nature methods

ARTICLES

https://doi.org/10.1038/s41592-019-0628-z

# Time-resolved serial femtosecond crystallography at the European XFEL

Suraj Pandey[1,14], Richard Bean[2,14], Tokushi Sato[2,14], Ishwor Poudyal[1], Johan Bielecki[2], Jorvani Cruz Villarreal[3], Oleksandr Yefanov[4], Valerio Mariani[4], Thomas A. White[4], Christopher Kupitz[5], Mark Hunter[5], Mohamed H. Abdellatif[4], Saša Bajt[6], Valerii Bondar[2], Austin Echelmeier[3], Diandra Doppler[3], Moritz Emons[2], Matthias Frank[7], Raimund Fromme[3], Yaroslav Gevorkov[4,8], Gabriele Giovanetti[2], Man Jiang[2], Daihyun Kim[3], Yoonhee Kim[2], Henry Kirkwood[2], Anna Klimovskaia[2], Juraj Knoska[4,9], Faisal H. M. Koua[4], Romain Letrun[2], Stella Lisova[10], Luis Maia[2], Victoria Mazalova[4], Domingo Meza[11], Thomas Michelat[2], Abbas Ourmazd[1], Guido Palmer[2], Marco Ramilli[2], Robin Schubert[11], Peter Schwander[1], Alessandro Silenzi[2], Jolanta Sztuk-Dambietz[2], Alexandra Tolstikova[4], Henry N. Chapman[4,9,12], Alexandra Ros[3], Anton Barty[4], Petra Fromme[3], Adrian P. Mancuso[2,13] and Marius Schmidt[1*]

**The European XFEL (EuXFEL) is a 3.4-km long X-ray source, which produces femtosecond, ultrabrilliant and spatially coherent X-ray pulses at megahertz (MHz) repetition rates. This X-ray source has been designed to enable the observation of ultrafast processes with near-atomic spatial resolution. Time-resolved crystallographic investigations on biological macromolecules belong to an important class of experiments that explore fundamental and functional structural displacements in these molecules. Due to the unusual MHz X-ray pulse structure at the EuXFEL, these experiments are challenging. Here, we demonstrate how a biological reaction can be followed on ultrafast timescales at the EuXFEL. We investigate the picosecond time range in the photocycle of photoactive yellow protein (PYP) with MHz X-ray pulse rates. We show that difference electron density maps of excellent quality can be obtained. The results connect the previously explored femtosecond PYP dynamics to timescales accessible at synchrotrons. This opens the door to a wide range of time-resolved studies at the EuXFEL.**

Time-resolved macromolecular crystallography (TRX) combines macromolecular structure determination with reaction dynamics[1,2]. Short and ultrashort light pulses are employed to enable snapshot observations that cope with the relevant timescales of biomolecular reactions. With TRX, biologically, biomedically and pharmacologically important reactions can be observed in real time with atomic or near-atomic spatial resolution. Hard X-ray free-electron lasers (XFELs) substantially changed the way TRX experiments were conducted[3,4], a direct consequence of the unprecedented brilliance of XFELs, and their short (femtosecond) X-ray pulses. Instead of examining macroscopically large crystals, microcrystals are injected into the X-ray beam[5,6] at room temperature. Although these microcrystals are often destroyed, the femtosecond X-ray pulse duration at the XFEL largely outruns radiation damage and the associated structural rearrangements[7–10]. Once exposed to the XFEL beam, the crystal must be replaced, demanding a serial approach where, for each new observation, a pristine microcrystal interacts with the subsequent X-ray pulse, a technique known as serial femtosecond crystallography (SFX)[5]. It has been demonstrated recently at the EuXFEL that SFX is possible with megahertz X-ray pulses[11,12]. In time-resolved SFX (TR-SFX) a reaction in a microcrystal is initiated with an optical laser during sample delivery into the X-ray interaction volume, and the progress of the reaction is probed after a time delay $\Delta t$ by the pulsed XFEL beam, as pioneered at the Linac Coherent Light Source (LCLS) at 120 Hz (refs. [3,4,13,14]). TR-SFX has the potential to take advantage of the megahertz peak rate of the European XFEL, to structurally map multiple stages of a reaction with a single experiment.

The experiments reported here examine the photocycle of PYP (Fig. 1a) using the MHz pulse structure of the EuXFEL (Supplementary Fig. 1). PYP is a bacterial photosensor, in which light triggers a reaction with several intermediates[15]. PYP is an excellent model system to establish TR-SFX at the EuXFEL, as it has been previously studied by TRX investigations at both synchrotrons and XFELs[4,14,16,17]. The photocycle is driven by the *trans* to *cis* isomerization of the central *para*-coumaric acid (pCA) chromophore[18] (Fig. 1b). In addition to being chemically highly important, its ultrafast dynamics displays similarities to other light-triggered

[1]Physics Department, University of Wisconsin-Milwaukee, Milwaukee, WI, USA. [2]European XFEL GmbH, Schenefeld, Germany. [3]School of Molecular Sciences, and Center for Applied Structural Discovery, The Biodesign Institute, Arizona State University, Tempe, AZ, USA. [4]Center for Free-Electron Laser Science, Deutsches Elektronen Synchrotron, Hamburg, Germany. [5]Linac Coherent Light Source, Stanford Linear Accelerator Center, National Accelerator Laboratory, Menlo Park, CA, USA. [6]Deutsches Elektronen Synchrotron, Hamburg, Germany. [7]Lawrence Livermore National Laboratory, Livermore, CA, USA. [8]Institute of Vision Systems, Hamburg University of Technology, Hamburg, Germany. [9]University of Hamburg, Hamburg, Germany. [10]Physics Department, Arizona State University, Tempe, AZ, USA. [11]Integrated Biology Infrastructure Life-Science Facility at the European XFEL, Schenefeld, Germany. [12]Centre for Ultrafast Imaging, Hamburg, Germany. [13]Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria, Australia. [14]These authors contributed equally: Suraj Pandey, Richard Bean, Tokushi Sato. *e-mail: smarius@uwm.edu

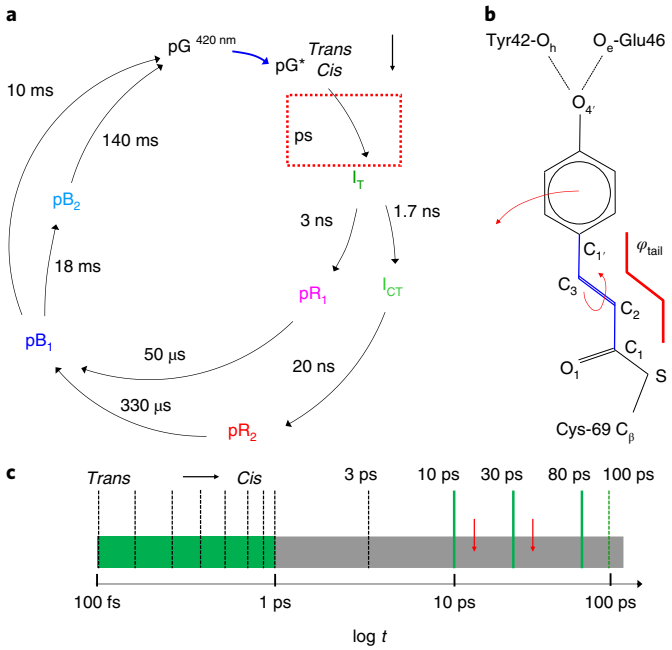


**Fig. 1 | The photocycle of PYP in crystals. a**, The photocycle (simplified) is initiated by blue light that excites the ground (dark) state pG to the electronic excited state pG*. After the *trans* to *cis* isomerization at 600 fs, several electronic ground-state intermediate states, called $I_T$, $pR_1$, $pR_2$, $pB_1$ and $pB_2$, are populated on various timescales until the photocycle completes. Approximate relaxation times are shown. Red dotted box: relaxations on the picosecond timescale. **b**, The chemical structure of the pCA chromophore bound to the Cys 69 sulfur. The *trans* configuration is shown. The torsional angle $\varphi_{tail}$ as defined by chromophore carbon atoms $C_1$–$C_2$=$C_3$–$C_{1'}$ is outlined in red. Hydrogen bonds between the pCA head and Glu 46 and Tyr 42 are marked. The rotation about the double bond as well as the head displacement at longer times are shown by arrows. **c**, The ultrafast timescale from 100 fs to 100 ps. Black dashed bars: time delays collected at the LCLS (Pande et al.[14]), green dashed bar: time delay collected at APS (Jung et al.[33]). Green solid bars: time delays in the 1 ps to 100 ps range (gray) as collected in this study. Red arrows: picosecond processes observed spectroscopically (Creelman et al. [23]).

reactions, including photoisomerization reactions in rhodopsin in the mammalian eye[19], and in other biologically relevant photoreceptors such as the phytochromes[20]. Furthermore, PYP has become a prominent optogenetic tool that can be used for the spatiotemporal optical control of complex biological processes, such as neural activity[21]. The photocycle of PYP has been extensively investigated from femtoseconds to seconds[14,17,22]. However, the time range between 1 ps and 100 ps has not been investigated in detail so far (Fig. 1c), with at least one more process observed by spectroscopy[23] (Fig. 1c, red arrows) for which there is currently no experimental structural evidence.

