

Received 21 July 2017

Accepted 13 September 2017

Edited by E. E. Lattman, University at Buffalo, USA

Keywords: drug discovery; protein structure; X-ray crystallography; serial crystallography; time-resolved studies; lysozyme.

PDB references: lysozyme, complex with chitotriose, 2 s mixing time, phased using PDB entry 1hw, 5npj; phased using PDB entry 4et8, 5njq; 50 s mixing time, phased using PDB entry 4et8, 5njj; phased using PDB entry 1hw, 5njs

Supporting information: this article has supporting information at www.iucrj.org

Mix-and-diffuse serial synchrotron crystallography

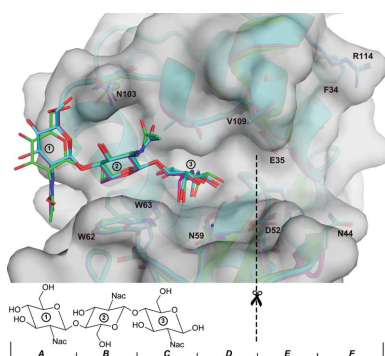
Kenneth R. Beyerlein,^{a,*} Dennis Dierksmeyer,^b Valerio Mariani,^a Manuela Kuhn,^b Iosifina Sarrou,^a Angelica Ottaviano,^c Salah Awel,^{a,d} Juraj Knoska,^{a,e} Silje Fuglerud,^{a,f} Olof Jönsson,^g Stephan Stern,^{a,h} Max O. Wiedorn,^{a,e} Oleksandr Yefanov,^a Luigi Adriano,^b Richard Bean,^a Anja Burkhardt,^b Pontus Fischer,^b Michael Heymann,^a Daniel A. Horke,^{a,d} Katharina E. J. Jungnickel,ⁱ Elena Kovaleva,^j Olga Lorbeer,^b Markus Metz,^a Jan Meyer,^b Andrew Morgan,^a Kanupriya Pande,^a Saravanan Panneerselvam,^b Carolin Seuring,^{a,d} Aleksandra Tolstikova,^a Julia Lieske,^b Steve Aplin,^a Manfred Roessle,^k Thomas A. White,^a Henry N. Chapman,^{a,d,e} Alke Meents^{a,b} and Dominik Oberthuer^{a,*}

^aCenter for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany, ^bPhoton Science, Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany, ^cDepartment of Physics, California State University, Northridge, California, USA, ^dThe Hamburg Centre for Ultrafast Imaging, University of Hamburg, 22761 Hamburg, Germany, ^eDepartment of Physics, University of Hamburg, Luruper Chaussee 149, 22607 Hamburg, Germany, ^fDepartment of Physics, Norwegian University of Science and Technology, Trondheim, Norway, ^gDepartment of Physics and Astronomy, Uppsala University, Uppsala, Sweden, ^hEuropean X-ray Free-Electron Laser Facility GmbH (XFEL), Schenefeld, Germany, ⁱDepartment of Biochemistry, University of Oxford, Oxford, England, ^jSSRL, SLAC National Accelerator Laboratory, Menlo Park, California, USA, and ^kFachhochschule Lübeck, Lübeck, Germany. *Correspondence e-mail: kenneth.beyerlein@cfel.de, dominik.oberthuer@cfel.de

Unravelling the interaction of biological macromolecules with ligands and substrates at high spatial and temporal resolution remains a major challenge in structural biology. The development of serial crystallography methods at X-ray free-electron lasers and subsequently at synchrotron light sources allows new approaches to tackle this challenge. Here, a new polyimide tape drive designed for mix-and-diffuse serial crystallography experiments is reported. The structure of lysozyme bound by the competitive inhibitor chitotriose was determined using this device in combination with microfluidic mixers. The electron densities obtained from mixing times of 2 and 50 s show clear binding of chitotriose to the enzyme at a high level of detail. The success of this approach shows the potential for high-throughput drug screening and even structural enzymology on short timescales at bright synchrotron light sources.

1. Introduction

The structural information gathered from X-ray crystallographic studies of proteins is incorporated into many stages of drug development (Congreve *et al.*, 2005; Blundell, 2017). For example, cancer, diabetes, inflammation and cardiovascular disease drugs that target kinase proteins have been developed by studying their binding sites (Parang & Sun, 2004), and anti-influenza drugs have been derived from the structure of neuraminidase (Varghese, 1999). More recently, the discovery of a new allosteric binding site in CCR9, where the ligand binds to the side of this chemokine receptor, is being explored as a new possible target for other G-protein coupled receptors (GPCRs; Oswald *et al.*, 2016). X-ray crystallographic fragment-screening studies have also shown the binding of small molecules that have an affinity that is too low to be detected by chemical assays (Schiebel *et al.*, 2016; Patel *et al.*, 2014; Erlanson *et al.*, 2016). Nonetheless, long-standing bottlenecks of solving ligand-bound protein structures include



OPEN ACCESS

(i) time-consuming screenings of substrate co-crystallization or soaking conditions to produce stable crystals; (ii) deviations from the native protein structure induced by cryogenic temperatures; and (iii) the lack of fast sample-exchange systems for efficient data collection.

Serial crystallography offers a high-throughput platform to overcome these bottlenecks. This method entails using a Bragg intensity set merged from many snapshot diffraction patterns of individual protein microcrystals to solve the protein structure (Chapman *et al.*, 2011; Boutet *et al.*, 2012). The change in paradigm of this approach allows the total scattering signal per crystal to be reduced many thousands of times, enabling measurements from small crystals (Gati *et al.*, 2017) without the need for cryogenic cooling. Its use with the brilliant femtosecond pulses of an X-ray free-electron laser has led to a number of GPCR structures (Liu *et al.*, 2013; Zhang *et al.*, 2015; Kang *et al.*, 2015), *de novo* phasing (Barends *et al.*, 2014; Yamashita *et al.*, 2015; Colletier *et al.*, 2016) and the study of fast light-induced dynamics of photoactive proteins (Tenboer *et al.*, 2014; Barends *et al.*, 2015; Nango *et al.*, 2016; Pande *et al.*, 2016).

However, most biological macromolecules are not triggered by light. Time-resolved studies of enzymatic reactions thus require different triggering methods, such as the introduction of photo-caged compounds or optogenetic methods (Moffat, 2014), both resulting in artificially light-active proteins. It was also proposed to use temperature jumps or THz-radiation to investigate protein dynamics (Neutze & Moffat, 2012; Moffat, 2014). Another approach is to simply mix one or many crystals of a biological macromolecule with a ligand or substrate (chemical triggering as opposed to physical triggering) (Schmidt, 2013). This was demonstrated in the past for protein crystals using flow-cells (Hajdu *et al.*, 1987). The further development of sample delivery methods (Calvey *et al.*, 2016; Oberthuer *et al.*, 2017) for short mixing times, and the use of very small crystals in serial crystallography leading to short ligand diffusion times into those crystals (Schmidt, 2013), enabled the first successful mix-and-diffuse experiments at the LINAC Coherent Light Source (LCLS) (Stagno *et al.*, 2016; Kupitz *et al.*, 2017), with the aim of unraveling dynamics of the respective biological macromolecules under investigation. Such continuous room-temperature sample delivery also avoids stopped-flow or freeze-trapping experiments in combination with heavy mutations of the native protein that have previously been used to study long-lived intermediates (Stoddard, 1996; Schlichting & Goody, 1997).

