

High-throughput *in situ* experimental phasingJoshua M. Lawrence,^{a,‡} Julien Orlans,^{a,b} Gwyndaf Evans,^a Allen M. Orville,^{a,c} James Foadi^{a,*}§ and Pierre Aller^{a,*}

^aDiamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, United Kingdom, ^bUMR0203, Biologie Fonctionnelle, Insectes et Interactions (BF2i); Institut National des Sciences Appliquées de Lyon (INSA Lyon); Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE), University of Lyon (Univ Lyon), F-69621 Villeurbanne, France, and ^cResearch Complex at Harwell, Rutherford Appleton Laboratory, Didcot OX11 0FA, United Kingdom. *Correspondence e-mail: j.foadi@bath.ac.uk, pierre.aller@diamond.ac.uk

Received 11 May 2020

Accepted 3 July 2020

Edited by Q. Hao, University of Hong Kong

‡ Current address: Department of Biochemistry, University of Cambridge, Hopkins Building, Downing Site, Cambridge CB2 1QW, United Kingdom.

§ Current address: Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom.

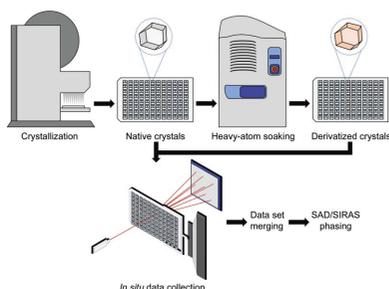
Keywords: *in situ* crystallography; experimental phasing; crystal soaking; high throughput; isomorphism.

Supporting information: this article has supporting information at journals.iucr.org/d

In this article, a new approach to experimental phasing for macromolecular crystallography (MX) at synchrotrons is introduced and described for the first time. It makes use of automated robotics applied to a multi-crystal framework in which human intervention is reduced to a minimum. Hundreds of samples are automatically soaked in heavy-atom solutions, using a Labcyte Inc. Echo 550 Liquid Handler, in a highly controlled and optimized fashion in order to generate derivatized and isomorphous crystals. Partial data sets obtained on MX beamlines using an *in situ* setup for data collection are processed with the aim of producing good-quality anomalous signal leading to successful experimental phasing.

1. Introduction

The exploitation of multiple crystals is becoming common in data collection and analysis at synchrotrons and X-ray free-electron lasers (XFELs); the encompassing methods are often termed serial macromolecular crystallography (SMX). Indeed, new scenarios have first been envisaged and then implemented to increase the probability of obtaining complete and good-quality sets of reflections by combining data from several to thousands of crystals in creative ways. These strategies have been applied in merging partial data sets from membrane proteins and large complexes (Arakawa *et al.*, 2015; Axford *et al.*, 2012, 2015; Huang *et al.*, 2015, 2016; Mylona *et al.*, 2017), in methodologies adopted to reduce radiation-induced global damage (Garman, 2010; Garman & Owen, 2006; Owen *et al.*, 2006, 2014) and in the enhancement of anomalous signal for SAD phasing (Giordano *et al.*, 2012; Liu *et al.*, 2011, 2012, 2013; Rose *et al.*, 2015; Terwilliger *et al.*, 2016). It has been particularly enlightening and encouraging to discover that the impact of radiation-induced alterations is minimized by distributing the dose required to solve high-resolution structures over many samples, even without the need for crystal cryocooling. While a clear advantage of working at 100 K versus room temperature (RT) is to slow the global radiation-induced alterations in the sample, flash-cooling often introduces stress and deformations into the crystals. If, on the one hand, a higher X-ray dose can be applied to crystals at 100 K, thus making it possible to collect complete data sets, then, on the other hand, serious non-isomorphism between samples can limit their utility for phasing with isomorphous derivatives. Typically, RT crystallography is used in the early stage of the crystallization process to screen potential hits *in situ* directly in the crystallization plate (Aller *et al.*, 2015; Axford *et al.*, 2012; Bingel-Erlenmeyer *et al.*, 2011; Douangamath *et al.*, 2013;



OPEN ACCESS

le Maire *et al.*, 2011; Sanchez-Weatherby *et al.*, 2019). However, the technique has been successful in facilitating high-throughput ligand-binding studies (Gelin *et al.*, 2015), in solving viral structures (Axford *et al.*, 2012) and those of membrane proteins (Axford *et al.*, 2015), and in providing more physiologically relevant RT structures of immunoglobulins (Davies *et al.*, 2017). In the present work, we demonstrate that the benefits of RT crystallography can be extended to experimental phasing by exploiting tens to hundreds of isomorphous crystals.

Experimental phasing techniques make use of heavy-atom (HA) derivatives, often in the form of halides, heavy metals or selenomethionyl (SeMet) protein crystals, alongside native data, to allow *de novo* phasing of structures. The technique is most effective if the various crystals involved are all isomorphous to each other (Crick & Magdoff, 1956; Rould, 2007). RT crystals do not experience the potential alterations induced by cryocooling and are therefore, in principle, better suited for experimental phasing. However, owing to limitations on the total X-ray dose, RT methods most often yield incomplete wedges of data from each crystal; thus, it is necessary to merge partial data sets from several crystals into complete data sets. The need to use many different crystals also introduces the significant possibility of non-isomorphism. To overcome this limitation, the *BLEND* program (Foadi *et al.*, 2013) can select several groups of isomorphous crystals and then scale and merge the related intensities into a single, complete reflection file. Data sets are grouped with hierarchical cluster analysis, using unit-cell parameters as a critical descriptor.

It is, of course, of the utmost importance to have tens to hundreds of isomorphous crystals at hand for SMX methods. Farley *et al.* (2014) have convincingly argued that the lack of reproducibility of unit-cell parameters is mostly caused by the manual handling associated with crystal preparation and HA soaks. Furthermore, micropipettes dispensing sub-microlitre volumes of HA solution imply a certain degree of experimental error which changes the concentration of HAs, thus potentially causing the distortion of unit-cell parameters during the process (Garman & Murray, 2003; Pike *et al.*, 2016; Lu & Sun, 2014). In this study, we have developed a high-throughput, robot-assisted workflow using a robotic liquid dispenser (TTP Labtech Mosquito Crystal) and an acoustic droplet ejector (Labcyte Inc. Echo 550 Liquid Handler) for the production of HA-derivatized isocrystals for *in situ* data collection (Newman *et al.*, 2005; Villaseñor *et al.*, 2012; Chilingaryan *et al.*, 2012; Fig. 1). This workflow has been applied to test crystals of lysozyme and proteinase K soaked in four different halides and four different heavy-metal solutions. A comparison with control data sets from crystals grown and derivatized using the same method but cryocooled for data collection at 100 K offers an effective way to determine the penalty of the lack of isomorphism caused by manual handling and cryocooling of samples. The overall workflow allows the *de novo* solution of protein crystal structures using SAD and SIRAS phasing on a shorter timescale than that for standard cryocooling experiments (Pape & Schneider, 2004; Sheldrick, 2010; Taylor, 2010). This procedure can be achieved at the new

VMXi beamline at Diamond Light Source (DLS), which is dedicated to *in situ* measurements (Sanchez-Weatherby *et al.*, 2019).

2. Materials and methods

2.1. Sample preparation

In order to ensure reproducibility and reduce human error, all sample crystallizations were handled using a TTP Labtech Mosquito Crystal with low-background 96-well Greiner CrystalQuick X crystallization plates optimized for *in situ* data collection (Aller *et al.*, 2015; Axford *et al.*, 2012, 2016).

2.1.1. Lysozyme crystallization. Commercial lyophilized lysozyme from hen egg white (Sigma–Aldrich) was resuspended in Milli-Q water to a concentration of 100 mg ml⁻¹. Sitting drops were dispensed with the TTP Labtech Mosquito Crystal by mixing 100 nl lysozyme solution and 100 nl reservoir solution (1.0 M NaCl, 50 mM sodium acetate pH 4.5), and were equilibrated against 40 µl reservoir solution at 20°C. Crystals (200 × 200 × 50 µm) grew overnight and were found to belong to space group *P*₄₃₂₁₂, with unit-cell parameters $a = b \simeq 79$, $c \simeq 38$ Å.

2.1.2. Proteinase K crystallization. Commercial lyophilized proteinase K (Sigma–Aldrich) was resuspended at 15 mg ml⁻¹ in 50 mM HEPES pH 7.0. Sitting drops were dispensed with the TTP Labtech Mosquito Crystal by mixing 100 nl proteinase K solution and 100 nl reservoir solution (1.2 M ammonium sulfate, 0.1 M Tris–HCl pH 8.0), and were equilibrated against 40 µl reservoir solution at 20°C. Microcrystals (10 × 10 × 10 µm) grew overnight and were found to belong to space group *P*₄₃₂₁₂, with unit-cell parameters $a = b \simeq 68$, $c \simeq 103$ Å.

2.2. Crystal soaking

Two crystallization wells per reservoir containing crystals were automatically soaked with a variety of halide and heavy-metal solutions using a Labcyte Inc. Echo 550 Liquid Handler in a similar way to its use by the XChem facility at DLS to inject structural fragments (Collins *et al.*, 2017). The liquid handler does not use pipette tips to transfer the soaking solution but uses a piezo-electric transducer, which generates a focused acoustic wave on the meniscus of the soaking solution to produce a 2.5 nl droplet (Hadimioglu *et al.*, 2016) directed on the crystal-containing drop. The Echo liquid handler allows highly reproducible and accurate crystal soaking for a high-throughput platform.

The halide solutions used were sodium bromide (NaBr), potassium bromide (KBr), sodium iodide (NaI) and potassium iodide (KI), all of which were purchased from Sigma–Aldrich. Stocks of 2 M halide solutions were prepared by dissolving the halide in the crystallization mother liquor. Halides tend to bind to patches on the surface of proteins through ionic interactions rather than through specific interactions with residue side chains (Pike *et al.*, 2016; Lu & Sun, 2014).