At the EuXFEL, X-rays arrive in pulse trains at 10 Hz (Fig. 2a and Supplementary Fig. 1). Each train consisting of bursts of X-ray pulses with an intratrain rate of up to 4.5 MHz (ref. [11]). In the current operational configuration, each train contains up to 176 pulses at a maximum rate of 1.13 MHz. This amounts to 1,760 pulses per second—already almost 15 times more pulses per unit time than the next highest repetition rate, hard X-ray FEL. A high-intensity MHz optical laser system has been commissioned recently and is now available at the SPB/SFX instrument of the EuXFEL[24]. Our experiments require the synchronization of each optical laser pulse with a corresponding X-ray pulse, in the end at MHz rates, and each time with picosecond precision. The high pulse repetition rate offers new opportunities for TR-SFX investigations at the EuXFEL; closely spaced time delays can be collected rapidly to cover processes in biomolecules in detail. At other XFELs, the low pulse repetition rate limits the amount of data that can be collected during sparsely available beamtimes. Additional parameters such as temperature[25], laser pulse duration and laser chirp may then be varied to control[26] and direct the biomolecular reaction.

## Results

**TR-SFX experiments.** A dense microcrystalline slurry of PYP was prepared and injected into the vacuum chamber of the SPB/SFX instrument[27] at the EuXFEL. The microcrystals were exposed to the trains of X-ray bursts. X-ray diffraction patterns were collected by the 'adaptive gain integrating pixel detector' (AGIPD)[28] operating with MHz frame rates (Supplementary Fig. 1). The PYP photocycle was initiated using laser pulses of 240 fs at a wavelength of 420 nm with a flux density of 1.6 mJ mm$^{-2}$ in a 42-µm (full-width half maximum, FWHM) focal spot. The viscosity of the dense microcrystalline slurry placed an upper limit to the achievable jet speed of 30 m s$^{-1}$ (determined in the laboratory under similar injection conditions). Initial measurements were conducted to establish optimal X-ray and laser pulse rates at the achievable jet speed. Laser pulse rates and X-ray pulse structures that were used here are shown in Fig. 2b,c.

First, we collected SFX data without any laser excitation at 1.13 MHz X-ray repetition rate to establish a suitable X-ray pulse rate to ensure that the sample was being refreshed between X-ray pulses (Supplementary Table 1, pure 'dark'). Next, we exposed crystals to the optical laser at 375 kHz repetition rate (every third X-ray pulse; Fig. 2b, control experiment) to determine when and whether the laser-excited jet volume had passed the X-ray interaction region. Data statistics are shown in Supplementary Table 2. With a jet velocity of 30 m s$^{-1}$ and a laser focus of 42 µm, the excited volume should leave the X-ray interaction region within 2 µs. Accordingly, the difference electron density (DED) map at the 2.67 µs time delay should be free of signal. However, as shown in Fig. 3, the resulting DED maps display signal at all time delays. The same density features are observed (α for negative features, and β1 and β2 for positive features) in all difference maps. All three maps are essentially identical, and contain a mixture of PYP intermediates pR1 and pR2 that persist in the early µs time range[17]. Compare, for example, the structure displayed in Fig. 3a determined at 1 µs delay at the LCLS[4] with those in Fig. 3b,c. In addition, the hit rate abruptly decayed from 2% at the first X-ray pulse in the train, to 1% (Supplementary Fig. 2a) in all subsequent pulses. This shows that jet velocities that are achievable with our dense PYP slurry do not reliably replace the sample at the X-ray interaction point at the 1.13 MHz X-ray repetition rate. Consequently, both the X-ray pulse repetition rate and the laser repetition rate are too high for our planned picosecond TR-SFX experiment. However, there is no indication that there are structural differences caused by the 1.13 MHz X-ray pulse rate. Values for $R$ / $R_{free}$ values of the reference model refined against the data collected in the dark are 17% / 24%, respectively, with no obvious differences in $F_{obs}$–$F_{calc}$ difference maps. It seems that, at least in the case of PYP, the crystals are not affected by acoustic shockwaves observed earlier[29] and the dominant effect is an absence of crystals at the interaction point.

Following these observations, the X-ray repetition rate was reduced to 564 kHz and the laser repetition rate was reduced to 141 kHz so that laser excitation was achieved before every fourth X-ray pulse (Fig. 2c). At the 0.56 MHz X-ray repetition rate, the hit rate across the entire pulse train remains essentially constant (Supplementary Fig. 2b). This shows that at 0.56 MHz the sample is sufficiently refreshed before the next X-ray pulse arrives. The pump–probe delay was set to a value of 10 ps, with subsequent X-ray pulses measuring delays of 1.78 µs, 3.56 µs and 5.33 µs

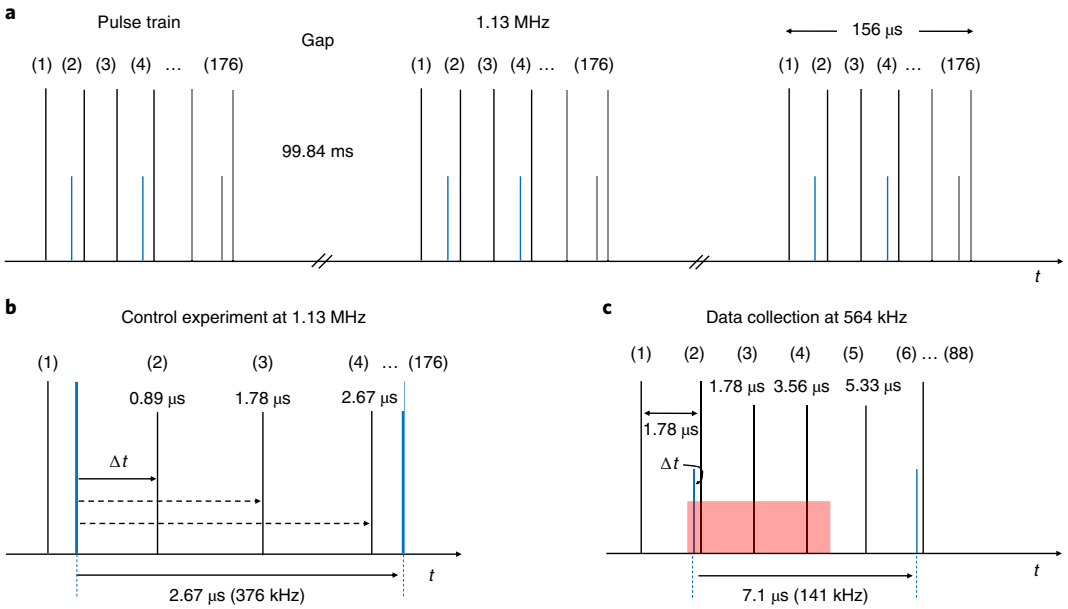


**Fig. 2 | Pulse train structure and laser excitation. a**, X-ray pulse trains (black vertical lines) at EuXFEL with 1.13-MHz pulse repetition rate. A pulse train is 156 µs long, contains 176 X-ray pulses and repeats ten times per second. There are 99.84-ms gaps between the pulse trains. Blue: laser pulses for a pump–probe dark TR-SFX data collection scheme. Note: when EuXFEL design specifications are reached, 2,700 pulses with up to 4.5 MHz pulse repetition rate are in a train. At 4.5 MHz, each pulse train is 600 µs long with 99.4-ms gaps between the trains. In total there are 27,000 pulses per second, a subset of which (about 3,520 pulses per second) can be stored in, and read out by, the AGIPD detector. **b**, The 1.13 MHz control experiment with 376 kHz laser excitation. After the laser pulse, subsequent X-ray pulses arrive at 887 ns, 1.78 µs and 2.67 µs. The sequence repeats until the end of the pulse train. **c**, The 564 kHz data collection with three interleaved X-ray pulses. There are 88 pulses in the train, only. The laser pulses are separated by 7.1 µs (141 kHz) to provide enough time for the laser-excited volume (red) to move out of the X-ray interaction region. A total of 519,336 diffraction patterns were averaged to determine the scheme.

(see Supplementary Table 3 for data statistics). At the 10 ps time delay, we observe a DED map with strong features (Fig. 4a) that resemble the DED map collected at a 3 ps delay at the LCLS (Fig. 5a)[14]. The DED map calculated with data from the following X-ray pulse at 1.78 µs (Fig. 4b) differs completely from the 10 ps DED map, and shows the same pattern of DED features as described in Fig. 3. It can be interpreted with the previously described mixture of intermediates with microsecond lifetimes. The microsecond DED signal decays to a spurious positive feature caused by the displacement of the electron rich Cys 69 sulfur at the 3.56 µs delay (Fig. 4c, blue arrow). This feature vanishes completely at the 5.33 µs delay, which means that a short time after 3.56 µs the laser-excited jet volume left the X-ray interaction region. Since the laser profile is Gaussian, with an FWHM of 42 µm, there is still substantial intensity in the flanks of the spot, which may cause contaminations up to the 3.56 µs time delay. These experiments demonstrate how the TR-SFX experiment must be set up when operating within our experimental conditions.