Serial crystallography has also been demonstrated at microfocus beamlines of third generation synchrotron sources using exposure times of a few milliseconds (Stellato *et al.*, 2014; Botha *et al.*, 2015; Nogly *et al.*, 2015; Martin-Garcia *et al.*, 2017). This exposure time matches the expected diffusion time of molecules in protein microcrystals (Schmidt, 2013), and might be suitable in many cases for fragment-based drug screening and time-resolved structural enzymology studies. As a first step in this direction, we investigated chitotriose binding to hen egg-white lysozyme *via* mix-and-diffuse serial crystallography at a synchrotron source.

Chitotriose (*N,N',N''*-triacetylchitotriose; CTO) is a natural product that competitively inhibits lysozyme (Johnson, 1998). Lysozymes degrade bacterial cell walls by cleaving the 1,4- β -linkages between *N*-acetylmuramic acid and *N*-acetyl-D-glucosamine (GlcNAc) in peptidoglycan (Blake *et al.*, 1967). In addition, some lysozymes, including that from hen egg white, can cleave between GlcNAc residues in chitodextrins such as chitin. The cleft-like polysaccharide-binding site of lysozyme can accommodate up to six GlcNAc residues. Enzymatic cleavage (Cheetham *et al.*, 1992; Vocadlo *et al.*, 2001) occurs at the linkage between the GlcNAc residues occupying subsites *D* and *E* (Maenaka *et al.*, 1995; Von Dreele, 2005); consequently, pentasaccharides or longer polysaccharides are preferably processed. Shorter saccharides such as chitotriose, chitobiose or monomeric GlcNAc still bind with reasonable affinity to lysozyme, but take longer to be released, making them a natural competitive inhibitor for structural studies (Blake *et al.*, 1967; Cheetham *et al.*, 1992; Hadfield *et al.*, 1994; Maenaka *et al.*, 1995; Vocadlo *et al.*, 2001; Von Dreele, 2005).

2. Experimental procedure

The lysozyme microcrystals used in this study were grown by the rapid-mixing batch method (Stellato *et al.*, 2014). Crystals with sizes of between 6 and 8 μm in diameter were obtained by adding three parts of precipitant [1 M NaCl, 40% (v/v) ethylene glycol, 15% (w/v) PEG 4000, 50 mM acetate buffer pH 3.5 filtered through a 450 nm filter] to one part of lysozyme (Sigma–Aldrich; dissolved to 126 mg ml⁻¹ in 50 mM acetate buffer pH 3.5 and filtered through a 100 nm filter) at 1°C (ThermoStat C, Eppendorf, Germany). The resulting mixture was immediately subjected to rapid mixing and incubated for 30 min at 1°C. Shortly before the X-ray measurements, the suspension of microcrystals was centrifuged at 2000g and the supernatant was removed to achieve a twofold higher concentration of crystals for higher diffraction-pattern collection rates. *N,N',N''*-Triacetylchitotriose {*O*-[2-Acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)]-*O*-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy-D-glucopyranose; Dextra, England} was dissolved in water and mixed with 5 M NaCl solution and 50% (w/v) PEG 4000 to achieve a concentration of 19.9 mM CTO in 1 M NaCl and 15% PEG 4000 before mixing with the microcrystal suspension.

The serial crystallography experiment was performed on the P11 beamline at PETRA III (Burkhardt *et al.*, 2016) using 13.5 keV photon energy X-rays focused to a spot of 4 \times 8 μm (width \times height) with a flux of 1.6 \times 10¹³ photons s⁻¹. A rotating beam chopper made of a 4 mm thick brass plate with holes for the X-rays to pass through was placed upstream of the focusing optics to generate X-ray pulses of 7.5 ms duration at a frequency of 25 Hz. The signal from a photodiode placed downstream of the chopper was used to trigger the readout of a PILATUS 6M detector, resulting in the collection of one diffraction image per pulse. The chopper was necessary to avoid sample heating caused by continuous exposure of the

buffer to the intense focus and to allow control of the exposure time independent of sample delivery.

Sample delivery of the microcrystals to the X-ray focus was performed using a specially developed tape-drive device. As shown in Fig. 1, the lysozyme microcrystal suspension was deposited onto the surface of polyimide tape that was continuously drawn from a feeder roll to a collector roll. The movement of the tape under the sample capillary formed a stream on the tape that was aligned with the X-ray focus. A set of rollers and motors controlled the tape speed and kept the tape under slight tension. During the experiment, a constant tape speed of 0.6 mm s^{-1} was used. This speed was chosen as it resulted in a translation distance of $24 \mu\text{m}$ between frames, which was fast enough to avoid repeated crystal exposures. As-received rolls of nonsticky polyimide tape with a width of 6 mm and a thickness of $12 \mu\text{m}$ (Caplinq, The Netherlands) were directly mounted on the device. The tape position was

