* **40,** 658–674, doi: 10.1107/S0021889807021206 (2007).

46. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography* **67,** 355–367, doi: 10.1107/S0907444911001314 (2011).

Acknowledgements

Author Contributions

D.O., H.N.C. and S.B. designed, coordinated and organized the experiment. J.K. and M.O.W. produced and tested the initial DFFN prototypes. K.R.B., D.O. and L.G. designed and developed the nozzles with J.K., M.O.W. and M.H. K.R.B., M.H., J.K., M.O.W., L.G. and L.A.: fabricated the nozzles and further improved the design. Nozzles were tested and analyzed by M.O.W., J.K. and L.A. LCLS data collection was executed by D.O., J.K., M.O.W., K.R.B., D.B., E.G.K., M.H., L.G., R.A.K., A.B., V.M., A.T., S.A., M.B., K.D., P.L.X., O.Y., D.J., G.N., D.W., G.C., Y.C., M.S., A.A., S.Bou., M.L., M.S.H., P.S., U.W., J.C.H.S., L.P. and H.N.C. X-ray radiography was carried out by F.W., K.R.B., L.A., S.A., M.O.W., J.K. and D.O. with input from M.H. and S.B. Nozzles were operated at LCLS by K.R.B., J.K., M.O.W., S.A. and D.O. CpGV was purified by K.D., M.B. and P.L.X. following a protocol developed by D.O. RNA polymerase II was produced, purified and crystallized by D.A.B. and P.J.R. *Re*MBH was prepared by A.S., S.F. and O.L., the crystallization protocol was developed by M.S., M.H. and P.S. and the project was coordinated by

Additional Information

Accession Codes: Coordinates and structure factors have been submitted to the Protein Data Bank under accession codes 5U5Q (RNA polymerase II), 5MND (CpGV) and 5TRX (HPCD).

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing Interests: The authors declare no competing financial interests.

How to cite this article: Oberthuer, D. *et al.* Double-flow focused liquid injector for efficient serial femtosecond crystallography. *Sci. Rep.* **7**, 44628; doi: 10.1038/srep44628 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.