**Picosecond time series.** In addition to the 10 ps time point, we collected TR-SFX datasets at 30 ps and 80 ps time delays, with the successful timing scheme described. Fluorinated oil was added to the crystalline slurry at various ratios (Supplementary Table 1) via a T-junction (Supplementary Fig. 1). The T-junction was located close to the outside end of the nozzle rod. The oil flowed through the long capillary across the rod and was injected together with the microcrystals. In addition to a slight increase of the liquid flow rate (Supplementary Table 1), the resulting jets became substantially elongated, from 100 µm to 700 µm. This allowed us to increase the distance from the nozzle to the interaction region to 400 µm, which helped to prevent the accumulation of debris on the nozzle tip that would otherwise quickly interfere with data collection. The 10 ps to 80 ps time delays cover the region previously unknown in the photocycle, probing a time region between the published LCLS data at 3 ps and synchrotron data at 100 ps. Supplementary Table 2 lists the data statistics.

In Fig. 5, the DED maps at 10 ps, 30 ps and 80 ps are shown together with those obtained previously at LCLS and at APS at 3 ps and 100 ps, respectively. In short, all time delays collected at the EuXFEL resulted in excellent DED maps that contain chemically meaningful positive and negative DED features (α and β in Fig. 5). The DED maps (Fig. 5b–d) are similar and comparable to those obtained at other X-ray sources on the ps timescale (Fig. 5a,e). The positive and negative DED features are interpreted by structural models using extrapolated maps[16,30–32] (see Methods; Supplementary Tables 4 and 5 list the refinement statistics). In addition to the newly collected time delays of 10, 30 and 80 ps, we also revisited the 3 ps (ref. [14]) and the 100 ps (ref. [33]) data collected previously at LCLS and APS, respectively (Supplementary Fig. 3). We subjected all data across the time window from 3 ps to 100 ps to our objective procedures described in the Methods, to ensure consistent results.

Population transfer (PT) for each time point in this experiment is approximately 7% (see Supplementary Table 6), which is lower compared to similar excitation schemes at other XFELs. With a femtosecond laser pulse we are limited to the primary photoexcitation yield, which is 20% at best when excitation is achieved into the absorption maximum[34]. More details on how to estimate the PT are given in the Methods. The yield is further diminished here by illuminating into the flanks of the absorption spectrum: excitation was achieved at 420 nm rather than into the central absorption peak at 450 nm (Supplementary Fig. 4b). Still, excellent data can be collected because the laser penetration depth matches the micrometer crystal size[4], leading to uniform sample excitation. From the side view of the pCA chromophore at various time delays in Fig. 5f–j,

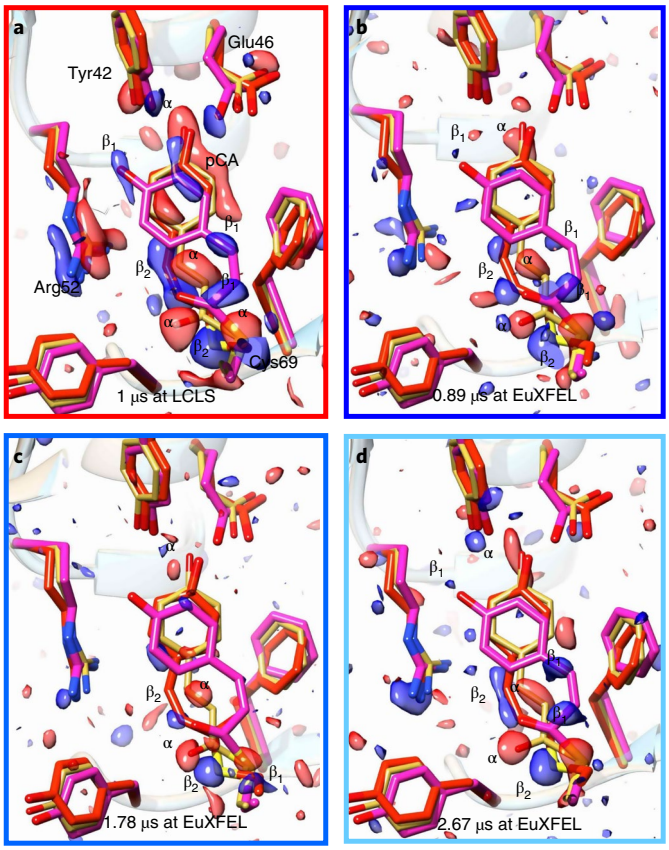


**Fig. 3 | TR-SFX experiments at LCLS and EuXFEL. a**, DED in the PYP chromophore pocket at 1 μs time delay, as determined at the LCLS (Tenboer et al.[4]). Red: negative DED, blue: positive DED on the $-3\sigma/3\sigma$ contour levels, respectively. Prominent features are labeled α (negative) or $\beta_1$ and $\beta_2$ (positive). Features labeled α are on top of the reference structure (yellow), $\beta_1$ and $\beta_2$ features correspond to intermediate structures called pR1 (magenta) and pR2 (red), respectively. The pattern of α and $\beta_1$, $\beta_2$ features persists in all maps at all times. **b**–**d**, Results of the control experiment with 1.13 MHz X-ray pulse repetition and 376 kHz laser excitation (see also Supplementary Fig. 2b): 0.89 μs after the laser pulse (**b**), 1.78 μs after the laser pulse (**c**), 2.67 μs after the laser pulse (**d**).

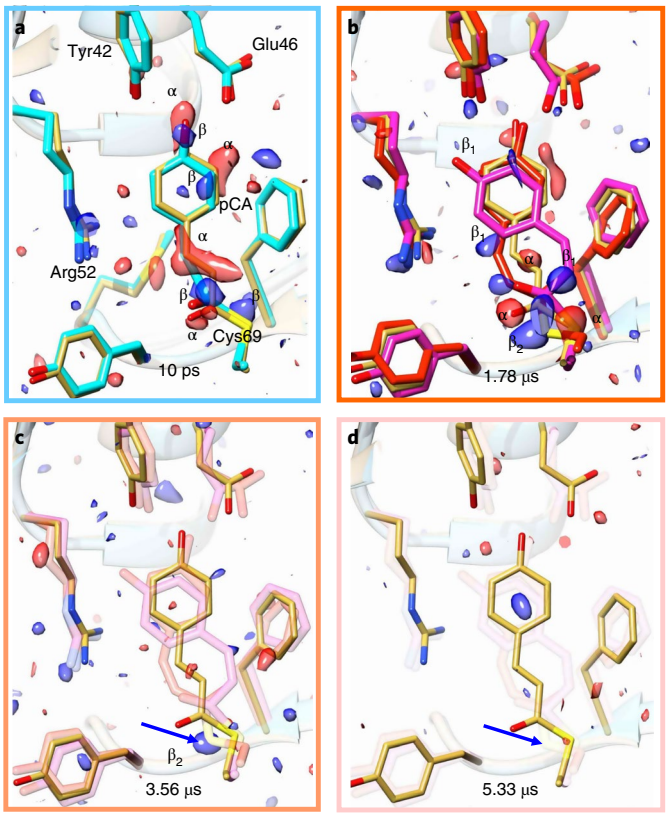


**Fig. 4 | DED and structures of the chromophore-binding region of PYP.** TR-SFX data were collected with 564 kHz X-ray and 141 kHz laser pulse repetition rates, respectively. **a**, A 10 ps time delay. Yellow: reference structure, green: 10 ps structure. Red features α: negative difference DED ($-3\sigma$ contour level), blue, β: positive DED ($3\sigma$ contour level). **b**, The pattern of DED features radically changes compared with **a**. Magenta and red: structures of the pR1 (DED features $\beta_1$) and pR2 (DED features $\beta_2$) intermediates, respectively. **c**, DED after 3.56 μs. Only the prominent feature of the Cys 69 sulfur remains (blue arrow). **d**, DED at 5.33 μs, the feature on the Cys 69 sulfur is absent (blue arrow). DED contour levels in **b**–**d** as in **a**.

one can see that the chromophore is in a twisted *cis* configuration throughout. The chromophore head is strongly tilted with the chromophore tail behind the plane of the unexcited (dark) pCA chromophore (see arrows in Fig. 5). Table 1 lists the torsional angle of the chromophore tail $\varphi_{tail}$ (Fig. 1b) obtained after refinement, as well as the lengths of the hydrogen bonds that fix the chromophore head to Tyr 42 and Glu 46. The revisited $\varphi_{tail}$ at 3 ps (39°) and 100 ps (30°) agree with results published earlier (35° and 33°, respectively)[14,22].