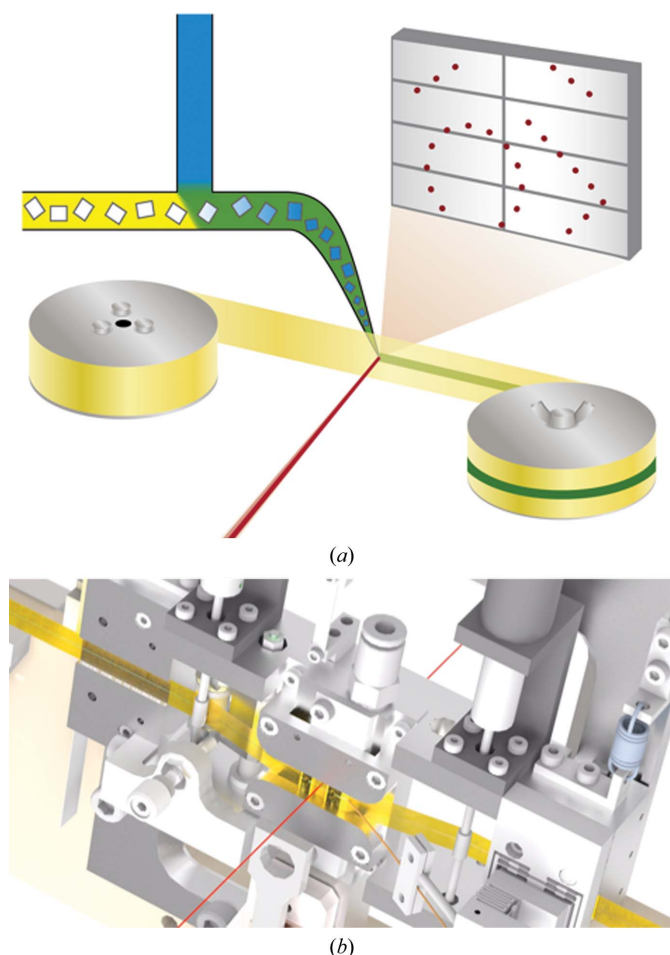


Figure 1

Schematics of the experimental setup. (a) The native protein crystal suspension (yellow) is mixed inline with the substrate solution (blue) before being deposited onto the detector side of the tape. The X-ray beam (red) is focused on the centre of the polyimide tape (yellow ribbon), which is being drawn from the feeder roll on the left to the collector roll on the right. A diffraction pattern measured in transmission is illustrated as red dots on the grey detector. (b) A technical drawing of the tape-drive device shows how tape is held under tension between two rollers near the interaction region and that X-rays pass through a hole in the centre of the device.

vertically confined by grooves in the tape-drive body that matched the width of the tape, as shown in Fig. 1(b). A photograph of the setup during the experiment can be seen in Supplementary Fig. S1. Our tape-drive design was derived from a previous device used at the National Synchrotron Light Source (Roessler *et al.*, 2013) and LCLS (Roessler *et al.*, 2016), in which a tape conveyor belt transported acoustically ejected protein crystal suspension droplets of $10\text{--}100 \mu\text{m}$ in diameter to a microfocused X-ray beam. In contrast to this device, our tape was oriented such that the X-ray beam was incident normal to the tape. Using this configuration, we had a combined thickness of tape and sample solution of only around $50 \mu\text{m}$, resulting in a low scattered X-ray background level.

The microcrystal suspension was placed in a reservoir that was attached to a motor programmed to rotate back and forth by 180° to prevent the protein crystals from settling during measurement. One side of the reservoir was connected to a water supply that was pressurized by an Elveflow OB1 flow controller to hydraulically actuate a plunger in the reservoir. The other side was connected to a borosilicate capillary with an internal diameter of $100 \mu\text{m}$ (Polymicro, USA), through which the microcrystal suspension flowed to a mixer before being deposited onto the tape. From a second reservoir, the aqueous ligand solution was also delivered to the mixer. Equal flow rates of microcrystal suspension and ligand solution of $0.6 \mu\text{l min}^{-1}$ were kept constant throughout the experiment, ensuring a 1:1 mixing ratio.

Two different mixing times, 2 and 50 s, were measured using different mixer configurations based on a three-port microfluidic T-junction (Upchurch). For the 50 s mixing time, mixing was performed by flowing crystal suspension and ligand

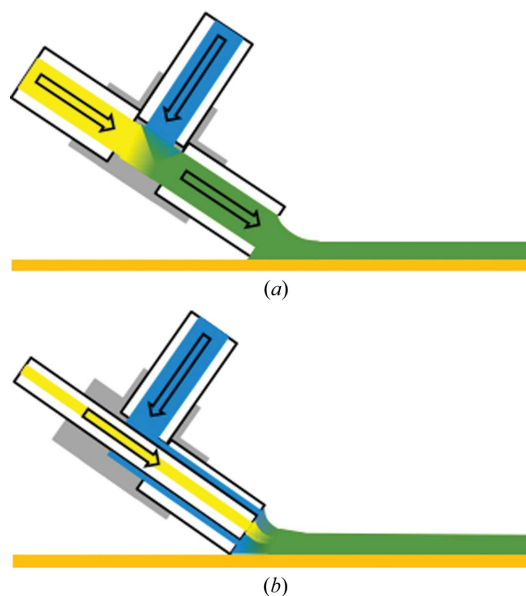


Figure 2

The mixer configurations for the (a) 50 s and (b) 2 s mixing times are illustrated, showing a side view of the sample-stream deposition on the tape (orange line). The colour scheme for the respective liquids is the same as in the experimental setup schematic (Fig. 1a). The images are not to scale.

solution into the T-junction and connecting an outlet capillary that transports the mixed solution to the tape (Fig. 2*a*). The length of the outlet capillary was 120 mm, resulting in a 48 s transport time of the mixed solutions to the tape surface. The tip of the sample-delivery capillary contacted the tape about 1 mm away from the X-ray focus. The translation time of the crystals on the tape was then about 2 s, resulting in a total mixing time of 50 s. For the 2 s mixing time, the protein crystal and ligand solutions were mixed at the capillary–tape contact point. This was accomplished by feeding the crystal suspension capillary (inner diameter 75 μm , outer diameter 150 μm) through the T-junction and into a square capillary with an internal edge length of 150 μm (Fig. 2*b*). The ligand solution then flowed around the crystal suspension capillary through the remaining space in the square capillary to mix with the protein solution on the tape. The tip of this concentric transport capillary was also positioned around 1 mm away from the X-ray focus, resulting in a mixing time of around 2 s.

Prior to the experiment, a series of tests were performed to understand the important parameters that determine the sample-stream width and thickness. Here, the width refers to the distance which the stream laterally spreads on the tape, while the thickness refers to the height of the stream normal to the tape surface. Assuming a micrometre focus and a frame rate of 25 Hz, tape speeds in the range of millimetres per second are sufficient to ensure that fresh sample is transported into the focus between frames. Capillaries of different sizes and contact angles were tested, but it was found that the sample-stream width for this slow tape speed was primarily determined by the sample flow rate and surface-wetting properties. A simple model assuming that the sample flow rate, f , is equal to the tape-translation speed, v , multiplied by the cross-sectional area of the stream was able to fit the dependence of the experimentally observed stream width on the flow rate and tape speed. Assuming a rectangular stream cross-section, the thickness of the sample stream is then given by $t = f/(vw)$, where w represents the stream width. Using this expression and the observed width, streams of the lysozyme crystal suspension on untreated polyimide tape were found to keep an aspect ratio (t/w) of 0.08, consistent with a constant wetting angle. The total flow rate of 1.2 $\mu\text{l min}^{-1}$ and tape speed of 0.6 mm s^{-1} used during the experiment resulted in a stream width of 500 μm measured by an inline microscope, which corresponds to a thickness of 40 μm .