The heavy-metal solutions used were potassium tetranitroplatinate(II) [K₂Pt(NO₂)₄], potassium tetrachloroaurate (KAuCl₄·H₂O), potassium hexachloroiridate (K₂IrCl₆) and

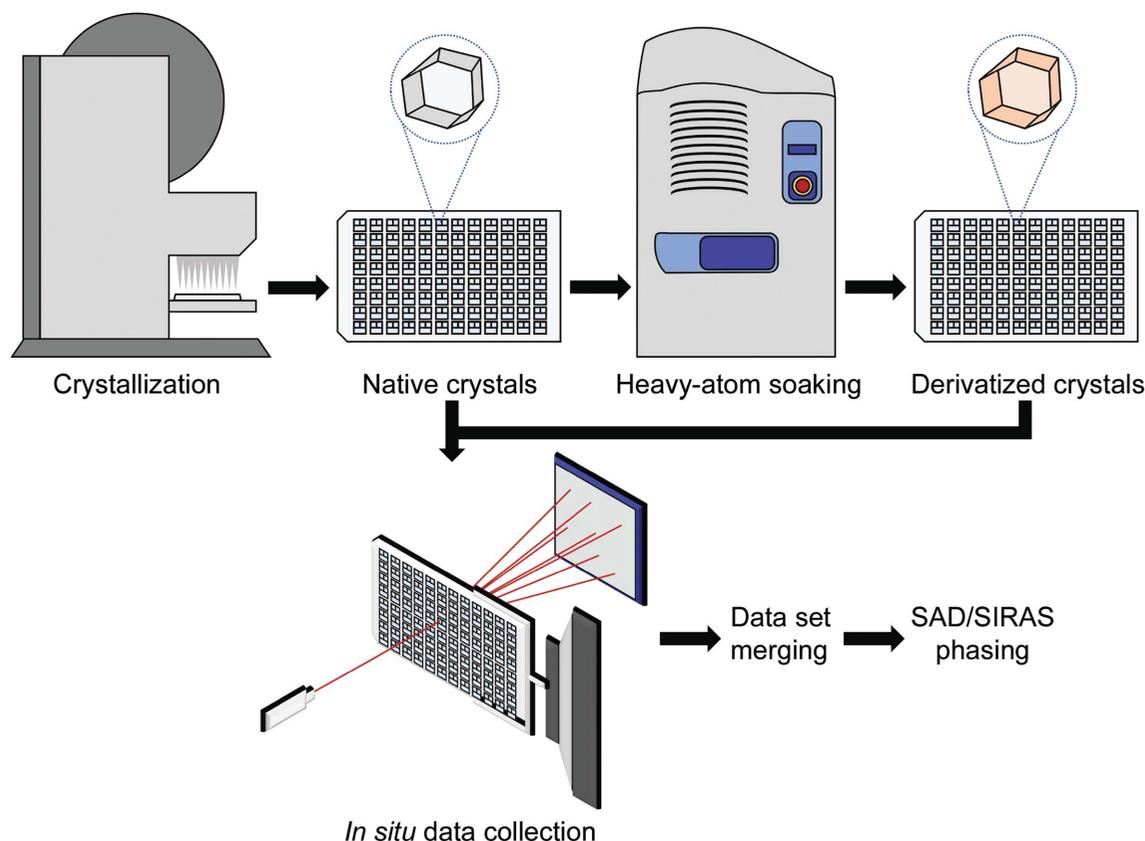


Figure 1
High-throughput workflow from crystallization to phasing. A video showing the workflow is available at <https://www.diamond.ac.uk/Instruments/Mx/XFEL-Hub/Staff/Aller/Research.html>.

samarium(III) chloride hexahydrate ($\text{SmCl}_3 \cdot 6\text{H}_2\text{O}$), all of which were purchased from Hampton Research. Stocks of heavy-metal solutions at 100 mM were prepared by dissolving the heavy-metal compounds in the crystallization mother liquor. This combination of HAs were chosen based on their high chance of success (HA binding), their binding sites and their nontoxicity (Agniswamy *et al.*, 2008; Joyce *et al.*, 2010).

The final soaking concentrations used ranged from a few millimolar for heavy-metal solutions to 50–180 mM for halide solutions. To minimize crystal damage and retain crystal isomorphism, we performed RT short soaks of approximately 15 min for heavy metals and approximately 30 min for halides.

Lysozyme crystals were soaked in halide and heavy-metal solutions with final concentrations of 180 and 9 mM, respectively. $\text{KAuCl}_4 \cdot \text{H}_2\text{O}$ was used at 4.5 mM because at 9 mM the lysozyme crystals start dissolving, likely owing to disruption of the crystal lattice on binding gold. For *in situ* data collection the soaking time varied between the first and the last data set collected. For cryogenic data collection, the soaking time was about 15 min for heavy-metal soaks and 30 min for halide soaks prior to rapidly plunge-cooling the samples in liquid N_2 .

Proteinase K crystals were soaked in halide and heavy-metal solutions with final concentrations of 50 and 2.5 mM, respectively. Half of the concentration was again used for $\text{KAuCl}_4 \cdot \text{H}_2\text{O}$, as a solution of $\text{KAuCl}_4 \cdot \text{H}_2\text{O}$ at 2.5 mM was detrimental to proteinase K crystals. For *in situ* data collection the soaking time varied between the first and the last data set

collected. For cryogenic data collection, the soaking time was about 15 min for heavy-metal soaks and 30 min for halide soaks.

The details of the soaking times and concentrations are shown in Table 1 for lysozyme and in Table 2 for proteinase K. Short soaks at a higher HA concentration were chosen based on the suggestion that they increase the chance of HA binding and yield more isomorphous crystals (Sun *et al.*, 2002). Short soaks also allowed for rapid data collection, ensuring the high-throughput nature of the workflow. Owing to the logistics of RT data collection, soak times for *in situ* samples tended to be slightly longer than for cryocooled samples as the crystals remain soaked in the HA throughout data collection.

We used a lower concentration for the heavy-metal conditions (less than 10 mM), as advised in the literature (Garman & Murray, 2003; Pike *et al.*, 2016). Also, heavy metals absorb large doses of X-rays owing to their greater number of excitable electrons, meaning that high concentrations would impact upon data quality owing to radiation damage (Pike *et al.*, 2016). For proteinase K, a quarter of the volume of each HA solution was used owing to the smaller size of the crystals.

2.3. Data collection

2.3.1. Cryo-data collection. A subset of the crystals (natives and derivatives) were cryocooled in order to probe crystal isomorphism and HA binding at 100 K.

Table 1

Concentration and soaking time of heavy-metal/halide solutions with 200 nl lysozyme crystal drops for data collected at RT.

For cryo data collection, the RT soaking time was about 15 min for heavy-metal soaks and 30 min for halide soaks prior to plunge-cooling in liquid N₂.

HA	Volume ejected (nl)	Stock solution (M)	Final concentration (mM)	RT soaking time (min)
NaBr	20	2	180.00	10–50
KBr	20	2	180.00	10–25
NaI	20	2	180.00	5–25
KI	20	2	180.00	5–20
K ₂ Pt(NO) ₄	20	0.1	9.00	10–50
KAuCl ₄ ·H ₂ O	10	0.1	4.75	5–25
K ₂ IrCl ₆	20	0.1	9.00	5–30
SmCl ₃ ·6H ₂ O	20	0.1	9.00	10–25

Lysozyme. The crystals were cryoprotected by adding ethylene glycol to 25% (v/v) to the crystal-containing droplets before flash-cooling in liquid nitrogen. For the derivatives, a heavy-metal/halide solution was added using the Echo liquid handler before adding the ethylene glycol solution prior to flash-cooling. At least five 360° rotation ‘complete’ data sets were collected from different crystals for the various native and derivative crystals, as shown in Supplementary Table S1. Data sets were collected on the I04 beamline at DLS using a beam size of 32 × 20 μm and a flux of ~1.7–2.5 × 10¹¹ photons per second.

Proteinase K. The crystals did not require cryoprotection before flash-cooling in liquid nitrogen. A slurry of micro-crystals was mounted on a micromesh (10 μm aperture) from MiTeGen after adding heavy-metal/halide solution with the Echo liquid handler for the derivatized crystals. About 20 data sets with 50° rotation were collected for each condition (native and derivatives; Supplementary Table S2). Data collection was performed on the I24 beamline at DLS using a beam size of 9 × 6 μm and a flux of ~2–4 × 10¹¹ photons per second.

2.3.2. In situ data collection. All *in situ* data sets were collected on the I24 beamline at DLS (Aller *et al.*, 2015; Axford *et al.*, 2012, 2015). The CrystalQuick X plates were mounted manually on the horizontal goniometer.

Lysozyme. About 25 ‘incomplete’ data sets of 20° rotation were collected for each condition (Supplementary Table S3) using a beam size of 50 × 50 μm and a flux of about 2 × 10¹¹ photons per second.

Proteinase K. About 50 ‘incomplete’ data sets of 2.5–5° rotation were collected for each condition (Supplementary Table S4) using a beam size of 9 × 6 μm and a flux of about 2 × 10¹¹ photons per second.

The energy used varied depending on the HA condition for the crystals (Supplementary Table S1). Native data were collected at 12 660 eV on beamline I04 and at 12 800 eV on beamline I24. Bromide data were collected at the *K* edge, whereas iodide data were collected at 7000 eV, the closest energy to the iodide *K* edge that could be reached. For lysozyme, data for the heavy-metal conditions were collected at the *L*_{II} edge, rather than the more suitable *L*_{III} edge, which has a greater anomalous signal. This less-favourable wavelength was opted for in order to create somewhat poorer conditions

Table 2

Concentration and soaking time of heavy-metal/halide solutions with 200 nl proteinase K crystal drops for data collected at RT.

For cryo data collection, the soaking time was about 15 min for heavy-metal soaks and 30 min for halide soaks.