## Discussion

**Structural dynamics.** When comparing results at different picosecond time delays, the torsional angle at 3 ps (39°) increases at 10 ps (51°) and 30 ps (54°) and relaxes through 80 ps to a final value (30°) at 100 ps. The torsional relaxations up to 3 ps occur in concert with an initial increase of the hydrogen bond distance from the pCA-$O_4$ to Glu 46-$O_\varepsilon$ (3.3 Å). After 80 ps the hydrogen bonds subsequently relax to shorter distances that approach those observed in the dark structure. Assuming that about half of the absorbed photon energy is stored in the near-*cis* chromophore configuration (of the order of 100 kJ mol$^{-1}$)[35], the release of the chromophore head from a network of two hydrogen bonds should be possible, since the energies of the hydrogen bonds are only about 10 kJ mol$^{-1}$ (ref. [36]) each. However, for the pCA head displacements to occur, chromophore pocket relaxations are required, which are not yet developed on fast timescales. Displacements of the $M_{41-71}$ moiety (Supplementary Fig. 4a), which wraps around the chromophore pocket peak at 10 ps, revert slightly at 30 ps, and slowly increase towards 100 ps (Supplementary Fig. 4c).

The initial displacements are reminiscent of ultrafast structural dynamics detected by time-resolved experiments on myoglobin in solution[37]. Although the faster, ps time range (between 1 ps and 100 ps) is sparsely covered by previous experiments[14,22,33] and the experiments at EuXFEL included here, direct structural evidence is provided to show how an energetically highly strained structure initially reorders and then only relaxes slowly for a longer period of time (Supplementary Figs. 4c and 5a,c). This is a direct visualization of a nonexponential, nonergodic, ultrafast relaxation from a high energy state towards a longer-lived thermal reservoir, which is structurally characterized by intermediate $I_T$ (Fig. 1a). Only after 100 ps, may the PYP molecules that populate this reservoir sample configurational space more comprehensively to find a reaction coordinate that permits transition to the next intermediate state. As shown previously, transitions to two intermediates ($I_{CT}$ and pR1) are possible[33], requiring reaction coordinates that likely arise from two different positions in configurational space.

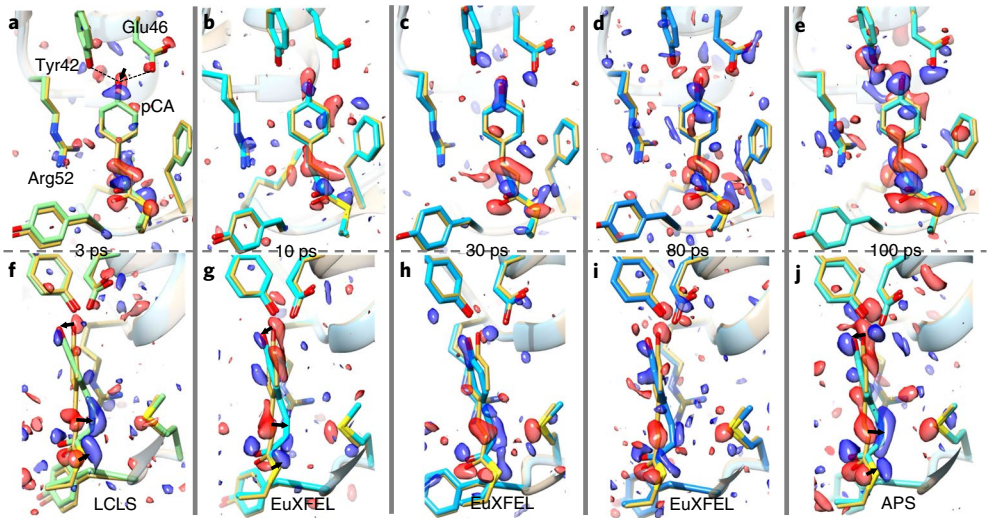


**Fig. 5 | Time series of TRX data from 3 ps to 100 ps collected at LCLS, EuXFEL and APS.** Structures and DED in the chromophore-binding region of PYP. Red: negative DED (−3$\sigma$ contour level), blue positive DED (3$\sigma$ contour level). Important residues and the pCA chromophore are marked in **a**. Yellow structure: structure of the (dark) reference state. Arrows depict structural displacements in **a**,**f** and **j**. Upper: front view, lower side view. **a**,**f**, With a 3 ps delay as collected at LCLS. Green: PYP structure at 3 ps (Pande et al.[14]). **b**,**g**, With a 10 ps time delay, this study, cyan: PYP structure at 10 ps. **c**,**h**, With a 30 ps time delay, this study, sky blue: PYP structure at 30 ps. **d**,**i**, With an 80 ps time delay, this study, blue: PYP structure at 80 ps. **e**,**j**, With a 100 ps time delay as determined at APS, light blue: PYP structure at 100 ps (Jung et al.[33]).

**Table 1 | Geometry of the pCA chromophore after refinement**

| | Dark | 3 ps | 10 ps | 30 ps | 80 ps | 100 ps |
|---|---|---|---|---|---|---|
| $N_{ext}$[a] | — | 16 | 30 | 29 | 29 | 40 |
| Hydrogen bond pCA to E46 (Å) | 2.55 | 3.30 | 2.82 | 2.87 | 2.86 | 2.79 |
| Hydrogen bond pCA to Y42 (Å) | 2.58 | 2.57 | 2.65 | 2.62 | 2.64 | 2.39 |
| $\varphi_{tail}$ (°) | 172 | 39 | 51 | 54 | 40 | 30 |

[a]Characteristic *N* used to calculate extrapolated maps The length of the hydrogen bonds from the pCA head hydroxyl to Glu 46 and Tyr 46, as well as the torsional angle $\varphi$ of the pCA tail defined by pCA carbon atoms C1, C2, C3 and C1′ is shown. For the tail in the *trans*-configuration $\varphi$ is close to 180°, in the *cis*-configuration $\varphi$ it is close to 0°.

**Data collection strategy.** In this particular experimental case, reduction of the X-ray pulse repetition rate to 564 kHz was necessary to perform TR-SFX experiments with these dense crystalline slurries, with sample jet speeds of about 30 m s$^{-1}$. The combination of jet speed and 42 µm (FWHM) laser spots demanded optical laser repetition rates not faster than 141 kHz, to minimize contamination from previous exposures. In this way, 22 laser pulses are accommodated per X-ray pulse train, which amounts to 220 laser excitations per second. This rate is between a factor 3.5 and 15 times faster than that achieved previously at other XFELs.

To push data collection rates towards the MHz range the liquid flow rate, as well as the gas flow rate that narrows the jet to boost its speed[38], must be further increased. If synthetic oil flows together with the crystalline slurry, clogging of, and debris deposition on, the gas dynamic virtual nozzle (GDVN) is largely reduced and an increase in the flow rate is possible. With higher gas flow rates, provisions such as increased pump rates to maintain high vacuum levels are required to protect the highly sensitive MHz X-ray detector. When the laser focus is reduced to about 20 µm FWHM, the laser-excited jet volume will also leave the X-ray interaction region faster. Then, 564 kHz laser pulses interleaved by 1.13 MHz X-ray pulses will push the speed of the pump–probe data collection rates to the limit, and a time delay can be collected of the order of minutes. Our results also pave the way for collecting X-ray data with femtosecond time delays at the EuXFEL. At the same time, meaningful datasets on the fast µs timescale may be obtained, which add important time delays and may serve as invaluable controls to assess signal levels in the DED maps. When longer time delays are explored, different strategies with largely reduced laser repetition rates (10 Hz) may be employed, where the reaction is initiated already within the nozzle capillary and probed, after injection, by the entire train of X-ray pulses. The MHz data acquisition rate by the pulse trains will not be affected, however, and the time delays can be swiftly collected.

**Future experiments.** The results of this experiment at the SPB/SFX instrument of the European XFEL demonstrate that TR-SFX experiments are feasible at high repetition rate X-ray lasers. High power optical laser sources are required that match the specific X-ray repetition rates and that are tunable to the photon energies needed to initiate reactions in biological macromolecules. The methods implemented here are generally applicable to comprehensively investigate macromolecular reactions within a dedicated experimental time, including built-in control measurements. The increased data rate at the EuXFEL may, in the future, support the collection of TR-SFX data within very limited experimental time (Supplementary Table. 6), assuming high uptime of all the necessary experimental apparatus. This opens the door to the deployment of enhanced analysis methods to extract macromolecular structures and their dynamics in the crystalline ensemble[39]. This has been achieved on slower timescales that were covered by dozens[25] of TRX datasets collected over multiple days, and often weeks, of beamtime at synchrotrons[17,25,40]. Close inspection and kinetic interpretation of the TR-SFX data allow an intimate view of mechanical aspects of signal transduction[41] and catalysis[42,43], with specific atomic displacements linked to specific protein function. The close relationship between structural dynamics and function established in this way provides new avenues for the control[26,44,45] and understanding of biological function, which then also paves the way to a deeper understanding of the mechanism of biomacromolecular reactions in biomedically and biologically important macromolecules.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-019-0628-z.