The average time necessary for a ligand molecule to diffuse to a lysozyme binding site was substantially different for each of the mixer configurations. The diffusion time of a ligand molecule in solution, t_{D} , can be calculated from the average diffusion distance, x , and diffusion coefficient, D , according to $t_{\text{D}} = \langle x \rangle^2/2D$. The diffusion coefficient of chitotriose in water has been measured to be $3.5 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ (Groves *et al.*, 2004). Considering the 50 s mixing case, mixing was performed in the T-junction, so the longest diffusion distance is approximately the capillary internal diameter of 100 μm . This results in a maximum time of 14 s for chitotriose to diffuse in solution for this mixer. For the 2 s case, mixing is performed on the tape, where the concentric geometry of the capillaries

should result in a layering of the solutions, as illustrated in Fig. 2(*b*). An estimate of the diffusion time is then found from half of the thickness of the protein crystal suspension layer, as the ligand is approaching the centre from all sides. We assume that the wicking and spreading of the solutions onto the tape quickly results in uniform layers. In this case, the relative thicknesses of the layers are given by the flow rates. Using the total stream thickness of 40 μm , and assuming ligand:crystal:ligand solution layer thicknesses of 10:20:10 μm , a maximum solution diffusion time of 150 ms is then found for this case. The diffusion time of chitotriose from the surface to the centre of a 10 μm lysozyme microcrystal should be around 5 ms (Schmidt, 2013). Therefore, for both mixer configurations diffusion through the solution is expected to take the most time. These diffusion times also serve as an estimate in the uncertainty in the starting time of the binding reaction for the two mixer configurations, as crystals at different positions in the stream would be exposed to the ligand molecules at different times in this range. In this regard, mixing on the tape allows a much more precisely defined reaction time.

Data were continuously collected for 8.5 h in each mixing configuration. This resulted in a total consumption of around 300 μl (18.9 mg) of lysozyme for each data set. The incoming data stream was monitored for hits using the *OnDA* software package (Mariani *et al.*, 2016) and the live hit rate was displayed for fast feedback during the experiment. An offline version of *OnDA* called *offDA* (Mariani *et al.*, 2016) was used to identify individual ‘hits’ from the complete set of collected diffraction patterns and convert them to HDF5 format. The detected hits were then indexed, integrated and merged in point group *4/mmm* to a resolution cutoff of 1.70 \AA using the *CrystFEL* analysis software for serial crystallographic data (White *et al.*, 2012, 2013, 2016; White, 2014).

MTZ files for crystallographic data processing were generated from *CrystFEL* merged reflection datafiles using *F2MTZ* from the *CCP4* suite (Winn *et al.*, 2011). Figures of merit were calculated using *compare_hkl* (R_{split} , $\text{CC}_{1/2}$ and CC^*) and *check_hkl* (SNR, multiplicity and completeness), which are both part of *CrystFEL*.

The lysozyme structure with PDB code 4et8 (Boutet *et al.*, 2012) was used as the starting model for both mixing cases. Owing to non-isomorphism of the collected data sets with that of 4et8, R_{free} flags were generated randomly using *phenix.refine* (Afonine *et al.*, 2012), and the same set of R_{free} flags was used for both data sets. Initial refinement was carried out using *phenix.refine*, with all isotropic atomic displacement parameters (ADPs) set to 20 \AA^2 and using simulated annealing. Ligands were then automatically placed in the initial $F_o - F_c$ map using *LigandFit* (Terwilliger *et al.*, 2006, 2007) in *PHENIX* (Adams *et al.*, 2010), giving only the three-letter code CTO for *N,N',N''*-triacylchitotriose as additional input. Restraints for CTO were generated using *ReadySet!* (Moriarty *et al.*, 2009) in *PHENIX*, followed by iterative cycles of restrained maximum-likelihood refinement using *phenix.refine* and manual model rebuilding using *Coot* (Emsley *et al.*, 2010). CTO grouped occupancy was refined in *phenix.refine*. After refinement, the isotropic ADPs of the tightest bound sugar

Table 1
Data-collection and refinement statistics for CTO bound to lysozyme.

Values in parentheses are for the highest resolution shell.

(a) Data collection.

	50 s	2 s
Temperature (K)	293	293
Crystal size (μm)	6–8	6–8
No. of collected images	527453	611182
Total measuring time (s)	30240	30480
Average acquisition rate (frames s^{-1})	25	25
No. of hits	169021	205181
Indexed patterns	142265	147407
Space group	$P4_32_12$	$P4_32_12$
Unit-cell parameters		
$a = b$ (\AA)	79.61	79.61
c (\AA)	38.32	38.32
$\alpha = \beta = \gamma$ ($^\circ$)	90	90
Resolution (\AA)	22.68–1.70 (1.761–1.70)	22.68–1.70 (1.761–1.70)
$\langle I/\sigma(I) \rangle$	16.65 (1.06)	14.53 (0.56)
Completeness (%)	100 (100)	100 (100)
Multiplicity	4932 (267)	3743 (165)
R_{split}	3.49 (112.45)	4.06 (215.9)
$\text{CC}_{1/2}$	0.99 (0.39)	0.99 (0.14)
CC^*	0.99 (0.75)	0.99 (0.50)
Wilson B factor (\AA^2)	25.77	29.12

(b) Refinement.

	50 s		2 s	
	4et8 phased	1hew phased	4et8 phased	1hew phased
PDB code	5njr	5njs	5njq	5njp
Resolution (\AA)	22.68–1.70 (1.76–1.70)	22.68–1.70 (1.76–1.70)	22.68–1.70 (1.76–1.70)	21.81–1.70 (1.76–1.70)
No. of reflections	14040 (1352)	14040 (1352)	14017 (1339)	14038 (1339)
Reflections used for R_{free}	1403 (136)	1403 (136)	1401 (134)	1401 (134)
R_{work}	0.1661 (0.3175)	0.1641 (0.3185)	0.1631 (0.3644)	0.1702 (0.3615)
R_{free}	0.1975 (0.3670)	0.1988 (0.3696)	0.2003 (0.3596)	0.2009 (0.3524)
No. of atoms				
Protein	1046	1001	1033	1009
Ligand/ion	47	46	46	46
Water	59	60	82	60
B factors (\AA^2)				
Protein	33.13	34.96	35.66	37.69
Ligand/ion	54.47	58.59	52.73	52.90
Water	51.37	41.37	45.58	44.43
Ramachandran favoured (%)	96.85	97.64	98.43	97.64
Ramachandran allowed (%)	3.15	1.57	1.57	2.36
Ramachandran outliers (%)	0	0.79	0	0
R.m.s. deviations				
Bond lengths (\AA)	0.004	0.004	0.005	0.004
Bond angles ($^\circ$)	0.64	0.62	0.70	0.64
Clashscore	0.93	0.00	1.41	2.43

ring of CTO matched the isotropic ADPs of the neighbouring lysozyme residues. Ordered solvent molecules and ions [chloride (Cl^-) and sodium (Na^+)] were placed in *Coot*, followed by further rounds of refinement in *phenix.refine* to convergence. *Polygon* (Urzhumtseva *et al.*, 2009) and *MolProbity* (Chen *et al.*, 2010) were used for validation of the final model and *ValLigURL* (Kleywegt & Harris, 2007) was used to check the geometry of the refined CTO molecule.