HA	Volume ejected (nl)	Stock solution (M)	Final concentration (mM)	RT soaking time (min)
NaBr	5.0	2	50.00	15–60
KBr	5.0	2	50.00	25–70
NaI	5.0	2	50.00	10–55
KI	5.0	2	50.00	25–70
K ₂ Pt(NO) ₄	5.0	0.1	2.50	15–60
KAuCl ₄ ·H ₂ O	2.5	0.1	1.25	5–45
K ₂ IrCl ₆	5.0	0.1	2.50	5–40
SmCl ₃ ·6H ₂ O	5.0	0.1	2.50	20–55

for lysozyme, the stable crystals of which are generally known to diffract very well. For proteinase K, data for the heavy-metal conditions were collected at the *L*_{III} edge for all of the conditions except samarium, data for which were collected at the *L*_{II} edge. This is because the *L*_{III} edge was not within the energy range of the beamline. To choose the corresponding energy at the *L*_{II} or *L*_{III} edge, an energy fluorescence scan was performed at the beamline followed by the use of *CHOOCH* (Evans & Pettifer, 2001).

2.4. Data processing

The data sets were indexed and integrated with *DIALS* (Winter *et al.*, 2018) and subsequently merged using *BLEND* (Axford *et al.*, 2015; Foadi *et al.*, 2013). The resulting merged data sets were scaled with *AIMLESS* (Evans, 2011; Evans & Murshudov, 2013); see Supplementary Tables S1, S2, S3 and S4. *HKL2MAP* and *SHELXC/D/E* (Pape & Schneider, 2004; Sheldrick, 2008; Usón & Sheldrick, 2018) were used to phase the derivatives using the SAD and SIRAS methods. The anomalous maps were calculated using *ANODE* (Thorn & Sheldrick, 2011) for each derivative, as shown in Figs. 2 and 3.

3. Results and discussion

The main goal of the present work is to demonstrate experimental phasing and structure solution with data from multiple RT crystals, as described above. In this article, the focus is on *de novo* phasing with a variety of popular heavy-atom derivatives and serial crystal data-collection strategies. Successful experimental phasing was achieved with four derivatives for lysozyme and one derivative for proteinase K.

3.1. Phasing at 100 K

Experimental phasing for both lysozyme and proteinase K was, in general, successful. Details are presented in Tables 3 and 4. SAD or SIRAS phasing was attempted using the *SHELX* family of programs (Sheldrick, 2008; Usón & Sheldrick, 2018; Pape & Schneider, 2004). There are three important indicators of success for experimental phasing when using *SHELXC/D/E*: (i) the cutoff resolution of the average anomalous signal, represented by *d*^{''} divided by its standard

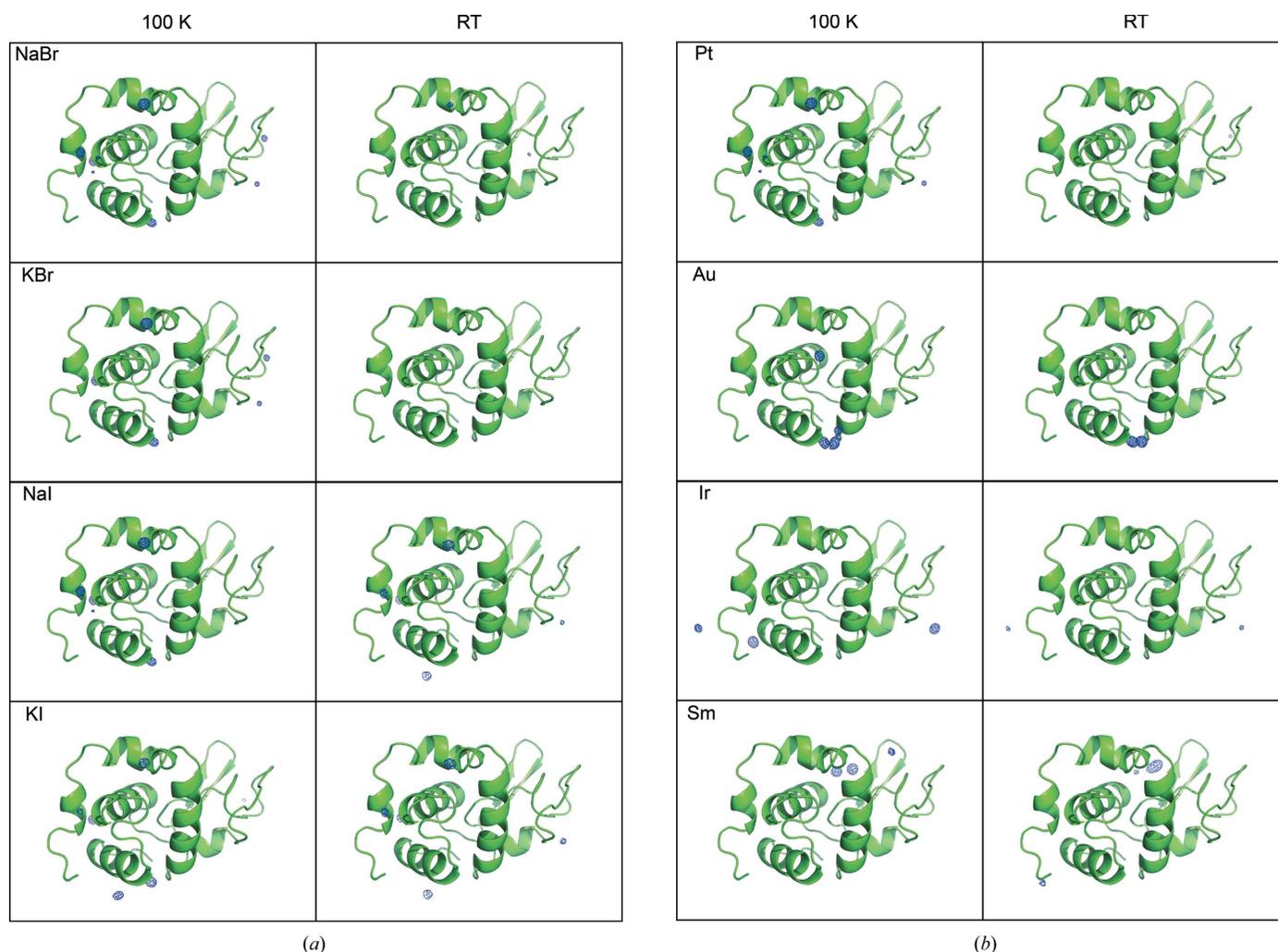


Figure 2 Heavy-atom anomalous signal in lysozyme. (a) Halide and (b) heavy-metal anomalous signal from diffraction data collected at 100 K (first column) and at RT (second column). The anomalous difference density map (shown in blue) was calculated using *ANODE* (Thorn & Sheldrick, 2011) and contoured at 10σ . The lysozyme protein backbone is represented as a green cartoon. The σ level was chosen to have strong visible peaks above the noise level. This figure was created using *PyMOL* (DeLano, 2008).

deviation, (ii) success or failure in the determination of the substructure within 1000 computational trials¹ and (iii) the fraction of residues fitted into the experimental electron-density map with the autotracing option. With regard to the first indicator, a higher resolution cutoff indicates a stronger anomalous signal derived from more reflections with significant anomalous differences. The third indicator reflects effective phasing because it quantitates the interpretability of an experimental electron-density map; better maps yield a larger number of fitted residues in the automated mode. The three indicators are sequential, interlocked thresholds. Low resolution or failure to find a suitable resolution cutoff as per the first indicator preclude subsequent substructure determination. Similarly, a failure in substructure determination (the second indicator) precludes model building (the fraction of autotracing; the third indicator) because interpretable electron density is not available. There is an obvious difference in

¹ The default in *SHELXD* is 100 trials. In our experience, 1000 trials proved to be a good compromise between efficiency and success. We tried up to 10 000 trials for a few cases, but this did not make any difference.

phasing effectiveness between lysozyme and proteinase K. Lysozyme is well known to produce high-quality crystals that provide good electron-density maps with nearly all experimental phasing methods. The resolution cutoff in the first indicator for lysozyme and proteinase K averaged 1.70 and 3.32 Å, respectively. Furthermore, anomalous signal for the KBr and iridium derivatives cannot be detected for proteinase K, and the value of the first indicator is very low, 8 Å, for the samarium derivative. Therefore, *SHELXD* cannot find substructures for these derivatives. However, phasing and model building are possible in all other cases except for the platinum derivative, where the first indicator has the relatively low value of 4 Å. In addition, the SIRAS methodology within *SHELX* appears to provide better maps than SAD.

3.2. Phasing at room temperature

The results of the phasing experiments for lysozyme and proteinase K data collected at RT are presented in Tables 5 and 6, respectively. Compared with cryogenic data, the anomalous signal is markedly lower for data collected at RT.

The average value of the first indicator for lysozyme is 2.30 Å (versus 1.70 Å) and that for proteinase K is 4.98 Å (versus 3.32 Å). The lower signal value can be explained by the higher background coming from a popular crystallization plate often used for RT data collection. This means that substructure determination is not possible for a few of the HA derivatives.

In general, higher resolutions in the first indicator are associated with substructure determination, but an exception to this rule is found, in the lysozyme case, for the iridium derivative. In contrast, the smaller proteinase K crystals used for RT data collection did not yield as many successes as the crystals at 100 K. Nevertheless, phasing is possible in at least

Table 3
Experimental phasing for lysozyme crystals using data sets collected at 100 K.