Received: 10 June 2019; Accepted: 3 October 2019;
Published online: 18 November 2019

## References

1. Moffat, K. Time-resolved biochemical crystallography: a mechanistic perspective. *Chem. Rev.* **101**, 1569–1581 (2001).
2. Schmidt, M. Time-resolved macromolecular crystallography at modern X-ray sources. *Methods Mol. Biol.* **1607**, 273–294 (2017).
3. Aquila, A. et al. Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Opt. Express* **20**, 2706–2716 (2012).
4. Tenboer, J. et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* **346**, 1242–1246 (2014).
5. Chapman, H. N. et al. Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011).
6. Boutet, S. et al. High-resolution protein structure determination by serial femtosecond crystallography. *Science* **337**, 362–364 (2012).
7. Lomb, L. et al. Radiation damage in protein serial femtosecond crystallography using an X-ray free-electron laser. *Phys. Rev. B* **84**, 214111 (2011).
8. Nass, K. et al. Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *J. Synchrotron Radiat.* **22**, 225–238 (2015).
9. Suga, M. et al. Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature* **543**, 131–135 (2017).
10. Chreifi, G. et al. Crystal structure of the pristine peroxidase ferryl center and its relevance to proton-coupled electron transfer. *Proc. Natl Acad. Sci. USA* **113**, 1226–1231 (2016).
11. Wiedorn, M. O. et al. Megahertz serial crystallography. *Nat. Commun.* **9**, 4025 (2018).
12. Grünbein, M. L. et al. Megahertz data collection from protein microcrystals at an X-ray free-electron laser. *Nat. Commun.* **9**, 3487 (2018).
13. Barends, T. R. et al. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350**, 445–450 (2015).
14. Pande, K. et al. Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science* **352**, 725–729 (2016).
15. Meyer, T. E., Yakali, E., Cusanovich, M. A. & Tollin, G. Properties of a water-soluble, yellow protein isolated from a halophilic phototrophic bacterium that has photochemical activity analogous to sensory rhodopsin. *Biochemistry* **26**, 418–423 (1987).
16. Genick, U. K. et al. Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science* **275**, 1471–1475 (1997).
17. Ihee, H. et al. Visualizing reaction pathways in photoactive yellow protein from nanoseconds to seconds. *Proc. Natl Acad. Sci. USA* **102**, 7145–7150 (2005).
18. Kort, R. et al. Evidence for *trans–cis* isomerization of the *p*-coumaric acid chromophore as the photochemical basis of the photocycle of photoactive yellow protein. *FEBS Lett.* **382**, 73–78 (1996).
19. Polli, D. et al. Conical intersection dynamics of the primary photoisomerization event in vision. *Nature* **467**, 440–443 (2010).
20. Mathes, T. et al. Femto- to microsecond photodynamics of an unusual bacteriophytochrome. *J. Phys. Chem. Lett.* **6**, 5 (2014).
21. Ali, A. M. et al. Optogenetic inhibitor of the transcription factor CREB. *Chem. Biol.* **22**, 1531–1539 (2015).
22. Schotte, F. et al. Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proc. Natl Acad. Sci. USA* **109**, 19256–19261 (2012).
23. Creelman, M., Kumauchi, M., Hoff, W. D. & Mathies, R. A. Chromophore dynamics in the PYP photocycle from femtosecond stimulated Raman spectroscopy. *J. Phys. Chem. B* **118**, 659–667 (2014).
24. Palmer, G. et al. Pump-probe laser system at the FXE and SPB/SFX instruments of the European X-ray free-electron laser facility. *J. Synchrotron Radiat.* **26**, 328–332 (2019).
25. Schmidt, M. et al. Protein energy landscapes determined by five-dimensional crystallography. *Acta Crystallogr. D* **69**, 2534–2542 (2013).
26. Prokhorenko, V. I. et al. Coherent control of retinal isomerization in bacteriorhodopsin. *Science* **313**, 1257–1261 (2006).
27. 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**, 660–676 (2019).
28. Allahgholi, A. et al. The adaptive gain integrating pixel detector at the European XFEL. *J. Synchrotron Radiat.* **26**, 74–82 (2019).
29. Stan, C. A. et al. Liquid explosions induced by X-ray laser pulses. *Nat. Phys.* **12**, 966–971 (2016).
30. Tripathi, S., Srajer, V., Purwar, N., Henning, R. & Schmidt, M. pH dependence of the photoactive yellow protein photocycle investigated by time-resolved crystallography. *Biophys. J.* **102**, 325–332 (2012).
31. Schmidt, M. in *Ultrashort Laser Pulses in Medicine and Biology* (eds Braun, M. et al.) 201–241 (Springer, 2008).
32. Schmidt, M. Time-resolved macromolecular crystallography at pulsed X-ray sources. *Int. J. Mol. Sci.* **20**, 1401 (2019).
33. Jung, Y. O. et al. Volume-conserving *trans–cis* isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nat. Chem.* **5**, 212–220 (2013).
34. Hutchison, C. D. M. & van Thor, J. J. Populations and coherence in femtosecond time resolved X-ray crystallography of the photoactive yellow protein. *Int, Rev. Phys. Chem.* **36**, 117–143 (2017).
35. Groenhof, G. et al. Photoactivation of the photoactive yellow protein: why photon absorption triggers a trans-to-cis isomerization of the chromophore in the protein. *J. Am. Chem. Soc.* **126**, 4228–4233 (2004).
36. Markovitch, O. & Agmon, N. Structure and energetics of the hydronium hydration shells. *J. Phys. Chem. A* **111**, 2253–2256 (2007).
37. Levantino, M. et al. Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nat. Commun.* **6**, 6772 (2015).
38. DePonte, D. P. et al. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Phys. D* **41**, 195505 (2008).
39. Schmidt, M., Rajagopal, S., Ren, Z. & Moffat, K. Application of singular value decomposition to the analysis of time-resolved macromolecular X-ray data. *Biophys. J.* **84**, 2112–2129 (2003).
40. Rajagopal, S., Schmidt, M., Anderson, S., Ihee, H. & Moffat, K. Analysis of experimental time-resolved crystallographic data by singular value decomposition. *Acta Crystallogr. D* **60**, 860–871 (2004).
41. Kang, Y. et al. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* **523**, 561–567 (2015).
42. Kupitz, C. et al. Structural enzymology using X-ray free electron lasers. *Struct. Dyn.* **4**, 044003 (2017).
43. Olmos, J. L. Jr. et al. Enzyme intermediates captured "on the fly" by mix-and-inject serial crystallography. *BMC Biol.* **16**, 59 (2018).
44. Paul, K. et al. Coherent control of an opsin in living brain tissue. *Nat. Phys.* **13**, 1111–1116 (2017).
45. Wang, J. et al. Time-resolved protein activation by proximal decaging in living systems. *Nature* **569**, 509–513 (2019).

## Methods

A step-by-step protocol for a successful TR-SFX experiment with PYP at the EuXFEL is available on the Protocol Exchange[46]. This protocol can be readily modified for other (photoreactive) biological macromolecules.

**Sample preparation.** PYP was overexpressed and purified as reported[4,18]. PYP microcrystals were grown with the stir method using 3.3 mol $l^{-1}$ malonate (pH 7) as precipitant[4]. PYP was concentrated to 100 mg $ml^{-1}$, and 4 mol $l^{-1}$ Na-malonate, pH 7, was added at once to a final concentration of 3.3 mol $l^{-1}$ under vigorous stirring. The suspension (20 ml) was stirred in a closed glass vial for 8 h and allowed to rest for an additional 24 h at room temperature. This method works equally well with smaller ~2 ml and larger ~20 ml volumes. Stirring is necessary to prevent the growth of crystals to sizes larger than 10 µm. The slurry was spun at 8,000*g* for 10 min. The microcrystals swim up. The clear solution below the microcrystals was removed and replenished by 2.8 mol $l^{-1}$ Na-malonate, pH 7. This resulted in a suspension containing about $10^9$ to $10^{10}$ crystals per ml, with most crystals being about 5 µm, as observed under a microscope with ×400 magnification using a Neubauer cell counting chamber. The dense PYP microcrystalline slurry was filtered twice by handpressing the slurry (placed in a 10 ml syringe) through a 10-µm stainless steel filter. The slurry was injected without further filtering into the vacuum chamber at the SPB/SFX imaging and serial crystallography instrument[27,47] using a GDVN[38] with 75 µm inner diameter (Supplementary Fig. 1). Smaller nozzle diameters led to clogging and reduced flow rate due to pressure limitations.