To assess the influence of the starting lysozyme model, structure refinement was also carried out in both mixing cases using *phenix.refine* starting with the structure of lysozyme co-crystallized with N,N',N'' -triacetylchitotriose (PDB entry

1hew; Cheetham *et al.*, 1992). Before starting refinement, all isotropic ADPs were reset to 20\AA^2 , but no further modification of the structure was made as the ordered solvent and NAG-NAG-NAG bound to the active centre were unchanged. After initial refinement (rigid body, xyz coordinates, isotropic ADP, simulated annealing and real space), the resulting model and maps were inspected using *Coot*, followed by iterative cycles of restrained maximum-likelihood refinement using *phenix.refine* and manual model rebuilding using *Coot*. After the addition of riding H atoms with *ReadySet!* in *PHENIX*, the structural model was edited manually in *Coot* to remove the automatically placed, but chemically incorrect, H atom from the glycosidic O atom between C1 and C4 of the individual NAG residues. *Polygon* and *MolProbity* were used for validation of the final model. Final models from starting models based on both PDB entries 4et8 and 1hew for both mixing cases were compared with each other and with PDB entry 1hew in *PyMOL*.

3. Results

A total of 147 407 and 142 265 indexable diffraction images were collected for the 2 and 50 s mixing times, respectively. This corresponded to 24% (2 s) and 27% (50 s) of the total collected detector exposures ('indexing fraction'), resulting in an effective data-collection rate of 6–7 Hz (see Table 1 for details). Compared with the earlier serial synchrotron experiment by Stellato and coworkers, we have reduced the sample consumption drastically from $6.25 \mu\text{g}$ to 89 ng of protein consumed per indexable detector frame. This was mainly owing to the markedly lower sample flow rate of 600 nl min^{-1} , combined with improvements in the data-analysis pipeline.

The electron-density maps obtained starting from the ligand-free model (PDB entry 4et8) described previously are shown in Fig. 3. In both mixing cases, the $2F_o - F_c$ and $F_o - F_c$ maps reveal the presence of a large molecule near the active site. Phasing with a starting model containing 1→4 linked NAG-NAG-NAG (PDB entry 1hew) shows that the mixed-in ligand bound to subsites *A*, *B* and *C* in both cases (see Fig. 4 and Supplementary Fig. S2), which agrees with previous crystallographic (Cheetham *et al.*, 1992) and NMR (Raftery *et al.*, 1969) studies.

To minimize model bias, we based our refinement on the initial electron-density maps (see Figs. 3*a* and 3*d*) obtained from the ligand-free starting model PDB entry 4et8 (Boutet *et al.*, 2012). Using the $F_o - F_c$ map, *LigandFit* from *PHENIX* (Terwilliger *et al.*, 2006, 2007) placed one molecule of CTO in the active site (see Figs. 3*b* and 3*e*). In both cases, the automatically placed orientation of the molecule agreed well with the known bound CTO positions in lysozyme (Cheetham *et al.*, 1992). Inspection of the map and model revealed only slight differences between the mixing cases, showing that a mixing time of only about 2 s is sufficient to populate 90% of the lysozyme active sites, found from grouped occupancy refinement. As reported in the literature (Cheetham *et al.*, 1992), the isotropic atomic displacement parameters of the atoms in the three NAG rings increased from the most tightly bound ring (close to binding site *C*) to the ring bound to binding site *A*. For the assessment of occupancy, we thus followed the rule that the isotropic ADP of the tightest bound NAG should correspond to that of the surrounding (binding) amino-acid residues. This resulted in CTO occupancies of 0.91 and 0.97 for the 2 and 50 s mixing times, respectively. The occupancy for the 50 s mixing time corresponds to that found in a bound single-crystal study (Cheetham *et al.*, 1992). Final rounds of refinement were carried out with these CTO occupancies; the resulting figures of merit are given in Table 1 and final maps and structural models can be seen in Figs. 3(*c*) and 3(*f*).

The overall root-mean-square deviation (r.m.s.d.) between the final refined models from both mixing times was 0.056 Å after alignment with *PyMOL*. The r.m.s.d.s between the starting model and final refined models were 0.152 Å (2 s) and 0.151 Å (50 s). Structural alignment with the model derived from co-crystallization of HEWL with CTO (PDB entry 1hew) gave r.m.s.d. values of 0.185 Å (2 s) and 0.189 Å (50 s).

Analysis of the hydrogen bonds between chitotriose and lysozyme using *PyMOL*, *PDBsum* (Laskowski, 2009) and *LigPlot+* (Laskowski & Swindells, 2011) revealed that substrate binding occurred in both mixing cases to binding subsites *A*, *B* and *C*, in agreement with a previous study (Cheetham *et al.*, 1992). A comparison of the 50 s structural model with the 1hew model is shown in Fig. 4, while a comparison of the 2 and 50 s mixing cases is shown in Supplementary Fig. S2. In both structures sugar ring 1, located in binding subsite *A*, binds to lysozyme *via* a hydrogen bond between OD2 of Asp101 and the acetamido N atom in ring 1 (see Supplementary Figs. S3 and S4), while in the corresponding 1hew structure NAG1 binds to Asn103 through a hydrogen bond between OD1 and the hydroxyl O atom at position 6 of the pyranose ring and Asp101 only forms a hydrogen bond to NAG2. This corresponds to an almost identical bond between Asp101 in the 2 and 50 s mixing cases and the same hydroxyl O atom in ring 2 of CTO. The slight difference is that in the 2 s mixing case Asn103 is connected to