HA	200 × 200 × 50 μm lysozyme crystals and data sets collected at 100 K					
	SAD			SIRAS/native cryo		
	$\langle d'/\text{sig} \rangle$ resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)	$\langle d'/\text{sig} \rangle$ resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)
NaBr	1.50	Yes	98.4	1.50	Yes	85.3
KBr	1.60	Yes	93.0	1.60	Yes	88.4
NaI	2.00	Yes	86.8	2.00	Yes	94.6
KI	1.86	Yes	96.1	1.86	Yes	93.8
K ₂ Pt(NO) ₄	1.70	Yes	94.6	1.70	Yes	89.9
KAuCl ₄ ·H ₂ O	1.62	Yes	83.7	1.62	Yes	93.0
K ₂ IrCl ₆	1.50	Yes	96.1	1.50	Yes	94.6
SmCl ₃ ·6H ₂ O	1.80	Yes	86.7	1.80	Yes	85.3

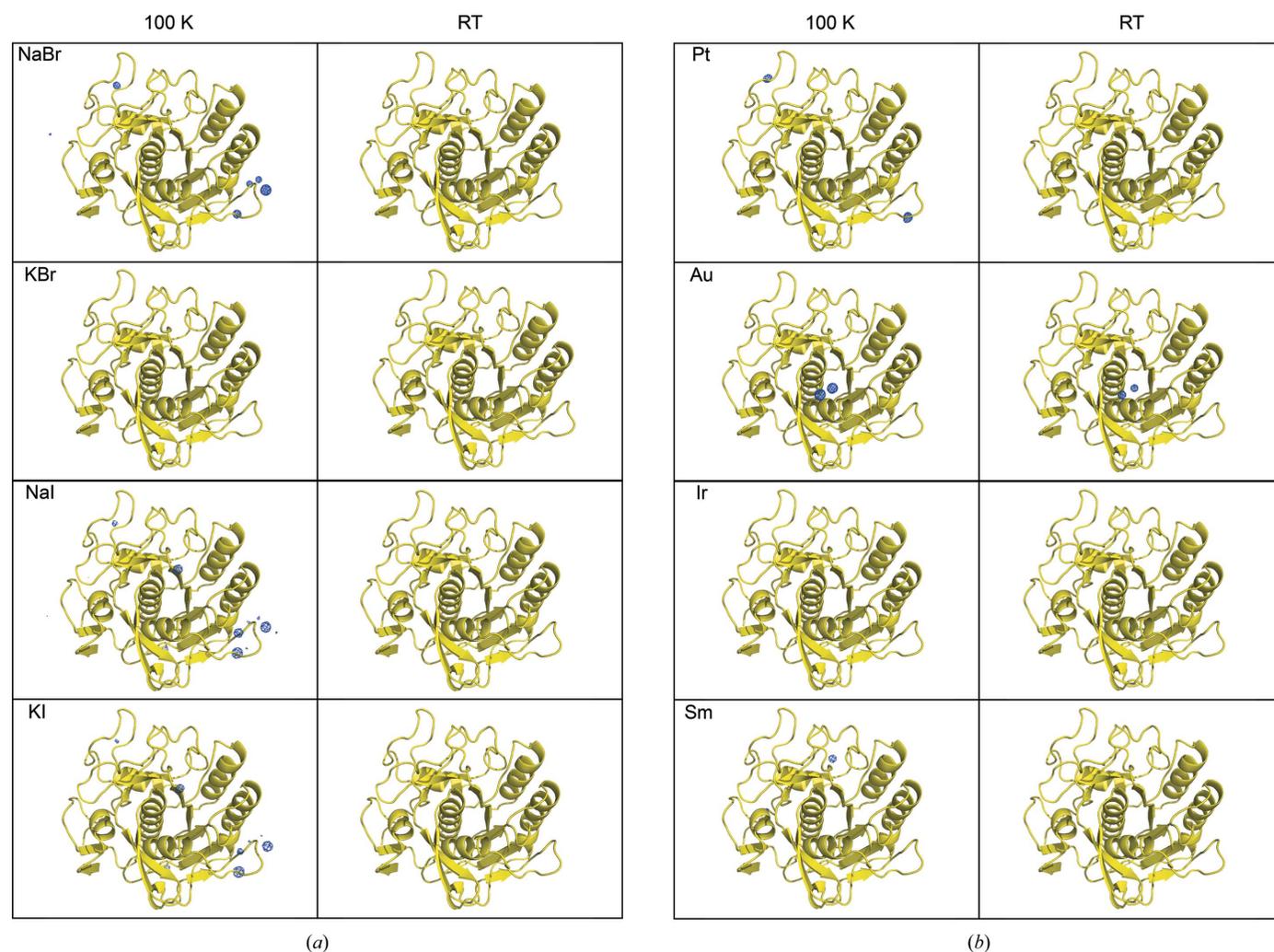


Figure 3
Heavy-atom anomalous signal in proteinase K. (a) Halide and (b) heavy-metal anomalous signal from diffraction data collected at 100 K (first column) and at RT (second column). The anomalous difference density map (shown in blue) was calculated using *ANODE* (Thorn & Sheldrick, 2011) and contoured at 10σ . The proteinase K backbone is represented as a yellow cartoon. The σ level was chosen to have strong visible peaks above the noise level. This figure was created using *PyMOL* (DeLano, 2008).

Table 4
Experimental phasing for proteinase K crystals using data sets collected at 100 K.

HA	10 × 10 × 10 μm proteinase K crystals and data sets collected at 100 K					
	SAD			SIRAS/native cryo		
	$\langle d'/\text{sig} \rangle$ anomalous signal resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)	$\langle d'/\text{sig} \rangle$ anomalous signal resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)
NaBr	1.90	Yes	98.9	1.90	Yes	97.8
KBr	No signal	N/A	N/A	No signal	N/A	N/A
NaI	2.15	Yes	62.3	2.15	Yes	96.0
KI	2.27	Yes	42.3	2.27	Yes	97.1
K ₂ Pt(NO) ₄	4.00	Yes	Failed	4.00	Yes	Failed
KAuCl ₄ ·H ₂ O	1.60	Yes	90.7	1.60	Yes	95.0
K ₂ IrCl ₆	No signal	N/A	N/A	No signal	N/A	N/A
SmCl ₃ ·6H ₂ O	8.00	No	N/A	8.00	No	N/A

Table 5
Experimental phasing for lysozyme crystals using data sets collected at RT.

HA	200 × 200 × 50 μm lysozyme crystals and <i>in situ</i> RT data sets					
	SAD			SIRAS/native <i>in situ</i>		
	$\langle d'/\text{sig} \rangle$ anomalous signal resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)	$\langle d'/\text{sig} \rangle$ anomalous signal resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)
NaBr	2.79	No	N/A	2.79	No	N/A
KBr	2.97	No	N/A	2.97	No	N/A
NaI	2.07	Yes	37.2	2.07	Yes	77.5
KI	2.07	Yes	39.5	2.07	Yes	87.6
K ₂ Pt(NO) ₄	2.84	No	N/A	2.84	No	N/A
KAuCl ₄ ·H ₂ O	1.74	Yes	81.4	1.74	Yes	76.0
K ₂ IrCl ₆	1.88	No	N/A	1.88	No	N/A
SmCl ₃ ·6H ₂ O	2.07	Yes	63.6	2.07	Yes	79.1

Table 6
Experimental phasing for proteinase K crystals using data sets collected at RT.

HA	10 × 10 × 10 μm proteinase K crystals and <i>in situ</i> RT data sets					
	SAD			SIRAS/native <i>in situ</i>		
	$\langle d'/\text{sig} \rangle$ anomalous signal resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)	$\langle d'/\text{sig} \rangle$ anomalous signal resolution cutoff (Å)	<i>SHELXD</i> solution	Residues autotraced with <i>SHELXE</i> (%)
NaBr	3.60	No	N/A	3.60	No	N/A
KBr	8.00	No	N/A	8.00	No	N/A
NaI	6.00	Partial	16.1	6.00	No	N/A
KI	3.50	No	N/A	3.50	No	N/A
K ₂ Pt(NO) ₄	8.00	No	N/A	8.00	No	N/A
KAuCl ₄ ·H ₂ O	2.60	Yes	9.7	2.60	Yes	95.7
K ₂ IrCl ₆	No signal	N/A	N/A	No signal	N/A	N/A
SmCl ₃ ·6H ₂ O	3.15	No	N/A	3.15	No	N/A

one case: that of the gold derivative. Finally, also at RT SIRAS seems to work better than SAD, still within *SHELX*.

3.3. Discussion

As clearly shown in Tables 3, 4, 5 and 6, the procedure documents the experimental phasing of crystal structures from serial MX data sets at RT. Not all of the samples or conditions were successful, most obviously because of the large contribution to background noise by the crystallization plates and the weaker diffraction data from the micrometre-sized samples within them. While this use case does not constitute an obstacle for the general application of the methodology, some comments are desirable in order to properly gauge

expectations. What can be anticipated from the following in-depth comments is that success or failure are governed by the interplay of several and diverse factors affecting the procedure. These include, but are not limited to, isomorphism of the crystals, the amount of inclusion of HAs in the derivatives, the size of the crystals, degradation owing to radiation damage and the anecdotal and well known intrinsic propensity of certain structures to be experimentally phased with respect to others.

3.3.1. Isomorphism. The number of crystals used for each derivative and the associated unit-cell variability parameter aLCV (absolute linear cell variation), calculated in *BLEND* and explained in Foadi *et al.* (2013), are included in Table 7. The aLCV summarizes the unit-cell size change and its

Table 7
Number of crystals and measure of unit-cell variability for all sets of derivatives.

HA	Lysozyme (200 × 200 × 50 μm)				Proteinase K (10 × 10 × 10 μm)			
	100 K		RT		100 K		RT	
	No. of crystals	aLCV† (Å)	No. of crystals	aLCV (Å)	No. of crystals	aLCV (Å)	No. of crystals	aLCV (Å)
NaBr	5	0.14	63	0.29	21	1.07	60	0.71
KBr	5	0.75	25	0.20	19	1.26	55	0.45
NaI	5	0.36	25	0.06	20	0.88	65	0.51
KI	6	0.31	24	0.11	21	0.89	55	0.44
K ₂ Pt(NO) ₄	5	0.16	54	0.11	21	0.52	98	0.18
KAuCl ₄ ·H ₂ O	5	0.58	24	0.21	47	1.16	116	0.57
K ₂ IrCl ₆	6	0.21	21	0.10	21	0.61	47	0.46
SmCl ₃ ·6H ₂ O	5	0.34	24	0.06	19	0.58	49	0.60

† Absolute linear cell variation calculated in *BLEND* (Foadi *et al.*, 2013).

numeric value is closely related to a linear expansion or contraction of the cell, measured in angstroms, with respect to the average unit cell of all crystals considered. This means that a higher aLCV value indicates lower isomorphism. Interesting conclusions from the data collected from RT crystals versus cryocooled crystals can be inferred by simple analysis of the mean and standard deviation of the aLCV for the presented cases (see Fig. 4). They have a distinct meaning in the present context. The mean aLCV indicates the variability within each of the four combinations: lysozyme at 100 K, lysozyme at RT, proteinase K at 100 K and proteinase K at RT. A higher mean aLCV, *i.e.* non-isomorphism, is more pronounced for proteinase K than for lysozyme both in the 100 K and the RT case. Within each structure, it is also evident that non-isomorphism is higher for cryocooled than for RT crystals; this corroborates the reasons that set us to develop this procedure for RT crystals. The other aLCV statistic, the standard deviation, provides an indication of the unit-cell variability caused by derivatization. Here also we can observe that the HA substrates cause lattice distortion in a greater measure in the cryocooled case. We can also observe that the substrates affect proteinase K crystals more than lysozyme crystals.