**Injection and alignment.** With the 75 µm inner diameter nozzles, we attempted to maximize slurry flow and gas pressure to produce a fast jet that may be able to cope with the enormous X-ray pulse rates. We measured the jet diameter to be about 5 µm, with a flow rate of 35 µl $min^{-1}$ under identical conditions to those used for the experiment (Supplementary Table. 1). This translates into a jet speed of approximately 30 m $s^{-1}$. To avoid clogging, for some time delays, we added an immiscible fluorinated oil mixture (perfluorodecalin and 1*H*,1*H*,2*H*,2*H*-perfluoro-1-octanol in a 10:1 ratio) through a T-junction (Supplementary Fig. 1) located upstream of the injector nozzle rod at various flow rate ratios (Supplementary Table 1). The GDVN tip surface, as well as the capillaries to the GDVN, was treated with Novec 1720 (ref. [48]) to stabilize the simultaneous injection with oil and avoid growth of stalactite-like debris on the nozzle, which originates from the jet explosion after X-ray exposure. With this treatment, the oil can form a lubricant layer between the walls and the aqueous jet, which minimizes interaction and sticking of crystals and effectively avoiding clogging. As observed earlier[11], exposure to the intense X-ray pulses resulted in gaps in the jet caused by Coulomb explosion. The PYP photocycle was initiated in the microcrystals with 240 fs laser pulses of wavelength 420 nm, which is 30 nm on the blue side of the PYP absorption maximum (Supplementary Fig. 4b). As established earlier by ultrafast spectroscopy on PYP crystals[49], laser excitation (into the absorption maximum) with 1 mJ $mm^{-2}$ is essentially free of nonlinear effects. This dispels concerns voiced recently about other systems[50]. Since the absorption at 420 nm is only 60% of that of the absorption maximum (Supplementary Fig. 4b), we used, accordingly, 1.6 mJ $mm^{-2}$ laser fluence in a 42 µm (FWHM) laser spot. Imaging the gaps in the jet[11] provides a convenient method to align the laser with the X-rays by centering the laser focal spot in the gap. Temperature was not controlled. Laser warming (about 10 °C)[25] and evaporative cooling effects (a few degrees) fortuitously compensate each other. Details of the experimental set-up, which includes laser set-up, laser alignment, the timing and the determination of the temporal overlap of X-ray and laser pulses are given in the following sections.

**Instrumentation.** Experiments were performed at the SPB/SFX instrument in March 2019 as a part of proposal 2166 using a similar configuration as that used in Wiedorn et al.[11] (Supplementary Fig. 1). The size of the mirror-focused focal spot[51] in the interaction region was estimated to be 2 × 3 $µm^2$ FWHM diameter based on optical imaging of single shots using a 20 µm thick Ce:YAG screen. The X-ray pulse energy was about 700 µJ. Diffraction from the sample was measured using an AGIPD[28,52] of 1 megapixel located 117.7–118.6 mm downstream of the sample interaction region, with the unused direct beam passing through a central hole in the detector to a beam stop further downstream (Supplementary Fig. 1). The resolution at the edge of the AGIPD was 1.8 Å, and 1.6-Å data were obtained by integrating Bragg reflections into the detector corner. Experiment control was provided by Karabo[53] and data acquisition was provided by dedicated technology to cope with the megahertz repetition rates[54].

**Laser set-up and timing.** Optical laser radiation, with wavelength 840 nm and 15 fs pulse duration, was used to generate a 420 nm second harmonic using a 0.5 mm thick beta barium borate (BBO) crystal. Additionally, fused silica windows and lenses stretched the pulse to 250 fs duration. The beam size at the sample position was measured by a scintillator and a simple in situ microscope, yielding a diameter of 42 µm FWHM. The average laser energy was about 2.3 µJ, which corresponds to 1.6 mJ $mm^{-2}$ fluence at the sample. The optical laser timing was synchronized with a radiofrequency (RF) signal, and timing at the experiment was controlled by a phase shifter. The T0 position, when both optical and X-ray beams arrive simultaneously, was estimated by the spatial encoding method using a $SrTiO_3$ crystal[55]. Both the inter-train timing and the intratrain timing were measured earlier in facility experiments. The train arrival time jitter was determined to be ~300 fs for the system in RF lock[55]. The intratrain jitter was measured to be significantly shorter than 1 ps, which is negligible in terms of the ps timing scheme employed here. In addition, we never observed more than 1 ps drift in our RF synchronization over a 12 h shift.

**Data processing.** Experiment progress was monitored online using OnDA[56] for serial crystallography. Diffraction images with Bragg reflections were found by Cheetah[57] (peakfinder8, minSNR = 8, minADC = 200, minPix = 1, minPeaks = 25) using the calibration process described by Wiedorn et al.[11]. Careful masking of shadowed and unreliable regions of the detector was performed on a run-by-run basis. Independent masks were used for peakfinding to avoid false peaks, for example, due to ice formation. Indexing was performed using CrystFEL v.0.8.0 (ref. [58]) with peaks found by Cheetah using the indexing package XGANDALF (ref. [59]). Detector geometry, especially the detector distance, was refined using the program Geoptimiser[60]. Merging and scaling of the Bragg peak intensities were performed using the partialator program from CrystFEL. To avoid the integration of noise for weakly scattering patterns, reflections were included up to 1.0 $nm^{-1}$ above a conservative resolution estimate for each crystal (--push-res = 1.0). Since PYP crystallizes in $P6_3$, an indexing ambiguity[61] is present[4]. This was corrected by the ambigator module in CrystFEL. Figures of merit were calculated using compare_hkl ($R_{split}$, $CC_{1/2}$, $CC^*$) and check_hkl (signal-to-noise ratio, multiplicity, completeness), both a part of CrystFEL. The intensities from all indexed patterns were scaled together irrespective of the delay setting. Individual time delays were separated, and their intensity merged. This separation has been achieved based on pulse identities in the train (see below and Supplementary Fig. 1), which are stored together with the diffraction patterns. The corresponding intensities were then merged to generate reference datasets, and datasets at different time delays, for each (Supplementary Tables 1–3).

**Pump–probe timing schemes.** X-ray pulses arrive in pulse trains with, currently, up to 176 X-ray pulses with a 1.13 MHz repetition rate within the train. Each train repeats with 10 Hz (Fig. 2a and Supplementary Fig. 1). The tunable, high-energy femtosecond laser system installed at the SPB/SFX instrument[24] is able to cope with the MHz pulse repetition rate. For the various experiments, two different pump–probe timing strategies were used. These are shown in Fig. 2. The laser was synchronized to X-ray pulse 1 in each train, whose radio frequency signal is delivered by the EuXFEL control room to the instrument. The X-ray fluence in pulse 1 has been very low for this experiment, and spurious diffraction patterns produced by it were not used. In scheme 1, the X-ray pulse rate was1.1 MHz (Fig. 2b). The laser was activated 2.3 ns after pulse 1. Accordingly, pulses 2, 3 and 4 probed the reaction after 0.89 µs (887 ns), 1.78 µs and 2.67 µs. This sequence repeats with laser activation after pulses 4, 7, 10 and so on, interleaved with the three µs time delays each. This results in a laser pulse repetition rate of 376 kHz, and 176/3 ≅ 58 laser activations per pulse train. The effective laser excitation rate is therefore 580 Hz. As demonstrated in Supplementary Figs. 2a and 3, both the X-ray repetition rate and the intratrain laser repetition rate are too large for the jet speeds achieved with the viscous PYP microcrystalline slurry. In scheme 2 the X-ray pulse repetition rate in the train is reduced to 564 kHz with 88 pulses per train (Fig. 2c). This time, the hit rate varies smoothly with the X-ray pulse energy across all X-ray pulses in the train (Supplementary Fig. 2b), no abrupt decay as observed for higher X-ray pulse rates (as in Supplementary Fig. 2a) was observed. The laser was synchronized again with respect to X-ray pulse 1. This time, the synchronization was precisely adjusted, so that the pump–probe delay $\Delta t$ between the laser pulse and X-ray pulse 2 was on the picosecond timescale (Fig. 2c), which is necessary to collect data for the 10 ps, 30 ps and 80 ps time delays. As shown for other XFELs, the XFEL-to-laser timing fluctuations are of the order of 300 fs (ref. [62]), which is negligible on the ps timescale, and a timing tool[63,64] was not required. The next laser pulse arrives after X-ray pulse 5. The laser pulse repetition rate was therefore 141 kHz. Accordingly pulses 2, 6, 10, etc. in the train probe a picosecond time delay, and three additional interleaving X-ray pulses probe time delays at 1.78 µs, 3.56 µs and 5.33 µs. This results in 22 pump–probe sequences with ps time delays per train, and 220 effective laser excitations per second. At the Spring-8 Angstrom Compact Laser and the LCLS, typical X-ray pulse rates are 30 Hz and 120 Hz, respectively. Pump–probe sequences with interleaving dark data collection require 15 Hz or 60 Hz laser pulse rates. Compared with these machines, even with the small number of 88 X-ray pulses in the train in these early experiments, the effective laser repetition rate of scheme 2 is a factor of 3.5 to 15 times faster. In the future, more than one order of magnitude more X-ray pulses will be available per train, which speeds up data collection accordingly. It must be decided on a case-by-case basis whether MHz pulse rates can be used when low viscosity slurries with small crystals are available that allow narrow GDVN orifices and enable fast jet speeds, or repetition rates of both X-ray and laser pulses must be reduced, as in the case of PYP. Of the order of 675 X-ray pulses per train, with a 1.13 MHz intratrain repetition rate, are planned to be available soon. A pump–probe data collection strategy shown in Fig. 2a that contains only one interleaved dark will be feasible. In this case 337 laser pulses per train result in the enormous effective laser

excitation rate of 3,370 Hz. As the AGIPD measures up to 352 pulses per train, 3,520 patterns (out of the 6,750) can be stored per second. With a low 2% hit rate and a 50% indexing rate, as demonstrated in this paper, the approximately 25,000 indexed diffraction patterns that are required to detect low levels of population transfer can be collected in about 20 min. This includes collection of the reference (dark) data. Protein consumption is about 10% of that expected at slower XFELs (Supplementary Table. 6), and should be between 20 mg and 40 mg per time delay, depending on the design of the experiment.