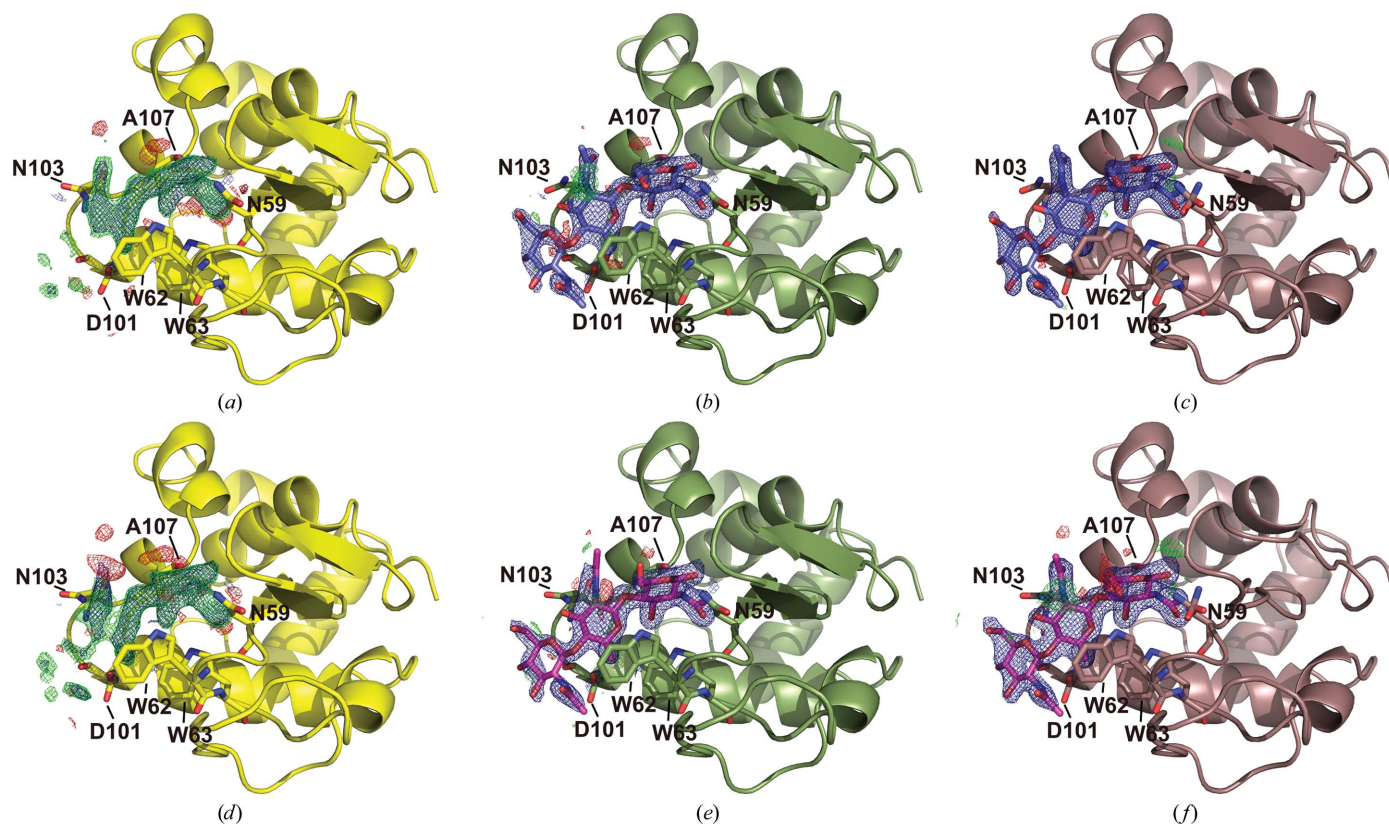


Figure 3

CTO bound to lysozyme: short-mixing case (*a*, *b*, *c*) and long-mixing case (*d*, *e*, *f*). (*a*) and (*d*) show the $2F_o - F_c$ (blue, 1σ) and $F_o - F_c$ (green and red, 2.5σ) maps of the lysozyme binding site after initial refinement with ligand-free phases. (*b*) and (*e*) show the maps at the same levels after automatic ligand-placement and refinement, and (*c*) and (*f*) show the respective final refined models and maps. Residues forming hydrogen bonds to CTO are shown.

CTO *via* a bound water molecule, whilst in the 50 s mixing case this water is missing and Asn103 only interacts with CTO through van der Waals forces.

Analysis of the extended solvent network between the bound ligand and lysozyme shows that for the 50 s mixing time four ordered solvent molecules are involved in binding (see Supplementary Fig. S3): three of these (O148, O170 and O180) directly form hydrogen bonds to CTO, while the fourth (O150) is connected to CTO through a hydrogen bond to O180. This fourth water stabilizes substrate binding through hydrogen bonds to Ile98, Asn103 and Gly104. O170 connects CTO to Gln57 and one of the alternate conformers of Asp52; together with Glu35 the latter forms the catalytic site of HEWL. O148 is connected to O170 and Ala107, both of which form hydrogen bonds to CTO and to Val109, which is part of binding subsite *D* (see Fig. 3). The water network of the 2 s case is far more extended (see Supplementary Fig. 4), especially along the cleft that forms the catalytic site. In the 50 s case only two waters adhere directly or indirectly (through other water molecules) to CTO along this cleft, whilst a total of eight water molecules (O151, O152, O166, O167, O171, O174, O194 and O210) form the extended solvent network in

the 2 s case (see Supplementary Fig. 4). These ordered solvent molecules connect CTO to residues Phe34, Glu35, Asp52, Gln57, Ala107 and Val109, which are all part of the sugar-binding cleft (subsites *C*, *D* and *E*). Three additional solvent molecules connect CTO to residues Ile98, Asp101, Asn103 and Gly104, such that a total of 11 water molecules are involved in CTO binding in the 2 s case, as opposed to four water molecules in the 50 s case.

Contrary to our findings, a powder diffraction study found that CTO preferentially binds to subsites *B*, *C* and *D* (*BCD*; Von Dreele, 2005). The authors attributed the difference from previous crystallographic studies to the fact that CTO bound to positions *A*, *B* and *C* (*ABC*). Thus, only CTO bound to subsites *ABC* is visible in crystallographic studies because these experiments take much longer from crystallization to data collection. In our experiment, we used a microcrystalline slurry and the time between mixing in CTO to probing the sample with X-rays was only a few seconds, as opposed to several minutes in the powder diffraction study. If binding of CTO to the *BCD* sites were indeed faster than to the *ABC* sites, we would expect to observe a mixture of both binding

modes. As already discussed, we observed a clear preferred binding of CTO to subsites *ABC*. The $2mF_o - DF_c$ and $mF_o - DF_c$ maps shown in Supplementary Fig. S5 contain only weak features at the *BCD* subsites at σ levels below 0.6 and 2.5, respectively. Furthermore, refinement starting from the powder diffraction structure (PDB entry 1sf6; Von Dreele, 2005) results in a model with a clear preference for the *ABC* binding mode. Supplementary Fig. S5 shows positive difference electron density at the position of sugar ring 1 of CTO from our model (bound to subsite *A*) and almost no electron density around the NAG molecule that would correspond to binding to subsite *D*.

4. Conclusion

We have demonstrated that serial synchrotron crystallography can be used for mix-and-diffuse studies of ligand binding just 2 s after the ligand is mixed inline with the protein microcrystal suspension. The small lysozyme crystal size in our experiment resulted in diffusion times that were sufficiently short that nearly full occupancy of the bound substrate could be achieved in this time. This overcomes many technological obstacles to native protein drug screening, and also opens up the possibility of investigating enzyme

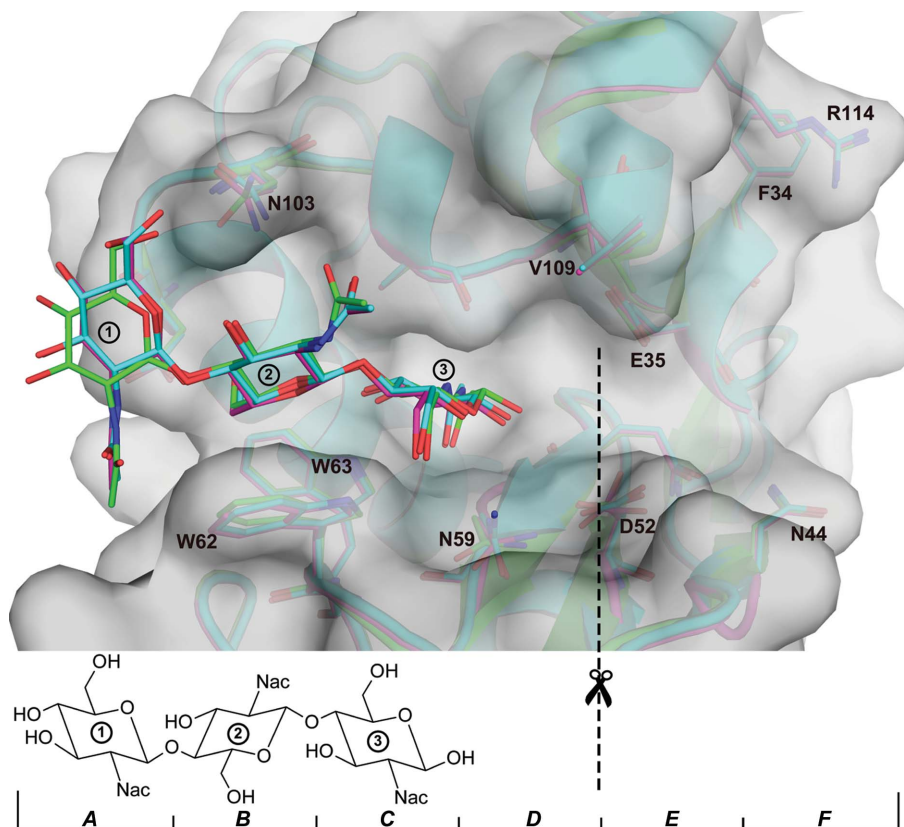


Figure 4
Comparison of ligand binding in the long-mixing case solved starting from PDB entries 4et8 (magenta) and 1hew (cyan) with that in the co-crystallized ligand structure PDB entry 1hew (green). It can be seen that all ligands bind to subsites *ABC* and that only the most flexible sugar ring 1 shows a significantly different orientation. Most notably, the different orientation of the hydroxyl O atom at position 6 of the pyranose ring and of the acetamido group of Asn103 results in a hydrogen bond between the ligand and protein in PDB entry 1hew but not in the mixing cases. Binding-site residues of lysozyme are displayed and the associated binding subsites *A–F* are indicated at the bottom of the figure. The dashed line indicates the active site for catalysis (cleavage site).

structural dynamics on a similar timescale. Further advances that may reduce the total collection time of a data set to seconds are on the horizon. Detectors are becoming available that are able to collect thousands of frames per second (Henrich *et al.*, 2011; Pennicard *et al.*, 2013; Redford *et al.*, 2016), which could be combined with higher speed crystal delivery at brighter diffraction-limited storage rings (Eriksson *et al.*, 2014). Recently, it was shown that by using the high flux of nonmonochromatic ‘pink-beam’ undulator radiation at a synchrotron facility, single patterns from protein microcrystals could be acquired from 100 ps X-ray pulses arising from single electron bunches (Meents *et al.*, 2017). The advantage of using the broader bandwidth was demonstrated by the requirement for only 50 such exposures to generate a full diffraction data set.

Acknowledgements

The authors would like to thank Rachel Beyerlein and Jelke Wibbeke for assistance with the illustrations. Support for AO was provided by the DESY Summer Student Program. The work of the authors affiliated with the Center for Free-Electron Laser Science was funded by the Helmholtz Association through programme-oriented funds.

Funding information

This work was supported by the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) through the Consolidator Grant COMOTION (ERC-614507-Küpper), by the excellence cluster ‘The Hamburg Center for Ultrafast Imaging – Structure, Dynamics and Control of Matter at the Atomic Scale’ of the Deutsche Forschungsgemeinschaft (CUI, DFG-EXC1074) and by the Helmholtz Gemeinschaft through the ‘Impuls- und Vernetzungsfond’.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
 Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* **D68**, 352–367.
 Barends, T. R. M. *et al.* (2015). *Science*, **350**, 445–450.
 Barends, T. R. M., Foucar, L., Botha, S., Doak, R. B., Shoeman, R. L., Nass, K., Koglin, J. E., Williams, G. J., Boutet, S., Messerschmidt, M. & Schlichting, I. (2014). *Nature (London)*, **505**, 244–247.
 Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967). *Proc. R. Soc. B Biol. Sci.* **167**, 378–388.
 Blundell, T. L. (2017). *IUCrJ*, **4**, 308–321.
 Botha, S., Nass, K., Barends, T. R. M., Kabsch, W., Latz, B., Dworkowski, F., Foucar, L., Panepucci, E., Wang, M., Shoeman, R. L., Schlichting, I. & Doak, R. B. (2015). *Acta Cryst.* **D71**, 387–397.
 Boutet, S. *et al.* (2012). *Science*, **337**, 362–364.
 Burkhardt, A., Pakendorf, T., Reime, B., Meyer, J., Fischer, P., Stübe, N., Panneerselvam, S., Lorbeer, O., Stachnik, K., Warmer, M., Rödig, P., Göries, D. & Meents, A. (2016). *Eur. Phys. J. Plus*, **131**, 56.
 Calvey, G. D., Katz, A. M., Schaffer, C. B. & Pollack, L. (2016). *Struct. Dyn.* **3**, 054301.
 Chapman, H. N. *et al.* (2011). *Nature (London)*, **470**, 73–77.

Cheatham, J. C., Artymiuk, P. J. & Phillips, D. C. (1992). *J. Mol. Biol.* **224**, 613–628.
 Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
 Colletier, J.-P. *et al.* (2016). *Nature (London)*, **539**, 43–47.
 Congreve, M., Murray, C. W. & Blundell, T. L. (2005). *Drug Discov. Today*, **10**, 895–907.
 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
 Eriksson, M., van der Veen, J. F. & Quitmann, C. (2014). *J. Synchrotron Rad.* **21**, 837–842.
 Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. (2016). *Nat. Rev. Drug Discov.* **15**, 605–619.
 Gati, C. *et al.* (2017). *Proc. Natl Acad. Sci. USA*, **114**, 2247–2252.
 Groves, P., Rasmussen, M. O., Molero, M. D., Samain, E., Cañada, F. J., Driguez, H. & Jiménez-Barbero, J. (2004). *Glycobiology*, **14**, 451–456.
 Hadfield, A. T., Harvey, D. J., Archer, D. B., MacKenzie, D. A., Jeenes, D. J., Radford, S. E., Lowe, G., Dobson, C. M. & Johnson, L. N. (1994). *J. Mol. Biol.* **243**, 856–872.
 Hajdu, J., Acharya, K. R., Stuart, D. I., McLaughlin, P. J., Barford, D., Oikonomakos, N. G., Klein, H. & Johnson, L. N. (1987). *EMBO J.* **6**, 539–546.
 Henrich, B. *et al.* (2011). *Nucl. Instrum. Methods Phys. Res. A*, **633**, S11–S14.
 Johnson, L. N. (1998). *Nat. Struct. Biol.* **5**, 942–944.
 Kang, Y. *et al.* (2015). *Nature (London)*, **523**, 561–567.
 Kleywegt, G. J. & Harris, M. R. (2007). *Acta Cryst.* **D63**, 935–938.
 Kupitz, C. *et al.* (2017). *Struct. Dyn.* **4**, 044003.
 Laskowski, R. A. (2009). *Nucleic Acids Res.* **37**, D355–D359.
 Laskowski, R. A. & Swindells, M. B. (2011). *J. Chem. Inf. Model.* **51**, 2778–2786.
 Liu, W. *et al.* (2013). *Science*, **342**, 1521–1524.
 Maenaka, K., Matsushima, M., Song, H., Sunada, F., Watanabe, K. & Kumagai, I. (1995). *J. Mol. Biol.* **247**, 281–293.
 Mariani, V., Morgan, A., Yoon, C. H., Lane, T. J., White, T. A., O’Grady, C., Kuhn, M., Aplin, S., Koglin, J., Barty, A. & Chapman, H. N. (2016). *J. Appl. Cryst.* **49**, 1073–1080.
 Martin-Garcia, J. M. *et al.* (2017). *IUCrJ*, **4**, 439–454.
 Meents, A. (2017). *Nat. Commun.* doi:10.1038/s41467-017-01417-3.
 Moffat, K. (2014). *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20130568.
 Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Cryst.* **D65**, 1074–1080.
 Nango, E. *et al.* (2016). *Science*, **354**, 1552–1557.
 Neutze, R. & Moffat, K. (2012). *Curr. Opin. Struct. Biol.* **22**, 651–659.
 Nogly, P. *et al.* (2015). *IUCrJ*, **2**, 168–176.
 Oberthuer, D. *et al.* (2017). *Sci. Rep.* **7**, 44628.
 Oswald, C., Rappas, M., Kean, J., Doré, A. S., Errey, J. C., Bennett, K., Deflorian, F., Christopher, J. A., Jazayeri, A., Mason, J. S., Congreve, M., Cooke, R. M. & Marshall, F. H. (2016). *Nature (London)*, **540**, 462–465.
 Pande, K. *et al.* (2016). *Science*, **352**, 725–729.
 Parang, K. & Sun, G. (2004). *Curr. Opin. Drug Discov. Devel.* **7**, 617–629.
 Patel, D., Bauman, J. D. & Arnold, E. (2014). *Prog. Biophys. Mol. Biol.* **116**, 92–100.
 Pennicard, D., Lange, S., Smoljanin, S., Hirsemann, H., Graafsma, H., Epple, M., Zuvic, M., Lampert, M.-O., Fritzsche, T. & Rothermund, M. (2013). *J. Phys. Conf. Ser.* **425**, 062010.
 Raftery, M. A., Dahlquist, F. W., Parsons, S. M. & Wolcott, R. G. (1969). *Proc. Natl Acad. Sci. USA*, **62**, 44–51.
 Redford, S. *et al.* (2016). *J. Instrum.* **11**, C11013.
 Roessler, C. G. *et al.* (2016). *Structure*, **24**, 631–640.
 Roessler, C. G., Kuczewski, A., Stearns, R., Ellson, R., Olechno, J., Orville, A. M., Allaire, M., Soares, A. S. & Héroux, A. (2013). *J. Synchrotron Rad.* **20**, 805–808.

- Schiebel, J., Radeva, N., Krimmer, S. G., Wang, X., Stieler, M., Ehrmann, F. R., Fu, K., Metz, A., Huschmann, F. U., Weiss, M. S., Mueller, U., Heine, A. & Klebe, G. (2016). *ACS Chem. Biol.* **11**, 1693–1701.
- Schlichting, I. & Goody, R. S. (1997). *Methods Enzymol.* **277**, 467–490.
- Schmidt, M. (2013). *Adv. Condens. Matter Phys.* **2013**, 1–10.
- Stagno, J. R. *et al.* (2016). *Nature (London)*, **541**, 242–246.
- Stellato, F. *et al.* (2014). *IUCrJ*, **1**, 204–212.
- Stoddard, B. L. (1996). *Pharmacol. Ther.* **70**, 215–256.
- Tenboer, J. *et al.* (2014). *Science*, **346**, 1242–1246.
- Terwilliger, T. C., Adams, P. D., Moriarty, N. W. & Cohn, J. D. (2007). *Acta Cryst.* **D63**, 101–107.
- Terwilliger, T. C., Klei, H., Adams, P. D., Moriarty, N. W. & Cohn, J. D. (2006). *Acta Cryst.* **D62**, 915–922.
- Urzhumtseva, L., Afonine, P. V., Adams, P. D. & Urzhumtsev, A. (2009). *Acta Cryst.* **D65**, 297–300.
- Varghese, J. N. (1999). *Drug Dev. Res.* **46**, 176–196.
- Vocadlo, D. J., Davies, G. J., Laine, R. & Withers, S. G. (2001). *Nature (London)*, **412**, 835–838.
- Von Dreele, R. B. (2005). *Acta Cryst.* **D61**, 22–32.
- White, T. A. (2014). *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20130330.
- White, T. A., Barty, A., Stellato, F., Holton, J. M., Kirian, R. A., Zatsepin, N. A. & Chapman, H. N. (2013). *Acta Cryst.* **D69**, 1231–1240.
- White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A. & Chapman, H. N. (2012). *J. Appl. Cryst.* **45**, 335–341.
- White, T. A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K. R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C. H., Diederichs, K. & Chapman, H. N. (2016). *J. Appl. Cryst.* **49**, 680–689.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yamashita, K. *et al.* (2015). *Sci. Rep.* **5**, 14017.
- Zhang, H. *et al.* (2015). *Cell*, **161**, 833–844.