3.3.2. The inclusion of HAs in the crystals. It is clear from the results presented on phasing (Table 5 and Table 6) that binding of the heavy atoms to lysozyme occurs at RT in the case of NaI, KI, gold and samarium. For proteinase K, binding can only be ascertained directly from crystallographic data in the case of gold at RT. Figs. 2 and 3 display the 10σ contour of the anomalous difference maps, thus calculated to unequivocally show the locations of the peaks. Our results match those reported by Sun *et al.* (2002), in which short soaking times allow HA ligand binding and are thought to minimize lattice disorder. In the same study, cryocooling was carried out quickly after completing the ligand soak to stop diffusion and lock the heavy atom in the binding position, thereby avoiding lattice degradation.

We compared cryocooled and RT *in situ* HA phasing procedures in this study. The important issue is: have the HAs bound to the structures even if they are not immediately visible in the 10σ anomalous maps from data collected at RT but are present in data collected at 100 K? In our data sets, the anomalous peaks are visible in most of the anomalous maps at

100 K, but are less prominent in RT maps. Anomalous peaks from all of the HAs are observable in lysozyme RT and/or cryogenic maps (Fig. 2, Tables 3 and 5). This provides clear evidence that heavy atoms do bind to the proteins within the crystal. However, only about half of the HA derivatives produced strong anomalous peaks in RT data sets, whereas all of the cryogenic data did yield *de novo* phasing. The failures to phase are correlated with poor quality of the anomalous signal (see Table 5). The situation for KBr is illustrative, since this derivative did not yield 10σ anomalous peaks from RT data sets, but does show an anomalous signal at 3.5σ (Fig. 5). Although the 3.5σ map is noisier, two of the prominent bromide-binding sites identified in the 100 K data sets are also occupied in the RT data sets (see the arrows in Fig. 5). A similar scenario occurs for all of the HA derivatives of proteinase K, with the exception of KBr and iridium, which both exhibited extremely poor-quality anomalous signal (Fig. 3, Tables 4 and 6). In many cases, the anomalous peaks are visible at the 3.5σ threshold. Derivatization of proteinase

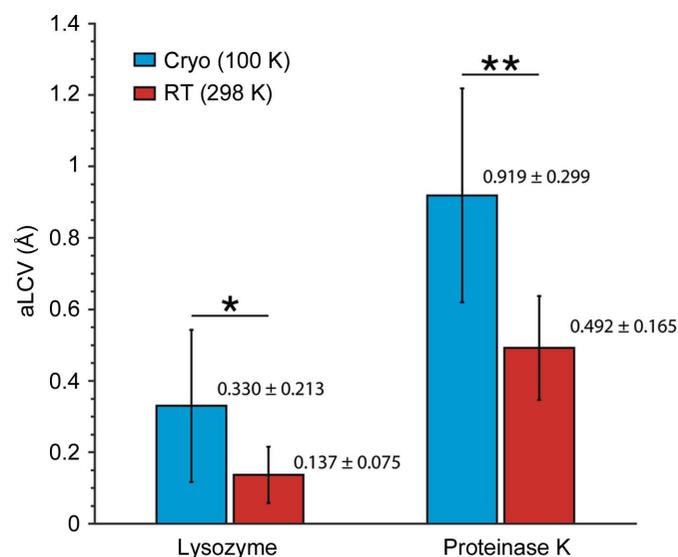


Figure 4
Mean and standard deviation of aLCV for lysozyme and proteinase K at 100 K and RT. Significant differences in aLCV are represented by * (95% confidence) and ** (99% confidence), as measured by a one-way ANOVA.

Table 8
Anomalous multiplicity for crystals of proteinase K.

HA	100 K	RT
NaBr	31.9 (12.0)	10.1 (9.6)
KBr	28.0 (10.4)	9.1 (9.0)
NaI	21.8 (7.7)	7.1 (5.5)
KI	24.8 (13.6)	7.8 (7.8)
K ₂ Pt(NO) ₄	34.0 (21.0)	11.0 (11.1)
KAuCl ₄ ·H ₂ O	82.9 (66.7)	13.9 (13.9)
K ₂ IrCl ₆	31.3 (9.3)	7.5 (7.2)
SmCl ₃ ·6H ₂ O	22.7 (7.4)	7.2 (7.3)

K with KBr or iridium appeared to fail since no anomalous peaks could be identified even in the 100 K data sets.

3.3.3. Crystal size. Proteinase K typically produced smaller crystals than lysozyme (see Section 2.1); the former were about 10 × 10 × 10 μm, whereas the latter were about 200 × 200 × 50 μm. This has obvious repercussions for the diffraction intensities (which are much lower for smaller crystals) and for the lifespan of each crystal during data collection (which is much shorter for smaller crystals) owing to the ensuing radiation damage. Thus, the smaller crystal size for proteinase K meant that the diffraction data were generally of poorer quality than those for lysozyme. An improvement in the quality is obtained when the number of crystals for a data set is substantially higher. Therefore, the number of crystals used for proteinase K was systematically larger than the number of lysozyme crystals. Still, at RT experimental phasing failed for all proteinase K derivatives, with the exception of the gold derivative. It is interesting to note that data from many more crystals (116; see Table 7) were indeed collected for the gold derivative. This seems to suggest that a larger number of successful cases could have been obtained if data from more

crystals had been collected for proteinase K. This observation is also corroborated by the anomalous multiplicity for each derivative at RT for proteinase K (see Table 8). We can observe that gold indeed has the highest anomalous multiplicity among all derivatives at RT.

3.3.4. Propensity for experimental phasing. There is ample anecdotal evidence that not all HA ligands enable the same propensity for experimental phasing in the same structure. As summarized in the review by Pike *et al.* (2016), HA binding depends on multiple factors such as the HA type and charge state, the availability of binding sites in the protein, the crystallization solution and the temperature. One of the major factors is the interaction of the HA with the protein crystallization solution. There is an affinity competition for the HA between the crystallization buffer and the potential binding sites on the protein. The affinity for the latter must be greater for a successful phasing experiment. We assume this to also be the case for the samples analysed in this article. For example, in the lysozyme RT case both iodide substrates led to successful phasing, while this did not happen with the bromide substrates. The iodide substrates also seem to perform better than the bromide substrates in the proteinase K RT case, although no phasing success was obtained. The main reason for including several diverse HA substrates in the procedure was in fact the different propensity for experimental phasing.

4. Conclusions

The procedure for experimental phasing tested and described in this article extends the emphasis on isomorphism for native data described in previous studies (Axford *et al.*, 2015; Foadi *et al.*, 2013). Classic derivatization procedures are lengthy and require sustained and substantial manual handling. This affects individual samples in a unique way, thus making the production of isomorphous crystals very difficult. The automated pipeline detailed above reduces human intervention to a minimum, making the production of similar crystals more likely. An important feature characterizing our method is the ready availability of a ‘derivatization kit’ made up of HA compounds which are nontoxic and are regularly used by the community. More hazardous compounds can be added to the kit if the liquid handler is exclusively dedicated to experimental phasing. In fact, the use of a liquid handler for HA soaking is, in such instances, even safer for the operator. For example, in this study, manual sample handling is minimized; the focus is on evaluating derivatization through a systematic analysis of the easy-to-assess indicators described in Section 3.1. To this end, the effect of different HAs on the crystals can be explored using the mean and standard deviation of the aLCV.

The success of this process ultimately depends on the data quality and redundancy. These, in turn, depend on crystal size. Evidently, a higher number of samples are required when the crystals are small, as in the case of proteinase K. However, this is not an insurmountable obstacle, especially in the light of modern serial crystallography at synchrotrons and XFELs.

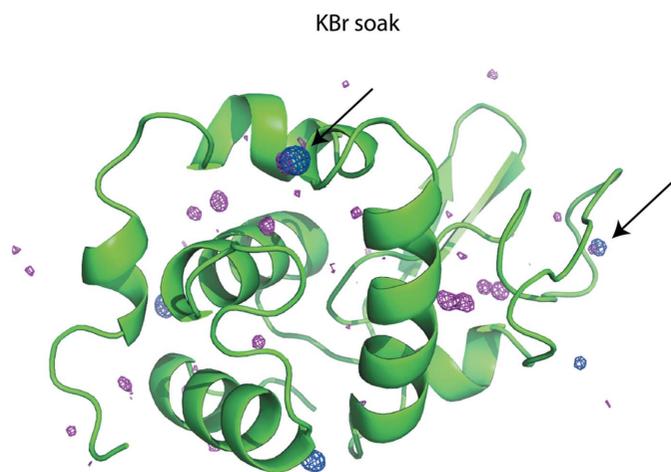


Figure 5
Bromide anomalous signal in lysozyme KBr soaks. Anomalous maps were calculated using ANODE (Thorn & Sheldrick, 2011). The anomalous signal of bromide is displayed at 10σ for diffraction data collected at 100 K (blue) and at 3.5σ for RT data (magenta). The lysozyme backbone is represented as a green cartoon. Black arrows represent anomalous peaks common to the RT and 100 K data sets. The sigma levels were chosen to have strong visible peaks above the noise level. This figure was created using PyMOL (DeLano, 2008).