**Difference map calculation.** A reference (dark state) model $M_{ref}$ was refined by using the program REFMAC (ref. [65]) against structure factor (SF) amplitudes collected in the dark $|F_{dark}|$ without laser excitation. To check for spurious features at µs delays, the pure dark data (Supplementary Table 2) were used as a reference. The dark3 data from the 30 ps time series display superior statistics (Supplementary Table 3), and served as reference for all ps time delays. Model structure factors were calculated from $M_{ref}$ with amplitude $|FC_{ref}|$ and phase $\phi_{ref}$. The measured $|F_{dark}|$ were brought to the absolute scale by scaling them to $|FC_{ref}|$ using the CCP4 program 'scaleit' (ref. [66]). The time-dependent SF amplitudes $|F_{obs}(t)|$ were then scaled to the $|F_{dark}|$ in a second run of 'scaleit'. As a result, both $|F_{obs}(t)|$ and $|F_{dark}|$ are on the absolute scale, and are scaled together. Difference structure factor (DSF) amplitudes were calculated as: $DSF_{obs} = |F_{obs}(t)| - |F_{dark}|$. A weighting factor, $w$, for the DSFs was determined to reduce the influence of outliers[67]. The DSFs were combined with phases $\phi_{ref}$. From the weighted DSFs, a weighted DED map was calculated using the program 'fft' from the CCP4 suite of programs[66]. Although the $|F_{obs}(t)|$ and the $|F_{dark}|$ are on the absolute scale, the difference map was, due to the difference Fourier approximation[68], only on half the absolute scale. The preserved absolute scale was necessary to estimate population transfer levels, as explained below. The DED maps were best contoured on the $3\sigma/-3\sigma$ levels (Supplementary Fig. 6)

**Refinement.** Meaningful negative features in the DED map were necessarily located on top of the reference model $M_{ref}$. However, contiguous, chemically sensible positive features in the DED map must be interpreted with a new structural model ($M_{TRX}$). To determine structures from DED maps, extrapolated, conventional electron density ($ED_{ext}$) maps[14,16] were used. For extrapolated structure factor ($SF_{ext}$) amplitudes a multiple of the DSFs were added to the $|FC_{ref}|$: $SF_{ext} = |FC_{ref}| + N \times DSF$ and combined with the reference state (dark) phases $\phi_{ref}$. Here the use of $|FC_{ref}|$ derived from an accurately refined dark state model was preferred over the $|F_{dark}|$, as explained by Terwilliger and Berendsen[69]. From the phased $SF_{ext}$, extrapolated electron density maps ($ED_{ext}$) were calculated with the CCP4 program 'fft'. A characteristic $N_{ext}$ was established when the electron density in the $ED_{ext}$ at the positions with strong negative features in the DED maps just vanishes. When $N$ is too large, false-negative features will appear in the $ED_{ext}$. This can be visualized by summing up negative values in the $ED_{ext}$ within a volume that contains strong DED features in the DED maps. Supplementary Fig. 7 shows results for such a summation for all our TR-SFX time delays collected at the EuXFEL and for TRX data selected from the literature[4,14,33]. The $N_{ext}$ is marked with an arrow in Supplementary Fig. 7. The value of $N_{ext}$ is approximately related to the PT: $PT \approx \frac{100}{N_{ext}} \times 2$ (%). The factor of two accounts for the difference Fourier approximation mentioned above. If the PT is small, $N_{ext}$ is large. A value of $N_{ext} = 30$ is not uncommon in TRX, especially with fs excitation, since the primary yield of photoactivation can be quite small[34] and is further diminished by experimental circumstances. Once $N_{ext}$ had been established, structural models were determined from the resulting $ED_{ext}$ maps. The $ED_{ext}$ map was displayed in a molecular modeling program such as 'coot'[70]. The reference model can be used as an initial model for a refinement. When structural changes were small, the initial model was altered by directly refining it against the $ED_{ext}$ map by a stepped real-space refinement in 'coot' with the torsional restraint switched off (default in 'coot'). For PYP, isomerization and structural changes were modeled automatically in this way, entirely without manual intervention. For other systems[71] structural changes can be modeled manually, analogous to conventional structure determination. A new structural model $M_{TRX}$ was obtained this way. From the real-space refined $M_{TRX}$ and $M_{ref}$ models, calculated $DSF_{calc}$ can be determined, this time with amplitude and phases $\phi_{\Delta}$. When the $DED_{calc}$ calculated from the phased difference structure factors is compared with the $DED_{obs}$, prominent DED features should match (Supplementary Fig. 8). The $\phi_{\Delta}$ were combined with the measured $DSF_{obs}$ and phased extrapolated SFs (pSF) were obtained by adding the (now phased) $DSF_{obs}$ to the $FC_{ref}$ as vectors in the complex plane[31,32]. The $M_{TRX}$ was refined against the |pSF| using restrained reciprocal space refinement, using, for example, 'refmac'[65]. Typically, $R$-factors were acceptable and did not deviate much from those of refinements against conventional X-ray data. Structural models and ED maps are shown in Supplementary Fig. 3. Refinement statistics are shown in Supplementary Tables 4 and 5. Selected model parameters are listed in Table 1 for the 3 ps to 100 ps time range.

**Displacements and difference distance matrices.** Structural differences were analyzed by calculating the root mean square displacements of like $C_{\alpha}$ atoms in the $M_{41-71}$ moiety (Supplementary Fig. 4a) between the structures determined at the various time delays and the reference structure. The root mean square displacements values in Supplementary Fig. 4c were fit by an empirical function consisting of an exponential term, a linear term and a strongly damped cosine function, which includes a phase shift:

$$\mathrm{RMSD}_{\mathrm{fit}} = A_0\left(1 - e^{\frac{-1}{\tau}t}\right) + bt + A_1\cos\left(\frac{2\pi}{T}t + \varphi\right)e^{-\beta t}$$

The fit values were $A_0 = 0.181$ Å, $\tau = 1.4$ ps, $b = 4.5 \times 10^{-3}$ Å ps$^{-1}$, $A_1 = 0.038$ Å, $T = 50$ ps, $\varphi = 257°$, $\beta = 1/50$ ps$^{-1}$. Some of the fit values are not unique. For example, $T$ in the cosine function can vary largely because only a few time delays are available across the 100 ps timescale. $T$ was selected so that the decrease of the root mean square displacements at 30 ps was modeled correctly, and the damping constant $\beta$ was selected so that the oscillation vanishes quickly. To show more global displacements, difference distance matrices (Supplementary Fig. 5a–d) were calculated[72] using the Fortran code of the DDMP program from the Center for Structural Biology at Yale University. The calculations include residues 42–92 and use the 3 ps structure as a reference (note: if the dark structure were to be used as a reference, small structural changes in the time window from 3 ps to 100 ps would not be observable). With these matrices one can also visually identify the decline in the magnitude of the difference distances between 10 ps and 30 ps, and their increase at 100 ps, in particular in the $M_{41-71}$ region (green bar), and then at 100 ps also more globally in the $H_{74-88}$ region (also depicted in Supplementary Fig. 4a).

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data has been deposited with the Coherent X-ray Imaging Data Bank[73] with CXIDB ID 100. This includes: stream files for all data and for data separated into each time delay, MTZ and PDB files for all time delays, including the dark/reference structures. We have deposited data (mtz-files and structures) for the 10 ps, 30 ps and 80 ps time delays, as well as the dark3 (30 ps) and pure dark reference structures, with the Protein Data Bank, with deposition codes 6P4I, 6P5D, 6P5E, 6P5G and 6P5F, respectively.

## Code availability

Linux scripts and Fortran source codes for the calculation of weighted difference maps, extrapolated electron density maps and the integration of negative densities within a spherical volume are included in a demonstration, which is available online as Supplementary Data.