Serial femtosecond crystallography (SFX) methods at XFEL sources very often exploit RT data-collection strategies using thousands of micrometre-sized crystals, from each of which only one still image is recorded (Chapman, 2019; Fromme, 2015; Schlichting, 2015; Spence, 2017). The crystals used for SFX are usually similar in size to the proteinase K samples used here. However, very high quality diffraction data from significantly smaller crystals, including submicrometre-sized samples, are often achievable using SFX methods (Gati *et al.*, 2017; Nass, Redecke *et al.*, 2020). Consequently, various sample-presentation strategies have been developed to minimize background noise for SFX, because nanometre-sized to micrometre-sized crystals produce weaker diffraction data compared with traditionally sized samples (Awel *et al.*, 2018; Beyerlein *et al.*, 2017; Calvey *et al.*, 2019; Dasgupta *et al.*, 2019; Davy *et al.*, 2019; Doak *et al.*, 2018; Ebrahim, Appleby *et al.*, 2019; Ebrahim, Moreno-Chicano *et al.*, 2019; Echelmeier *et al.*, 2019; Fuller *et al.*, 2017; Grunbein & Nass Kovacs, 2019; Lieske *et al.*, 2019; Martiel *et al.*, 2019; Meents *et al.*, 2017; Mehrabi *et al.*, 2019; Monteiro *et al.*, 2019; Nass, Gorel *et al.*, 2020; Nogly *et al.*, 2016; Oberthuer *et al.*, 2017; Orville, 2017; Owen *et al.*, 2017; Roedig *et al.*, 2017; Schulz *et al.*, 2018, 2019; Shelby *et al.*, 2020; Sierra *et al.*, 2016; Stagno *et al.*, 2017; Suga *et al.*, 2020; Sugahara *et al.*, 2017; Weinert *et al.*, 2017; Wiedorn *et al.*, 2018; Zhao *et al.*, 2019). In addition, X-ray pulse durations of tens of femtoseconds and the tight focus and intensity of XFEL pulses often produce very high-quality data sets with little or no radiation-induced effects in the data and refined atomic models. Progress with SFX methods applying *de novo* phasing demonstrates the value of phasing with larger numbers of samples (Gorel, Motomura, Fukuzawa, Doak, Grünbein, Hilpert, Inoue, Kloos, Kováčsová *et al.*, 2017; Gorel, Motomura, Fukuzawa, Doak, Grünbein, Hilpert, Inoue, Kloos, Nass Kovacs *et al.*, 2017; Nass *et al.*, 2016; Yamashita *et al.*, 2017). However, the extreme limits in XFEL beamtime also highlight the need to develop appropriate methods for *de novo* phasing with serial MX approaches at synchrotrons using smaller samples and potentially RT methods. It is therefore appropriate to extend the method described in this paper to the context of serial crystallography. This will necessitate special attention to the analysis of still images because of the uncertainty in the determination of unit-cell parameters.

5. Related literature

The following reference is cited in the supporting information for this article: Zeldin *et al.* (2013).

Acknowledgements

We would like to thank Dr Agata Butryn for her critical review of the manuscript, and the I24 and I04 teams. People from the XCEM facility have been invaluable in providing guidance and support using the Echo. Diffraction data were measured on beamlines I24 and I04 at DLS (proposal Nos. NR14449 and NR16818). All data sets used in this research are freely available for research purposes. Please contact us to

obtain access. A video showing the workflow is available at <https://www.diamond.ac.uk/Instruments/Mx/XFEL-Hub/Staff/Aller/Research.html>.

Funding information

This work was supported by Diamond Light Source, which receives funding from the UK Government and the Wellcome Trust. PA and AMO were also partially supported by funding from the Biotechnology and Biological Sciences Research Council and a Strategic Award from the Wellcome Trust (102593); AMO was also partially supported by a Royal Society Wolfson Fellowship (RSWF/R2\182017) and a Wellcome Investigator Award in Science (210734/Z/18/Z).

References

- Agniswamy, J., Joyce, M. G., Hammer, C. H. & Sun, P. D. (2008). *Acta Cryst. D* **64**, 354–367.
- Aller, P., Sanchez-Weatherby, J., Foadi, J., Winter, G., Lobley, C. M. C., Axford, D., Ashton, A. W., Bellini, D., Brandao-Neto, J., Culurgioni, S., Douangamath, A., Duman, R., Evans, G., Fisher, S., Flaig, R., Hall, D. R., Lukacik, P., Mazzorana, M., McAuley, K. E., Mykhaaylyk, V., Owen, R. L., Paterson, N. G., Romano, P., Sandy, J., Sorensen, T., von Delft, F., Wagner, A., Warren, A., Williams, M., Stuart, D. I. & Walsh, M. A. (2015). *Methods Mol. Biol.* **1261**, 233–253.
- Arakawa, T., Kobayashi-Yurugi, T., Alguel, Y., Iwanari, H., Hatae, H., Iwata, M., Abe, Y., Hino, T., Ikeda-Suno, C., Kuma, H., Kang, D., Murata, T., Hamakubo, T., Cameron, A. D., Kobayashi, T., Hamasaki, N. & Iwata, S. (2015). *Science*, **350**, 680–684.
- Awel, S., Kirian, R. A., Wiedorn, M. O., Beyerlein, K. R., Roth, N., Horke, D. A., Oberthür, D., Knoska, J., Mariani, V., Morgan, A., Adriano, L., Tolstikova, A., Xavier, P. L., Yefanov, O., Aquila, A., Barty, A., Roy-Chowdhury, S., Hunter, M. S., James, D., Robinson, J. S., Weierstall, U., Rode, A. V., Bajt, S., Küpper, J. & Chapman, H. N. (2018). *J. Appl. Cryst.* **51**, 133–139.
- Axford, D., Aller, P., Sanchez-Weatherby, J. & Sandy, J. (2016). *Acta Cryst. F* **72**, 313–319.
- Axford, D., Foadi, J., Hu, N.-J., Choudhury, H. G., Iwata, S., Beis, K., Evans, G. & Alguel, Y. (2015). *Acta Cryst. D* **71**, 1228–1237.
- Axford, D., Owen, R. L., Aishima, J., Foadi, J., Morgan, A. W., Robinson, J. I., Nettleship, J. E., Owens, R. J., Moraes, I., Fry, E. E., Grimes, J. M., Harlos, K., Kotecha, A., Ren, J., Sutton, G., Walter, T. S., Stuart, D. I. & Evans, G. (2012). *Acta Cryst. D* **68**, 592–600.
- Beyerlein, K. R., Dierksmeyer, D., Mariani, V., Kuhn, M., Sarrou, I., Ottaviano, A., Awel, S., Knoska, J., Fuglerud, S., Jönsson, O., Stern, S., Wiedorn, M. O., Yefanov, O., Adriano, L., Bean, R., Burkhardt, A., Fischer, P., Heymann, M., Horke, D. A., Jungnickel, K. E. J., Kovaleva, E., Lorbeer, O., Metz, M., Meyer, J., Morgan, A., Pande, K., Panneerselvam, S., Seuring, C., Tolstikova, A., Lieske, J., Aplin, S., Roessle, M., White, T. A., Chapman, H. N., Meents, A. & Oberthuer, D. (2017). *IUCrJ*, **4**, 769–777.
- Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X., Ebner, S. G., Isenegger, A., Schneider, R., Schneider, J., Gletting, W., Pradervand, C., Panepucci, E. H., Tomizaki, T., Wang, M. & Schulze-Briese, C. (2011). *Cryst. Growth Des.* **11**, 916–923.
- Calvey, G. D., Katz, A. M. & Pollack, L. (2019). *Anal. Chem.* **91**, 7139–7144.
- Chapman, H. N. (2019). *Annu. Rev. Biochem.* **88**, 35–58.
- Chilingaryan, Z., Yin, Z. & Oakley, A. J. (2012). *Int. J. Mol. Sci.* **13**, 12857–12879.
- Collins, P. M., Ng, J. T., Talon, R., Nekrosiute, K., Krojer, T., Douangamath, A., Brandao-Neto, J., Wright, N., Pearce, N. M. & von Delft, F. (2017). *Acta Cryst. D* **73**, 246–255.
- Crick, F. H. C. & Magdoff, B. S. (1956). *Acta Cryst.* **9**, 901–908.

- Dasgupta, M., Budday, D., de Oliveira, S. H. P., Madzelan, P., Marchany-Rivera, D., Seravalli, J., Hayes, B., Sierra, R. G., Boutet, S., Hunter, M. S., Alonso-Mori, R., Batyuk, A., Wierman, J., Lyubimov, A., Brewster, A. S., Sauter, N. K., Applegate, G. A., Tiwari, V. K., Berkowitz, D. B., Thompson, M. C., Cohen, A. E., Fraser, J. S., Wall, M. E., van den Bedem, H. & Wilson, M. A. (2019). *Proc. Natl Acad. Sci. USA*, **116**, 25634–25640.
- Davies, A. M., Rispens, T., Ooijevaar-de Heer, P., Aalberse, R. C. & Sutton, B. J. (2017). *Mol. Immunol.* **81**, 85–91.
- Davy, B., Axford, D., Beale, J. H., Butryn, A., Docker, P., Ebrahim, A., Leen, G., Orville, A. M., Owen, R. L. & Aller, P. (2019). *J. Synchrotron Rad.* **26**, 1820–1825.
- DeLano, W. L. (2008). *PyMOL*. <http://www.pymol.org>.
- Doak, R. B., Nass Kovacs, G., Gorel, A., Foucar, L., Barends, T. R. M., Grünbein, M. L., Hilpert, M., Kloos, M., Roome, C. M., Shoeman, R. L., Stricker, M., Tono, K., You, D., Ueda, K., Sherrell, D. A., Owen, R. L. & Schlichting, I. (2018). *Acta Cryst. D* **74**, 1000–1007.
- Douangamath, A., Aller, P., Lukacik, P., Sanchez-Weatherby, J., Moraes, I. & Brandao-Neto, J. (2013). *Acta Cryst. D* **69**, 920–923.
- Ebrahim, A., Appleby, M. V., Axford, D., Beale, J., Moreno-Chicano, T., Sherrell, D. A., Strange, R. W., Hough, M. A. & Owen, R. L. (2019). *Acta Cryst. D* **75**, 151–159.
- Ebrahim, A., Moreno-Chicano, T., Appleby, M. V., Chaplin, A. K., Beale, J. H., Sherrell, D. A., Duyvesteyn, H. M. E., Owada, S., Tono, K., Sugimoto, H., Strange, R. W., Worrall, J. A. R., Axford, D., Owen, R. L. & Hough, M. A. (2019). *IUCrJ*, **6**, 543–551.
- Echelmeier, A., Kim, D., Cruz Villarreal, J., Coe, J., Quintana, S., Brehm, G., Egatz-Gomez, A., Nazari, R., Sierra, R. G., Koglin, J. E., Batyuk, A., Hunter, M. S., Boutet, S., Zatsepin, N., Kirian, R. A., Grant, T. D., Fromme, P. & Ros, A. (2019). *J. Appl. Cryst.* **52**, 997–1008.
- Evans, G. & Pettifer, R. F. (2001). *J. Appl. Cryst.* **34**, 82–86.
- Evans, P. R. (2011). *Acta Cryst. D* **67**, 282–292.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst. D* **69**, 1204–1214.
- Farley, C., Burks, G., Siebert, T. & Juers, D. H. (2014). *Acta Cryst. D* **70**, 2111–2124.
- Foadi, J., Aller, P., Alguet, Y., Cameron, A., Axford, D., Owen, R. L., Armour, W., Waterman, D. G., Iwata, S. & Evans, G. (2013). *Acta Cryst. D* **69**, 1617–1632.
- Fromme, P. (2015). *Nat. Chem. Biol.* **11**, 895–899.
- Fuller, F. D., Gul, S., Chatterjee, R., Burgie, E. S., Young, I. D., LeBrette, H., Srinivas, V., Brewster, A. S., Michels-Clark, T., Clinger, J. A., Andi, B., Ibrahim, M., Pastor, E., de Lichtenberg, C., Hussein, R., Pollock, C. J., Zhang, M., Stan, C. A., Kroll, T., Fransson, T., Weninger, C., Kubin, M., Aller, P., Lassalle, L., Bräuer, P., Miller, M. D., Amin, M., Koroidov, S., Roessler, C. G., Allaire, M., Sierra, R. G., Docker, P. T., Glowina, J. M., Nelson, S., Koglin, J. E., Zhu, D., Chollet, M., Song, S., Lemke, H., Liang, M., Sokaras, D., Alonso-Mori, R., Zouni, A., Messinger, J., Bergmann, U., Boal, A. K., Bollinger, J. M. Jr, Krebs, C., Högbom, M., Phillips, G. N. Jr, Vierstra, R. D., Sauter, N. K., Orville, A. M., Kern, J., Yachandra, V. K. & Yano, J. (2017). *Nat. Methods*, **14**, 443–449.
- Garman, E. & Murray, J. W. (2003). *Acta Cryst. D* **59**, 1903–1913.
- Garman, E. F. (2010). *Acta Cryst. D* **66**, 339–351.
- Garman, E. F. & Owen, R. L. (2006). *Acta Cryst. D* **62**, 32–47.
- Gati, C., Oberthuer, D., Yefanov, O., Bunker, R. D., Stellato, F., Chiu, E., Yeh, S. M., Aquila, A., Basu, S., Bean, R., Beyerlein, K. R., Botha, S., Boutet, S., DePonte, D. P., Doak, R. B., Fromme, R., Galli, L., Grotjohann, I., James, D. R., Kupitz, C., Lomb, L., Messerschmidt, M., Nass, K., Rendek, K., Shoeman, R. L., Wang, D., Weierstall, U., White, T. A., Williams, G. J., Zatsepin, N. A., Fromme, P., Spence, J. C. H., Goldie, K. N., Jehle, J. A., Metcalf, P., Barty, A. & Chapman, H. N. (2017). *Proc. Natl Acad. Sci. USA*, **114**, 2247–2252.
- Gelin, M., Delfosse, V., Allemand, F., Hoh, F., Sallaz-Damaz, Y., Pirocchi, M., Bourguet, W., Ferrer, J.-L., Labesse, G. & Guichou, J.-F. (2015). *Acta Cryst. D* **71**, 1777–1787.
- Giordano, R., Leal, R. M. F., Bourenkov, G. P., McSweeney, S. & Popov, A. N. (2012). *Acta Cryst. D* **68**, 649–658.
- Gorel, A., Motomura, K., Fukuzawa, H., Doak, R. B., Grünbein, M. L., Hilpert, M., Inoue, I., Kloos, M., Kováčsová, G., Nango, E., Nass, K., Roome, C. M., Shoeman, R. L., Tanaka, R., Tono, K., Joti, Y., Yabashi, M., Iwata, S., Foucar, L., Ueda, K., Barends, T. R. M. & Schlichting, I. (2017). *Nat. Commun.* **8**, 1170.
- Gorel, A., Motomura, K., Fukuzawa, H., Doak, R. B., Grünbein, M. L., Hilpert, M., Inoue, I., Kloos, M., Nass Kovacs, G., Nango, E., Nass, K., Roome, C. M., Shoeman, R. L., Tanaka, R., Tono, K., Foucar, L., Joti, Y., Yabashi, M., Iwata, S., Ueda, K., Barends, T. R. M. & Schlichting, I. (2017). *Sci Data*, **4**, 170188.
- Grünbein, M. L. & Nass Kovacs, G. (2019). *Acta Cryst. D* **75**, 178–191.
- Hadimioglu, B., Stearns, R. & Ellson, R. (2016). *J. Lab. Autom.* **21**, 4–18.
- Huang, C.-Y., Olieric, V., Ma, P., Howe, N., Vogeley, L., Liu, X., Warshamange, R., Weinert, T., Panepucci, E., Kobilka, B., Diederichs, K., Wang, M. & Caffrey, M. (2016). *Acta Cryst. D* **72**, 93–112.
- Huang, C.-Y., Olieric, V., Ma, P., Panepucci, E., Diederichs, K., Wang, M. & Caffrey, M. (2015). *Acta Cryst. D* **71**, 1238–1256.
- Joyce, M. G., Radaev, S. & Sun, P. D. (2010). *Acta Cryst. D* **66**, 358–365.
- Lieske, J., Cerv, M., Kreida, S., Komadina, D., Fischer, J., Barthelmess, M., Fischer, P., Pakendorf, T., Yefanov, O., Mariani, V., Seine, T., Ross, B. H., Crosas, E., Lorbeer, O., Burkhardt, A., Lane, T. J., Guenther, S., Bergtholdt, J., Schoen, S., Törnroth-Horsefield, S., Chapman, H. N. & Meents, A. (2019). *IUCrJ*, **6**, 714–728.
- Liu, Q., Dahmane, T., Zhang, Z., Assur, Z., Brasch, J., Shapiro, L., Mancina, F. & Hendrickson, W. A. (2012). *Science*, **336**, 1033–1037.
- Liu, Q., Liu, Q. & Hendrickson, W. A. (2013). *Acta Cryst. D* **69**, 1314–1332.
- Liu, Q., Zhang, Z. & Hendrickson, W. A. (2011). *Acta Cryst. D* **67**, 45–59.
- Lu, J. & Sun, P. D. (2014). *FEBS J.* **281**, 4021–4028.
- Maire, A. le, Gelin, M., Pochet, S., Hoh, F., Pirocchi, M., Guichou, J.-F., Ferrer, J.-L. & Labesse, G. (2011). *Acta Cryst. D* **67**, 747–755.
- Martiel, I., Müller-Werkmeister, H. M. & Cohen, A. E. (2019). *Acta Cryst. D* **75**, 160–177.
- Meents, A., Wiedorn, M. O., Srajer, V., Henning, R., Sarrou, I., Bergtholdt, J., Barthelmess, M., Reinke, P. Y. A., Dierksmeyer, D., Tolstikova, A., Schaible, S., Messerschmidt, M., Ogata, C. M., Kissick, D. J., Taft, M. H., Manstein, D. J., Lieske, J., Oberthuer, D., Fischetti, R. F. & Chapman, H. N. (2017). *Nat. Commun.* **8**, 1281.
- Mehrabi, P., Schulz, E. C., Agthe, M., Horrell, S., Bourenkov, G., von Stetten, D., Leimkohl, J. P., Schikora, H., Schneider, T. R., Pearson, A. R., Tellkamp, F. & Miller, R. J. D. (2019). *Nat. Methods*, **16**, 979–982.
- Monteiro, D. C. F., Vakili, M., Harich, J., Sztucki, M., Meier, S. M., Horrell, S., Josts, I. & Trebbin, M. (2019). *J. Synchrotron Rad.* **26**, 406–412.
- Mylona, A., Carr, S., Aller, P., Moraes, I., Treisman, R., Evans, G. & Foadi, J. (2017). *Crystals*, **7**, 242.
- Nass, K., Gorel, A., Abdullah, M. M. V., Martin, A., Kloos, M., Marinelli, A., Aquila, A., Barends, T. R. M., Decker, F. J., Doak, R. B., Foucar, L., Hartmann, E., Hilpert, M., Hunter, M. S., Jurek, Z., Koglin, J. E., Kozlov, A., Lutman, A. A., Kovacs, G. N., Roome, C. M., Shoeman, R. L., Santra, R., Quiney, H. M., Ziaja, B., Boutet, S. & Schlichting, I. (2020). *Nat. Commun.* **11**, 1814.
- Nass, K., Meinhart, A., Barends, T. R. M., Foucar, L., Gorel, A., Aquila, A., Botha, S., Doak, R. B., Koglin, J., Liang, M., Shoeman, R. L., Williams, G., Boutet, S. & Schlichting, I. (2016). *IUCrJ*, **3**, 180–191.
- Nass, K., Redecke, L., Perbandt, M., Yefanov, O., Klinge, M., Koopmann, R., Stellato, F., Gabdulhakov, A., Schönherr, R., Rehders, D., Lahey-Rudolph, J. M., Aquila, A., Barty, A., Basu, S., Doak, R. B., Duden, R., Frank, M., Fromme, R., Kassemeyer, S.,

- Katona, G., Kirian, R., Liu, H., Majoul, I., Martin-Garcia, J. M., Messerschmidt, M., Shoeman, R. L., Weierstall, U., Westenhoff, S., White, T. A., Williams, G. J., Yoon, C. H., Zatsepin, N., Fromme, P., Duzsenko, M., Chapman, H. N. & Betzel, C. (2020). *Nat. Commun.* **11**, 620.
- Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I. & Perrakis, A. (2005). *Acta Cryst. D* **61**, 1426–1431.
- Nogly, P., Panneels, V., Nelson, G., Gati, C., Kimura, T., Milne, C., Milathianaki, D., Kubo, M., Wu, W., Conrad, C., Coe, J., Bean, R., Zhao, Y., Båth, P., Dods, R., Harimoorthy, R., Beyerlein, K. R., Rheinberger, J., James, D., DePonte, D., Li, C., Sala, L., Williams, G. J., Hunter, M. S., Koglin, J. E., Berntsen, P., Nango, E., Iwata, S., Chapman, H. N., Fromme, P., Frank, M., Abela, R., Boutet, S., Barty, A., White, T. A., Weierstall, U., Spence, J., Neutze, R., Schertler, G. & Standfuss, J. (2016). *Nat. Commun.* **7**, 12314.
- Oberthuer, D., Knoška, J., Wiedorn, M. O., Beyerlein, K. R., Bushnell, D. A., Kovaleva, E. G., Heymann, M., Gumprecht, L., Kirian, R. A., Barty, A., Mariani, V., Tolstikova, A., Adriano, L., Awel, S., Barthelmeß, M., Dörner, K., Xavier, P. L., Yefanov, O., James, D. R., Nelson, G., Wang, D., Calvey, G., Chen, Y., Schmidt, A., Szczepek, M., Frielingsdorf, S., Lenz, O., Snell, E., Robinson, P. J., Sarler, B., Belšák, G., Maček, M., Wilde, F., Aquila, A., Boutet, S., Liang, M., Hunter, M. S., Scheerer, P., Lipscomb, J. D., Weierstall, U., Kornberg, R. D., Spence, J. C. H., Pollack, L., Chapman, H. N. & Bajt, S. (2017). *Sci. Rep.* **7**, 44628.
- Orville, A. M. (2017). *X-Ray Free Electron Lasers: Applications in Materials, Chemistry and Biology*, edited by U. Bergmann, V. Yachandra & J. Yano, pp. 348–364. London: Royal Society of Chemistry.
- Owen, R. L., Axford, D., Sherrell, D. A., Kuo, A., Ernst, O. P., Schulz, E. C., Miller, R. J. D. & Mueller-Werkmeister, H. M. (2017). *Acta Cryst. D* **73**, 373–378.
- Owen, R. L., Paterson, N., Axford, D., Aishima, J., Schulze-Briese, C., Ren, J., Fry, E. E., Stuart, D. I. & Evans, G. (2014). *Acta Cryst. D* **70**, 1248–1256.
- Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4912–4917.
- Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
- Pike, A. C. W., Garman, E. F., Krojer, T., von Delft, F. & Carpenter, E. P. (2016). *Acta Cryst. D* **72**, 303–318.
- Roedig, P., Ginn, H. M., Pakendorf, T., Sutton, G., Harlos, K., Walter, T. S., Meyer, J., Fischer, P., Duman, R., Vartiainen, I., Reime, B., Warmer, M., Brewster, A. S., Young, I. D., Michels-Clark, T., Sauter, N. K., Kotecha, A., Kelly, J., Rowlands, D. J., Sikorsky, M., Nelson, S., Damiani, D. S., Alonso-Mori, R., Ren, J., Fry, E. E., David, C., Stuart, D. I., Wagner, A. & Meents, A. (2017). *Nat. Methods*, **14**, 805–810.
- Rose, J. P., Wang, B.-C. & Weiss, M. S. (2015). *IUCrJ*, **2**, 431–440.
- Rould, M. A. (2007). *Methods Mol. Biol.* **364**, 159–182.
- Sanchez-Weatherby, J., Sandy, J., Mikolajek, H., Lobley, C. M. C., Mazzorana, M., Kelly, J., Preece, G., Littlewood, R. & Sørensen, T. L.-M. (2019). *J. Synchrotron Rad.* **26**, 291–301.
- Schlichting, I. (2015). *IUCrJ*, **2**, 246–255.
- Schulz, E. C., Mehrabi, P., Müller-Werkmeister, H. M., Tellkamp, F., Jha, A., Stuart, W., Persch, E., De Gasparo, R., Diederich, F., Pai, E. F. & Miller, R. J. D. (2018). *Nat. Methods*, **15**, 901–904.
- Schulz, J., Bielecki, J., Doak, R. B., Dörner, K., Graceffa, R., Shoeman, R. L., Sikorski, M., Thute, P., Westphal, D. & Mancuso, A. P. (2019). *J. Synchrotron Rad.* **26**, 339–345.
- Shelby, M. L., Gilbale, D., Grant, T. D., Seuring, C., Segelke, B. W., He, W., Evans, A. C., Pakendorf, T., Fischer, P., Hunter, M. S., Bartyuk, A., Barthelmeß, M., Meents, A., Coleman, M. A., Kuhl, T. L. & Frank, M. (2020). *IUCrJ*, **7**, 30–41.
- Sheldrick, G. M. (2008). *Acta Cryst. A* **64**, 112–122.
- Sheldrick, G. M. (2010). *Acta Cryst. D* **66**, 479–485.
- Sierra, R. G., Gati, C., Laksmono, H., Dao, E. H., Gul, S., Fuller, F., Kern, J., Chatterjee, R., Ibrahim, M., Brewster, A. S., Young, I. D., Michels-Clark, T., Aquila, A., Liang, M., Hunter, M. S., Koglin, J. E., Boutet, S., Junco, E. A., Hayes, B., Bogan, M. J., Hampton, C. Y., Puglisi, E. V., Sauter, N. K., Stan, C. A., Zouni, A., Yano, J., Yachandra, V. K., Soltis, S. M., Puglisi, J. D. & DeMirici, H. (2016). *Nat. Methods*, **13**, 59–62.
- Spence, J. C. H. (2017). *IUCrJ*, **4**, 322–339.
- Stagno, J. R., Liu, Y., Bhandari, Y. R., Conrad, C. E., Panja, S., Swain, M., Fan, L., Nelson, G., Li, C., Wendel, D. R., White, T. A., Coe, J. D., Wiedorn, M. O., Knoška, J., Oberthuer, D., Tuckey, R. A., Yu, P., Dyba, M., Tarasov, S. G., Weierstall, U., Grant, T. D., Schwieters, C. D., Zhang, J., Ferré-D'Amaré, A. R., Fromme, P., Draper, D. E., Liang, M., Hunter, M. S., Boutet, S., Tan, K., Zuo, X., Ji, X., Barty, A., Zatsepin, N. A., Chapman, H. N., Spence, J. C. H., Woodson, S. A. & Wang, Y.-X. (2017). *Nature*, **541**, 242–246.
- Suga, M., Shimada, A., Akita, F., Shen, J. R., Toshi, T. & Sugimoto, H. (2020). *Biochim. Biophys. Acta*, **1864**, 129466.
- Sugahara, M., Nakane, T., Masuda, T., Suzuki, M., Inoue, S., Song, C., Tanaka, R., Nakatsu, T., Mizohata, E., Yumoto, F., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Numata, K., Nango, E. & Iwata, S. (2017). *Sci. Rep.* **7**, 703.
- Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst. D* **58**, 1092–1098.
- Taylor, G. L. (2010). *Acta Cryst. D* **66**, 325–338.
- Terwilliger, T. C., Bunkóczi, G., Hung, L.-W., Zwart, P. H., Smith, J. L., Akey, D. L. & Adams, P. D. (2016). *Acta Cryst. D* **72**, 359–374.
- Thorn, A. & Sheldrick, G. M. (2011). *J. Appl. Cryst.* **44**, 1285–1287.
- Usón, I. & Sheldrick, G. M. (2018). *Acta Cryst. D* **74**, 106–116.
- Villaseñor, A. G., Wong, A., Shao, A., Garg, A., Donohue, T. J., Kuglstatter, A. & Harris, S. F. (2012). *Acta Cryst. D* **68**, 893–900.
- Weinert, T., Olieric, N., Cheng, R., Brünle, S., James, D., Ozerov, D., Gashi, D., Vera, L., Marsh, M., Jaeger, K., Dworkowski, F., Panepucci, E., Basu, S., Skopintsev, P., Doré, A. S., Geng, T., Cooke, R. M., Liang, M., Prota, A. E., Panneels, V., Nogly, P., Ermler, U., Schertler, G., Hennig, M., Steinmetz, M. O., Wang, M. & Standfuss, J. (2017). *Nat. Commun.* **8**, 542.
- Wiedorn, M. O., Awel, S., Morgan, A. J., Ayer, K., Gevorkov, Y., Fleckenstein, H., Roth, N., Adriano, L., Bean, R., Beyerlein, K. R., Chen, J., Coe, J., Cruz-Mazo, F., Ekeberg, T., Graceffa, R., Heymann, M., Horke, D. A., Knoška, J., Mariani, V., Nazari, R., Oberthür, D., Samanta, A. K., Sierra, R. G., Stan, C. A., Yefanov, O., Rompotis, D., Correa, J., Erk, B., Treusch, R., Schulz, J., Hogue, B. G., Gañán-Calvo, A. M., Fromme, P., Küpper, J., Rode, A. V., Bajt, S., Kirian, R. A. & Chapman, H. N. (2018). *IUCrJ*, **5**, 574–584.
- Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Cryst. D* **74**, 85–97.
- Yamashita, K., Kuwabara, N., Nakane, T., Murai, T., Mizohata, E., Sugahara, M., Pan, D., Masuda, T., Suzuki, M., Sato, T., Kodan, A., Yamaguchi, T., Nango, E., Tanaka, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Manya, H., Endo, T., Kato, R., Senda, T., Kato, H., Iwata, S., Ago, H., Yamamoto, M., Yumoto, F. & Nakatsu, T. (2017). *IUCrJ*, **4**, 639–647.
- Zeldin, O. B., Gerstel, M. & Garman, E. F. (2013). *J. Appl. Cryst.* **46**, 1225–1230.
- Zhao, F.-Z., Zhang, B., Yan, E.-K., Sun, B., Wang, Z.-J., He, J.-H. & Yin, D.-C. (2019). *FEBS J.* **286**, 4402–4417.