## References


46. Pandey, S., Bean, R., Sato, T., Mancuso, A. P. & Schmidt, M. Time-resolved serial femtosecond crystallography at the European X-ray free electron laser. *Prot. Exch.* https://doi.org/10.21203/rs.2.14634/v1 (2019).
47. Mancuso, A. P., Aquila, A., Borchers, G., Giewekemeyer, K. & Reimers, N. *Technical Design Report: Scientific Instrument Single Particles, Clusters, and Biomolecules (SPB)* https://doi.org/10.3204/XFEL.EU/TR-2013-004 (XFEL.EU, 2013).
48. Echelmeier, A. et al. 3D printed droplet generation devices for serial femtosecond crystallography enabled by surface coating. *J. Appl. Cryst.* **52**, 997–1008 (2019).
49. Hutchison, C. D. M. et al. Photocycle populations with femtosecond excitation of crystalline photoactive yellow protein. *Chem. Phys. Lett.* **654**, 63–71 (2016).
50. Nass Kovacs, G. et al. Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin. *Nat. Comm.* **10**, 3177 (2019).
51. Bean, R. J., Aquila, A., Samoylova, L. & Mancuso, A. P. Design of the mirror optical systems for coherent diffractive imaging at the SPB/SFX instrument of the European XFEL. *J. Opt.* **18**, 074011 (2016).
52. Greiffenberg, D. The AGIPD detector for the European XFEL. *J. Instrum.* **7**, CO1103 (2012).
53. Fangohr, H. et al. Data analysis support in Karabo at European XFEL. *In Proc. of International Conference on Accelerator and Large Experimental Control Systems* (eds Costa, I. et al.) 245–252 (inSPIRE, 2017).
54. Boukhelef, D., Szuba, J., Wrona, K. & Youngman, C. *Software Development for High Speed Data Recording and Processing* (JACoW2014).
55. Kirkwood, H. J. et al. Initial observations of the femtosecond timing jitter at the European XFEL. *Opt. Lett.* **44**, 1650–1653 (2019).
56. Mariani, V. et al. OnDA: online data analysis and feedback for serial X-ray imaging. *J. Appl. Crystallogr.* 49, (1073–1080 (2016).
57. 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).
58. White, T. A. et al. Recent developments in CrystFEL. *J. Appl. Cryst.* **49**, 680–689 (2016).
59. Gevorkov, Y. et al. XGANDALF – extended gradient descent algorithm for lattice finding. *Acta Cryst. Found. Adv.* **75**, 694–704 (2019).

60. Yefanov, O. et al. Accurate determination of segmented X-ray detector geometry. *Opt. Express* **23**, 28459–28470 (2015).
61. Brehm, W. & Diederichs, K. Breaking the indexing ambiguity in serial crystallography. *Acta Crystallogr. D* **70**, 101–109 (2014).
62. Glownia, J. M. et al. Time-resolved pump-probe experiments at the LCLS. *Opt. Express* **18**, 17620–17630 (2010).
63. Harmand, M. et al. Achieving few-femtosecond time-sorting at hard X-ray free-electron lasers. *Nat. Photonics* **7**, 215–218 (2013).
64. Bionta, M. R. et al. Spectral encoding of X-ray/optical relative delay. *Opt. Express* **19**, 21855–21865 (2011).
65. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
66. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
67. Ren, Z. et al. A molecular movie at 1.8 Å resolution displays the photocycle of photoactive yellow protein, a eubacterial blue-light receptor, from nanoseconds to seconds. *Biochemistry* **40**, 13788–13801 (2001).
68. Drenth, J. *Principles of Protein X-Ray Crystallography* (Springer, 1999)..
69. Terwilliger, T. C. & Berendzen, J. Bayesian difference refinement. *Acta Crystallogr. D* **52**, 1004–1011 (1996).
70. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
71. Nogly, P. et al. Retinal isomerization in bacteriorhodopsin captured by a femtosecond X-ray laser. *Science* **361**, eaat0094 (2018).
72. Richards, F. M. & Kundrot, C. E. Identification of structural motifs from protein coordinate data—secondary structure and 1st-level supersecondary structure. *Proteins* **3**, 71–84 (1988).
73. Maia, F. R. The coherent X-ray imaging data bank. *Nat. Methods* **9**, 854–855 (2012).

## Acknowledgements

We acknowledge European XFEL in Schenefeld, Germany, for provision of X-ray free-electron laser beamtime at Scientific Instrument SPB/SFX and thank the instrument group and facility staff for their assistance. This work was supported by NSF Science and Technology Centers (grant no. NSF-1231306; Biology with X-ray Lasers) to H.N.C., P.F., A.O., A.R., P.S. and M.S.). CFEL (H.N.C.) is supported by the Gottfried Wilhelm Leibniz Program of the DFG; the project 'X-probe' funded by the European Union's 2020 Research and Innovation Program under the Marie Sklodowska-Curie grant agreement (no. 637295); the European Research Council, 'Frontiers in Attosecond X-ray Science: Imaging and Spectroscopy (AXSIS)' (no. ERC-2013-SyG 609920, together with P.F.); and the Human Frontiers Science Program grant (no. RGP0010 2017). P.F. and A.R. acknowledge the support of funding from the Biodesign Center for Applied Structural Discovery at Arizona State University and NSF award (no. 1565180). Funding from the National Institutes of Health (grant nos. R01GM095583 to P.F. and R01GM117342 to M.F.) is also acknowledged.

## Author contributions

S.P., I.P. and M.S. expressed, purified and crystallized the protein. R.B., T.S., J.B., V.B., M.E., G.G., M.J., Y.K., H.K., A.K., R.L., L.M., T.M., G.P., M.R., A.S., J.S.-D. and A.P.M. operated the SPB/SFX instrument. S.L., J.K., R.S. and H.N.C. provided injector nozzles. J.C.V., C.K., M.H., M.H.A., J.K., F.H.M.K., S.L., V.Maz., D.M., R.S. and A.T. collected the data. S.P., I.P., O.Y., V.Mar., T.A.W., Y.G., A.O., P.S., A.T. and A.B. processed the data. S.P., I.P., P.S., A.O. and M.S. analyzed the data. C.K., M.H.A., R.F. and P.F. logged the experiment. J.C.V., A.E., D.D., D.K. and A.R. conceived and operated the oil co-flow. R.B., T.S., M.F., H.N.C., A.R., A.B., P.F., A.P.M. and M.S. designed the experiment. S.P., S.B., A.B., P.F., A.P.M. and M.S. wrote the manuscript with input from all other authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41592-019-0628-z.

**Correspondence and requests for materials** should be addressed to M.S.

**Peer review information** Arunima Singh and Allison Doerr were the primary editors on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team

**Reprints and permissions information** is available at www.nature.com/reprints.

nature research

Corresponding author(s): Marius Schmidt

Last updated by author(s): Sep 11, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed |
|---|---|
| ☐ | ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☒ | ☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ A description of all covariates tested |
| ☒ | ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☒ | ☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Data is collected using Karabo 2.4.0. and Cheetah. Karabo is EuXFEL's homogeneous software framework available on-site . It is used for data acquisition, data management and data analysis. Cheetah is open source software to collect the data and perform preliminary data processing on-site at XFELS. The manuals for Cheetah is available on-line. |
| Data analysis | CrystFEL version 0.8.0., CCP4 suite of programs v7.0.075 including Coot v0.8.9.2 and refmac version 5.0.32, self written Fortran programs (available as supplementary material online). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data has been deposited with the Coherent X-ray Imaging data bank with CXIDB ID 100. This includes: Stream files for all data and for data separated into each time-delay, MTZ and PDB files for all time-delays including the dark/reference structures. We have deposited data (mtz-files and structures) for the 10 ps, 30 ps and 80 ps time-delays, as well as the dark3 {30ps) and pure dark reference structures to the protein data bank, with deposition codes 6P41, 6P5D, 6PSE, 6PSG and 6PSF, respectively.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | c.a. 1.5 g of recombinant photoactive yellow protein that yields on the order of 1.5 trillion microcrystals. The number of microcrystals were counted using the Neubauer counting chamber using well established method. |
| Data exclusions | Diffraction patterns without Bragg reflections were excluded by a "hit-finding" algorithm implemented in the Cheetah program. Exclusion criteria was determined on-site using the parameters peakfinder8, minSNR=8, minADC=200, minPix=1, minPeaks=25 in Cheetah program. |
| Replication | The experiment is a result from four 12 hrs beam time at EuXFEL. Millions of diffraction patterns were collected in total and, out of them, 421,532 patterns were meaningful and were used to determine the results. The robustness of the data were tested using various well established methods like R-split, CC-half etc. This verified the reproducibility of the results. The results can be reproduced using the data provided in CXI data bank. |
| Randomization | Orientations of individual microcrystals are naturally random caused by the injection process. A diffraction pattern is produced due to the interaction between a X-ray pulse and a random crystal. So, each diffraction pattern is also naturally random. Millions of these kinds of patterns are collected. They are indexed, integrated and merged using different programs of CrystFEL. So, a data set for every time point is obtained from of tens of thousands of random diffraction patterns. |
| Blinding | This is a macromolecular crystallographic experiment. Blinding does not apply. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |