JOURNAL OF SYNCHROTRON RADIATION
ISSN 1600-5775

In situ X-ray data collection and structure phasing of protein crystals at Structural Biology Center 19-ID

Karolina Michalska,[a,b] Kemin Tan,[a,b] Changsoo Chang,[a,b] Hui Li,[a] Catherine Hatzos-Skintges,[a] Michael Molitsky,[b] Randy Alkire[b] and Andrzej Joachimiak[a,b,c]*

[a]Midwest Center for Structural Genomics, Biosciences Division, Argonne National Laboratory, USA, [b]Structural Biology Center, Biosciences Division, Argonne National Laboratory, USA, and [c]Department of Biochemistry and Molecular Biology, University of Chicago, USA. *Correspondence e-mail: andrzejj@anl.gov

Received 29 June 2015
Accepted 5 September 2015

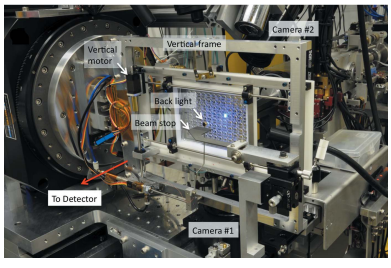
Edited by G. E. Ice, Oak Ridge National Laboratory, USA

Keywords: *in situ*; synchrotron data collection; crystallization plate.

A prototype of a 96-well plate scanner for *in situ* data collection has been developed at the Structural Biology Center (SBC) beamline 19-ID, located at the Advanced Photon Source, USA. The applicability of this instrument for protein crystal diffraction screening and data collection at ambient temperature has been demonstrated. Several different protein crystals, including selenium-labeled, were used for data collection and successful SAD phasing. Without the common procedure of crystal handling and subsequent cryo-cooling for data collection at T = 100 K, crystals in a crystallization buffer show remarkably low mosaicity (<0.1°) until deterioration by radiation damage occurs. Data presented here show that cryo-cooling can cause some unexpected structural changes. Based on the results of this study, the integration of the plate scanner into the 19-ID end-station with automated controls is being prepared. With improvement of hardware and software, *in situ* data collection will become available for the SBC user program including remote access.

1. Introduction

In macromolecular X-ray crystallography (MX) successful structure determination depends upon obtaining good quality single crystals. Current practice consists of a semi-exhaustive search of crystallization space performed to identify conditions for growth of the best crystal (Carter & Carter, 1979; Jancarik *et al.*, 2004; Delucas *et al.*, 2005). The initial exploration of conditions suitable for crystal growth involves screening hundreds of chemical formulations with varying buffers, precipitating agents, ligands, additives, *etc*. The selection of a particular screening set is often supported by databases linking protein properties with crystallization success rates and selecting protein- or complexes-specific screens (Lasala *et al.*, 2015; Grimm *et al.*, 2010; Hoggan *et al.*, 2003; Bulek *et al.*, 2012). The initial crystallization hit is then refined by varying concentrations of constituents, pH, droplet volumes, temperatures and other parameters. In a high-throughput crystallography laboratory, the crystallization screening process is largely automated; for example, liquid handlers mix and dispense solutions and generate crystallization plates, which are subsequently incubated in plate hotels and auto-imaged. These robotized processes are followed by one of the few remaining manual steps in the MX pipeline, *i.e.* crystal preparation for an X-ray diffraction experiment. This is performed either for diffraction screening purposes to identify the most promising crystal-



lization conditions appropriate for further optimization or, if crystals are suitable, collection of a complete diffraction data set.

A crystal is harvested from the crystal growth device, typically the crystallization plate, and is mounted in a capillary (or a special loop holder) for ambient temperature data collection. However, in the past 20 years, most commonly, the crystal is cryo-protected with a suitable solution, mounted in a loop and flash-cooled in liquid nitrogen or sometimes in propane to facilitate cryo-temperature data collection. Virtually all macromolecular crystals are treated this way for use at the high X-ray flux third-generation light sources. This is because cryo-temperatures mitigate X-ray radiation damage to the crystal by slowing the diffusion of free radicals. This allows a much greater X-ray dose to be delivered to the crystal, significantly prolonging its lifetime in the X-ray beam and often giving higher resolution and redundancy data. As a result, the number of crystals required for a complete diffraction data set (or multiple data sets for phasing using anomalous signal) is often reduced to one. It also allows data collection from much smaller specimens including crystals grown *in cellulo* (Ginn *et al.*, 2015; Sawaya *et al.*, 2014; Zhang *et al.*, 2015).

Cryo-crystallography has been a triumphant approach in the last two decades, even though it is not completely flawless. During the flash-cooling procedure, a fragile crystal is subjected to significant stresses, starting with mechanical manipulation through washing with the cryo-protectant solution and ending with rapid transfer from ambient to cryo-temperatures. Alternative approaches, including changing the properties of solvent under high pressure prior to flash-cooling, are not stress-free either (Kim *et al.*, 2013). Such a shock sometimes leads to sample destruction; therefore, the entire process of crystal handling may require careful and time-consuming optimization. In special cases, successful but much more elaborate cryo-procedures have been developed (Luger *et al.*, 1997). Still, some very fragile crystals do not tolerate any manipulation (Wikoff *et al.*, 1998), leading to false negative results in the search for diffraction quality crystals. If no diffraction is observed after crystal cryo-cooling it does not necessarily mean that the crystal is poorly ordered, but rather it may have been mistreated. Therefore, evaluation of crystal quality prior to any handling must be given an important consideration. Moreover, there are a number of reports suggesting that cryo-cooling may in some cases introduce changes in the protein structure leading to incorrect structural or chemical interpretation, affecting biochemically relevant conclusions (Juers *et al.*, 2007). In fact, it has been shown that cryo-cooling not only reduces intrinsic conformational heterogeneity of protein molecules but it also introduces new conformational states (Keedy *et al.*, 2014; Fraser *et al.*, 2011). These changes are associated with or even induced by crystal shrinking, leading to an increase in the intermolecular contact interface (Juers & Matthews, 2004, 2001) or movement of domains within the protein molecule (Petrova *et al.*, 2009, 2010). In some extreme cases, the crystals undergo phase transition upon cooling and the molecules re-pack into a different space group (Campobasso *et al.*, 1998; Skrzypczak-Jankun *et al.*, 1996).

In recent years, there has been a growing interest in collecting X-ray diffraction data directly from the crystallization devices to completely eliminate the crystal harvesting step. A number of successful tests have been already performed at different light sources and with an in-house instrument (PX Scanner, Agilent Technologies, Santa Clara, CA, USA). These experiments examined the usability of crystallization microcapillaries (Yadav *et al.*, 2005), microfluidic devices (Heymann *et al.*, 2014; Maeki *et al.*, 2012; Perry *et al.*, 2013), chips (Kisselman *et al.*, 2011), micromeshes (Yin *et al.*, 2014) and regular crystallization plates (Axford *et al.*, 2012; Bingel-Erlenmeyer *et al.*, 2011). At some synchrotrons, *in situ*-dedicated platforms already exist. For example, the Swiss Light Source operates a beamline adjacent to the crystallization facility that enables seamless transition from crystals to diffraction (Bingel-Erlenmeyer *et al.*, 2011). At other light sources, such facilities are being constructed or planned (VMXi at the Diamond Light Source or MX3D at the Australian Synchrotron). Currently the Protein Data Bank (PDB) contains only a few deposits that could be identified as resulting from *in situ* data collection (le Maire *et al.*, 2011). However, success of these initial attempts, which include not only collection of complete data sets but also experimental structure phasing through single-wavelength anomalous dispersion (SAD) (Kisselman *et al.*, 2011; Perry *et al.*, 2013), strongly suggest that in the near future *in situ* data collection at ambient temperatures may be routinely applied. Such an approach will reduce cooling-introduced bias, increase throughput and possibly the success rate of macromolecular crystallography and will allow studies of macromolecules closer to physiological conditions. In addition, *in situ* data collection at ambient temperature is being used almost routinely in serial crystallography at free-electron lasers and third-generation light sources (Spence, 2015; Coquelle *et al.*, 2015).

At the Advanced Photon Source, one of the most productive facilities for macromolecular crystallography worldwide (Kuller *et al.*, 2002), the majority of data from macromolecular crystals are collected at cryo-temperatures (90%, X-ray PDB entries, temperature $\leq$120 K). There were some initial attempts to collect data *in situ* at the Structural Biology Center (SBC), LS- and GM/CA-CATs (Perry *et al.*, 2013; Yadav *et al.*, 2005), but a complete system available for the user community does not exist. Of particular interest is the possibility of data collection directly from crystallization plates and microfluidic devices. With the development of various robotic systems for plate setup and storage, the crystallization plates have been standardized to achieve compatibility with automatic platforms. Good progress has been made to provide plates with low X-ray absorption profile, and scattering properties, such as the Microlytic Crystal Former, the Greiner CrystalQuick™X Plate, or the MiTeGen *In Situ*–1™ Plate (Aller *et al.*, 2015; Bingel-Erlenmeyer *et al.*, 2011).

Therefore, at the SBC we have designed and built a prototype of a crystal plate scanner that can accept any

commercially available 96-well Society for Biomolecular Screening (SBS) format crystallization plates. Here we demonstrate applicability of this instrument for diffraction screening and data collection using three different protein crystals, including selenium-labeled for SAD phasing. One protein is β-*N*-acetylhexosaminidase from *Bacteroides fragilis* belonging to glycosyl hydrolase (GH) family 20 (GH20) (Lombard *et al.*, 2014) and will be referred to as *Bf*GH20 in the following description and discussions. The structure of *Bf*GH20 was determined earlier under cryo-temperature and deposited in the PDB with entry code 4PYS. The second protein is a functionally uncharacterized GH3 family member of periplasmic β-glucosidases from *Beutenbergia cavernae* strain DSM 12333. It is closely related to NagZ β-*N*-acetylglucosaminidases (Litzinger *et al.*, 2010). It carries a sequence motif KHFPGLGRVTENTD that is similar to the conserved sequence motif KH(F/I)PG(H/L)GXXXXD(S/T)H of NagZ β-*N*-acetylglucosaminidase but it lacks a key Asp–His dyad for catalysis (Litzinger *et al.*, 2010). For descriptive convenience, we named the protein *Bc*GH3. The crystal structure of *Bc*GH3 determined with data collected at cryo-temperature was deposited in the PDB with entry code 5BU9. The prototype scanner also enabled atomic-resolution data collection (1.23 Å) from lysozyme crystals, which are commonly used as a standard to evaluate an X-ray diffraction system.

2. Experimental procedures

2.1. Plate scanner for *in situ* data collection

The prototype crystal plate scanner, designed to accept 96-well SBS format plates, was installed on the SBC 19-ID omega axis (Fig. 1) replacing the existing spindle used for traditional crystallography. Data were collected and stored using the data acquisition system and software that are already in place. In order to accommodate the SBS crystallization plates, the scanner was designed with 100 mm × 140 mm (V × H) travel, with an additional ±5 mm movement along the beam direction to ensure complete focus control within the plates. Rotation of the device about the omega axis is constrained by the on-axis camera mirror and the beam stop, located on either side of the scanner. The total available rotation range was −5 to +40°. The scanner uses stepper motors capable of better than 0.1 µm step resolution. The *x* and *y* axes are spring-loaded using torsion springs to minimize backlash. Each plate is loaded into the scanner by hand and held in place using spring-loaded blocks in a custom plastic frame. The scanner is constructed from lightweight materials to reduce positional errors caused by weight imbalance.

In order to put each individual crystal on the ω axis rotation center, each crystal was aligned using two cameras (camera No. 1 and No. 2), located 37.5° apart (Fig. 1*a*). A moveable LED backlight was also provided to enhance visualization of the crystals within droplets during alignment. The scanner was operated manually through a pendant-type device, which allowed access to all three axes. To accommodate long moves within the plate, the pendant has three different resolution

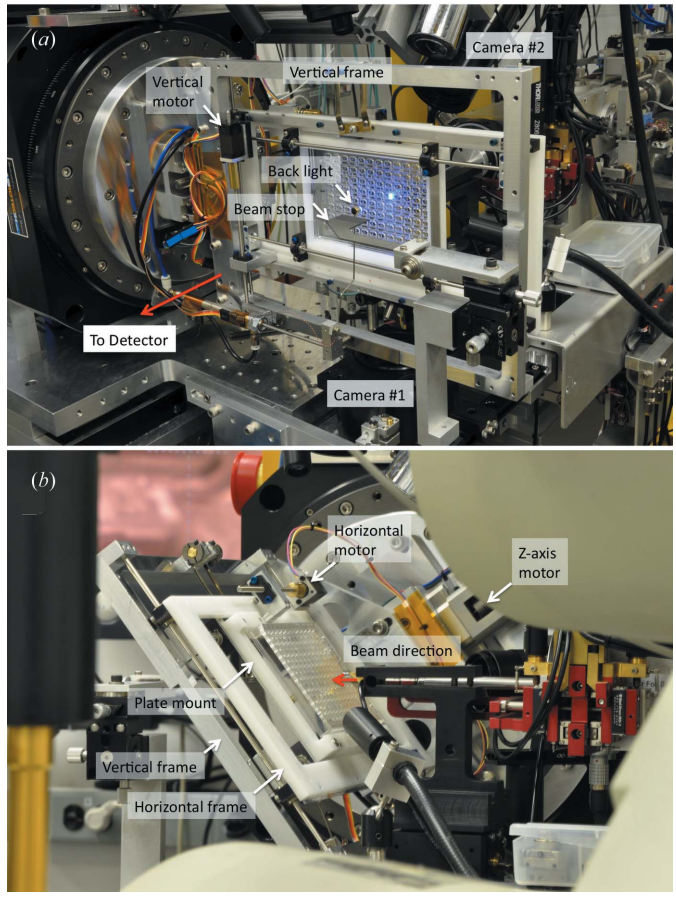


Figure 1
The setup of a 96-well plate scanner for *in situ* data collection. (*a*) A back view of the plate scanner mounted on the beamline goniometer (omega axis) at the rotation angle of $\omega = 0°$. (*b*) A side front view of the plate scanner at its maximum tilt angle, $\omega = 40°$. The total rotation range established was 45°, from −5° to +40°.

settings, allowing coarse-, fine- and high-resolution positioning accuracy, with a proportionate increase in alignment speed depending upon the resolution setting.

The diffraction data were recorded on the Area Detector Systems Corporation Quantum 315r (Q315r) 3 × 3 CCD area detector in the hardware-binned mode. This detector has an active area of 315 mm × 315 mm, with a resolution of 6144 × 6144 pixels and 51 µm × 51 µm pixel size. It has 1 s readout time (16 bit) and a dynamic range of 65535. When operated in full image mode it generates ∼75 Mb images and in the hardware-binned mode ∼18 Mb images. Calibrations of the detector parameters have been performed initially using silicon powder and refined using single-crystal lysozyme diffraction with standard lattice constants and processed *via* *HKL3000* software (Minor *et al.*, 2006).

2.2. Protein cloning, expression and purification

The *Bf*GH20 gene was synthesized chemically and initially cloned into vector pUC57. Subsequently, the construct with sequence range 16–518 was amplified with KOD DNA polymerase using conditions and reagents provided by Novagen (Madison, WI, USA) and cloned into the pMCSG68 vector

according to the ligation-independent procedure (Aslanidis & de Jong, 1990; Eschenfeldt *et al.*, 2009) and transformed into the *E. coli* BL21(DE3)-Gold strain (Agilent Technologies, Santa Clara, CA, USA). This construct provided a system to produce a fusion protein containing an N-terminal His_6-tag followed by a TEV protease cleavage site and a target protein. Similarly, a *Bc*GH3 gene fragment encoding residues 72–408 was cloned into the pMCSG68 vector and subsequently transformed into the *E. coli* BL21(DE3) Magic cells.

The starter cultures were grown at 310 K overnight in 500 ml polyethylene terephthalate beverage bottles containing 25 mL of non-sterile modified M9 salts 'pink' medium (Donnelly *et al.*, 2006). Each culture was then transferred to a 2 L polyethylene terephthalate beverage bottle containing 1 L of M9 'pink' media. Then, the bacterial cultures were grown at 310 K, 200 rpm until they reached OD_{600} = 1.0 (for *Bf*GH20) or 1.4 (for *Bc*GH3). Methionine biosynthetic inhibitory amino acids (25 mg L^{-1} each of L-valine, L-isoleucine, L-leucine, L-lysine, L-threonine, L-phenylalanine) and 90 mg L^{-1} of L-selenomethionine (SeMet, Medicillin, catalog number MD045004D) were added. After air-cooling the cultures down to 277 K for 60 min, protein expression was induced by 0.5 m*M* or 1 m*M* isopropyl-β-D-thiogalactoside (IPTG) for *Bf*GH20 and *Bc*GH3 proteins, respectively. The cells were incubated overnight at 291 K, harvested and re-suspended in 30 mL of lysis buffer [500 m*M* NaCl, 5% (*v*/*v*) glycerol, 50 m*M* HEPES pH 8.0, 20 m*M* imidazole, and 10 m*M* β-mercaptoethanol plus one protease inhibitor cocktail tablet (Complete, Roche, Indianapolis, IN, USA)]. Cells were disrupted by lysozyme treatment (1 mg mL^{-1}) and sonication, and the insoluble cellular material was removed by centrifugation, 30000 g for 60 min, followed by filtration through 0.45 µm syringe filters. The SeMet-labeled protein was purified from other contaminating proteins using Ni-NTA affinity chromatography and the ÄKTAxpress system (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) with the addition of 10 m*M* β-mercaptoethanol in all buffers as described previously. This was followed by the cleavage of the His_6-tag using recombinant His_7-tagged TEV protease and an additional step of Ni-NTA affinity chromatography was performed to remove the protease, uncut protein and affinity tag. For the *Bc*GH3 protein, size-exclusion chromatography was also performed. The pure proteins were concentrated using Amicon Ultra centrifugal filters (Millipore, Bedford, MA, USA) in 20 m*M* HEPES pH 8.0 buffer, 250 m*M* NaCl, and 2 m*M* dithiothreitol (DTT). Protein concentrations were determined from the absorbance at 280 nm using a molar absorption coefficient calculated by the method developed by Gill & von Hippel (1989). The concentration of protein samples used for crystallization was ~75 mg mL^{-1} and 113 mg mL^{-1} for *Bf*GH20 and *Bc*GH3, respectively.

2.3. Protein crystallization

The MiTeGen (Ithaca, NY, USA) *In Situ*-1™ sitting-drop vapor-diffusion crystallization 96-well SBS format plate was chosen for crystallization experiments and subsequently for diffraction data collection at room temperature (T = 295 K) because of low X-ray absorption and scattering in comparison with similar commercial plates (Aller *et al.*, 2015). The crystallization setting was performed with the help of the Mosquito liquid dispenser (TTP Labtech, Cambridge, MA, USA) using the sitting-drop vapor-diffusion technique at T = 289 K. Our standard crystallization set-up was adapted for the *In Situ*-1™ crystallization plates with 35 µL of crystallization buffer in the reservoir for equilibration. These plates are very similar to 96-well SBS plates from Greiner (Greiner Bio-One, Monroe, NC, USA) that are typically used for crystallization screening. Within each of the 96 wells, 0.4 µL of protein and 0.4 µL of reservoir solution were mixed. The volume of 0.4 µL of protein solution was used to ensure the liquid dispenser performs consistently without missing any drops; the droplets do not evaporate during the procedure, a few nucleations are promoted and a minimum amount of protein is used without sacrificing the size of the protein crystals. This follows the standard protocol developed for the Midwest Center for Structural Genomics (MCSG) crystallization screens (Kim *et al.*, 2011). Such volume of droplets often results in crystals of sufficient size for the collection of complete data sets under cryo-conditions. The geometry of the *In Situ*-1™ plate prevents any crystallization drop movement out of the wells and prohibits any cross-contamination of reservoir solutions. The crystallization condition of each SeMet-labeled protein was adapted from one of the known conditions identified in the MCSG database. All crystals were grown at T = 289 K *via* vapor diffusion in the *In Situ*-1™ plates. For *Bf*GH20, the protein concentration was 75 mg mL^{-1} and the crystallization formulation consisted of 0.2 *M* $MgCl_2$, 0.1 *M* sodium cacodylate:HCl, pH 6.5, 10% PEG3350. The data collection quality crystals grew within two days and reached sizes of approximately 0.20 mm × 0.12 mm × 0.20 mm. For *Bc*GH3, the crystallization formulation contained 0.2 *M* sodium malonate, pH 7.0 and 20% PEG3350 and crystals of data collection quality reached sizes of approximately 0.52 mm × 0.25 mm × 0.25 mm within a few days. The lysozyme was purchased from Sigma Inc. and was crystallized with a protein concentration of 75 mg mL^{-1} in 0.1 *M* sodium acetate, pH 4.8 and a reservoir containing 0.1 *M* sodium acetate, pH 4.8 and 5–7.75% NaCl – conditions used to grow high-quality crystals for beamline testing. The crystals reached sizes of approximately 1.00 mm × 0.40 mm × 0.40 mm within a few days. In prior diffraction data collection at T = 100 K, crystals of *Bf*GH20 and *Bc*GH3 were cryo-protected with amended crystallization solution containing 25% glycerol. The *Bf*GH20 crystals were washed with cryo-solution for 60–120 s prior to plunging into liquid nitrogen while *Bc*GH3 was treated with cryo-solution for a couple of seconds. Both crystals were mounted into LithoLoops (Molecular Dimensions, Altamonte Springs, FL, USA).

2.4. Data collection

Data were collected at 19-ID using the *SBCcollect* data acquisition system. Crystal alignment was performed manually at two ω orientations. First, at ω = 0°, the desired crystal is

Table 1
Data collection protocol, image processing, structure phasing and refinement statistics.

Refinement was performed only for more complete *in situ* data sets. N/A = not applicable.

Column	1	2	3	4	5	6	7
Project	*Bf*GH20	*Bf*GH20	*Bf*GH20	*Bc*GH3	*Bc*GH3	*Bc*GH3	Lysozyme
Data collection	*In situ*	*In situ*	Cryo-cooling	*In situ*	*In situ*	Cryo-cooling	*In situ*
Crystal(s)/data set(s)†	1/1	3/3	1/1	1/11	2/7	1/1	3/8
Rotation (°)/frames‡	1/30	1/105	0.2/300	1/163	1/104	1/180	82
Temperature (K)	295	295	100	295	295	100	295
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P3_121$	$P3_121$	$P3_121$	$P4_32_12$
Cell dimensions (Å)	$a = b = 97.93$, $c = 317.8$	$a = b = 98.00$, $c = 318.0$	$a = b = 96.8$, $c = 316.0$	$a = b = 147.9$, $c = 89.85$	$a = b = 147.8$, $c = 89.92$	$a = b = 146.5$, $c = 85.49$	$a = b = 79.3$, $c = 37.9$
Wavelength (Å)	0.9792	0.9792	0.9791	0.9793	0.9793	0.9792	0.9792
Beam size (µm)	50 × 80	50 × 80	75 × 75	50 × 50	50 × 50	75 × 75	50 × 80
Flux (photons s^{-1})	6.62×10^{10}	3.93×10^{10}	2.64×10^{11}	4.43×10^{10}	4.31×10^{10}	1.05×10^{11}	6.16×10^{10}
Resolution (Å)§	2.45 (2.45–2.49)	2.45 (2.45–2.49)	1.82 (1.82–1.85)	2.85 (2.85–2.90)	2.85 (2.85–2.90)	2.25 (2.25–2.29)	1.23 (1.23–1.25)
Unique reflections	50483	57593	134406	22436	26403	50001	35273
R_{merge}¶	0.102 (0.737)	0.102 (0.671)	0.094 (0.634)	0.088 (0.795)	0.092 (0.835)	0.106 (0.814)	0.067 (0.834)
R_{pim}	0.068 (0.546)	0.050 (0.367)	0.050 (0.341)	0.026 (0.236)	0.041 (0.383)	0.036 (0.275)	0.030 (0.503)
$CC_{1/2}$	0.521	0.715	0.769	0.829	0.613	0.877	0.557
$\langle I\rangle/\langle\sigma I\rangle$	8.19 (1.40)	11.42 (2.03)	32.00 (2.60)	28.28 (3.80)	17.46 (2.21)	21.46 (2.67)	25.43 (1.37)
Completeness (%)	86.9 (91.8)	99.0 (98.7)	99.6 (99.9)	85.3 (83.8)††	98.6 (99.7)	100.0 (100.0)	98.9 (92.3)
Redundancy	2.7 (2.7)	4.6 (4.1)	4.5 (4.3)	11.9 (11.7)	5.7 (5.5)	9.6 (9.7)	5.6 (3.0)
Mosaicity (°)	0.04–0.20	0.04–0.16	0.18–0.30	0.06–0.23	0.06–1.11	1.16–1.76	0.016–0.072
Solvent (%)	62.2	62.3	60.7	70.2	70.2	68.2	
SAD phasing††							
CullR_ano (%)	0.90	0.87	0.73	0.74	0.82	0.80	N/A
FOM‡‡	0.12	0.15	0.26	0.28	0.24	0.25	N/A
Model building							
Mainchain (%)	81	97	91	98	99	99	N/A
Sidechain (%)	30	94	86	94	97	97	N/A
R_{work}/R_{free}§§	0.470/0.533	0.212/0.262	0.262/0.287	0.224/0.293	0.192/0.237	0.234/0.274	N/A
FOM¶¶	0.504	0.793	0.811	0.820	0.862	0.811	N/A
Refinement†††							
R_{work}/R_{free}	N/A	0.149/0.186	0.165/0.190	N/A	0.137/0.186	0.158/0.198	N/A
No. of atoms protein/water	N/A	7912/206	7689/604	N/A	4699/10	4829/415	N/A
Average *B*-factor (Å^2) protein/water	N/A	39.4/38.8	38.1/43.5	N/A	53.54/55.73	36.6/44.4	N/A
RMSD							
Bond lengths (Å)	N/A	0.008	0.007	N/A	0.008	0.014	N/A
Bond angles (°)	N/A	1.083	1.071	N/A	1.091	1.346	N/A
Ramachandran statistics of φ/ψ angles (%)							
Most favored/outliers	N/A	97.1/0.2	96.8/0.2	N/A	96.0/0.0	96.8/0.0	N/A

† Number of crystal(s) and the total number of spots on the crystal(s) used for data collection. ‡ Rotation angle per each frame and the total number of frames used for data merging. § Values in parentheses correspond to the highest resolution shell. ¶ $R_{merge} = \Sigma h \Sigma j |Ihj - \langle Ih\rangle| / \Sigma h \Sigma j Ihj$, where *Ihj* is the intensity of observation *j* of reflection *h*. †† Full data set was used for phasing. ‡‡ Figure of merit (FOM) value before density modification (DM) and non-crystallographic symmetry (NCS) averaging. §§ $R = \Sigma h |F_o| - |F_c| / \Sigma h |F_o|$ for all reflections, where F_o and F_c are observed and calculated structure factors, respectively. R_{free} is calculated analogously for the test reflections, randomly selected and excluded from the refinement. ¶¶ FOM value after automatic model building using the program *Buccaneer*. †††Full data set was used for refinement, in which 5% of reflections were set aside for R_{free} calculation.

moved into position (Fig. 1*a*). At $\omega = 0°$, the plate is positioned at a 90° angle to the on-axis camera. This allows crystal orientation to be set perpendicular to the beam axis and parallel to the translation direction of the motorized plate assembly. Next, ω is rotated to 37°, allowing the vertically offset camera to view the sample at a 90° angle to the plate. Adjusting crystal alignment with this viewing angle allows corrections to the translation direction along the beam. One or two iterations are necessary to achieve optimal alignment due to the limited rotation angles available.

The diffraction experiments were carried out at the ambient temperature of the 19-ID beamline ($T \approx 295$ K operating temperature) and X-ray photon energy of 12.66 keV using beam sizes 50 × 50, 50 × 80 or 75 × 75 (V × H) (µm) and photon fluxes listed in Table 1. The images were recorded on the Q315r CCD area detector in the hardware-binned mode and 1 s per degree exposure. For each project, an initial set of data at full rotation (45°) would be recommended to be collected for an assessment of radiation damage of the crystal (Fig. 2). For data collection the X-ray flux was selected based on initial tests for each crystal type and the beam was attenuated as necessary. In the case of a large crystal, data were collected from an array of multiple sites on the single crystal in a staggering grid mode for starting spindle angle (Fig. 3).

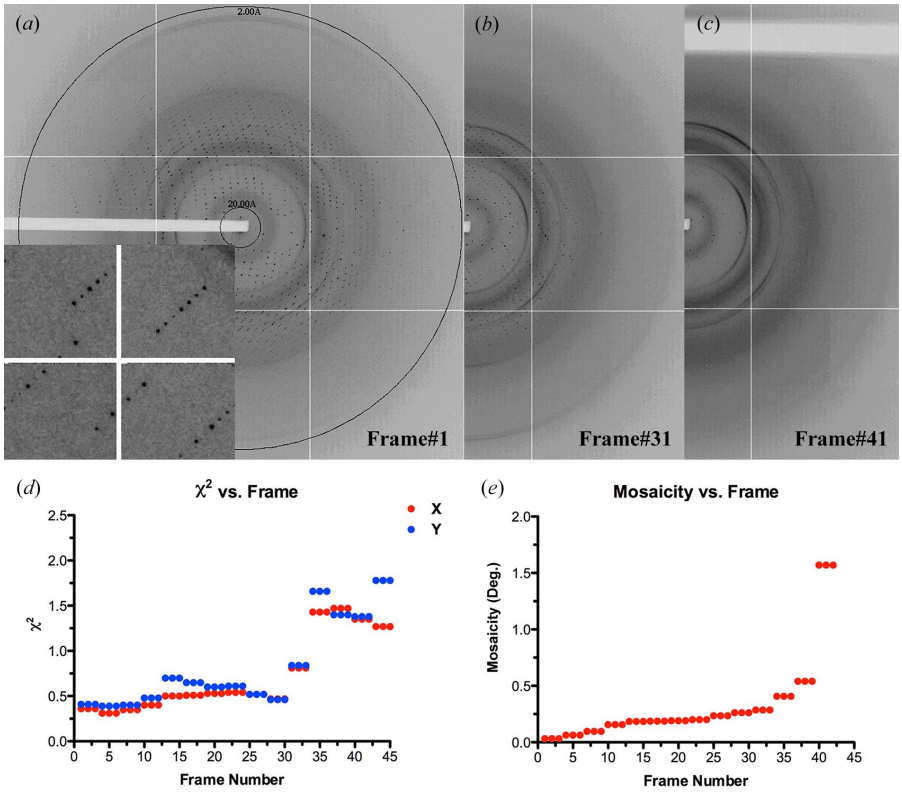


Figure 2
A typical test run of *in situ* data collection. (*a*)–(*c*) Selected diffraction patterns of *Bf*GH20 in a sweep of exposures of 45° rotation. The first frame represents the initial diffraction pattern of the crystal. The 31st frame records an onset of radiation damage. The 41st frame shows that the crystal was nearly destroyed by radiation damage. (*d*) χ^2 values and (*e*) mosaicity *versus* frame number. These are two examples of refined parameters obtained from *in situ* data integration and indicators of crystal radiation damage.

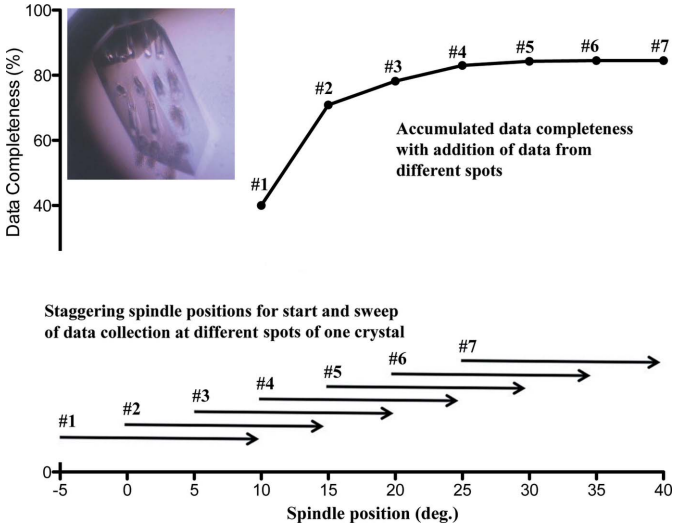


Figure 3
A data collection strategy from multiple spots on one crystal. On each spot of a crystal, a partial data set covering only a 15° rotation was collected to avoid severe radiation damage. The starting spindle position was staggered with a 5° increment on the next spot. Data from seven spots cover the maximum rotation of 45° for the plate scanner. The top of the figure shows the increased data completeness with data from each spot being added for a typical run of a *Bc*GH3 crystal. The insert photograph shows an array of burned spots resulting from radiation damage after sequential data collection on them.

Commonly, data sets from several crystals in different orientations within one droplet or from different droplets were collected to achieve data completeness. Such a strategy was necessary for two reasons: (i) fast crystal decay due to radiation damage, which is much faster at T = 295 K than at the typical cryotemperature of T = 100 K, and (ii) hardware limitations imposing the reduced rotation range. The setup for cryogenic data collection of *Bf*GH20 and *Bc*CH3 mounted on loops is the standard setup at the 19-ID beamline (Rosenbaum *et al.*, 2006).

2.5. Data processing and structure solution

Data were processed with the *HKL3000* software suite (Minor *et al.*, 2006). Data sets were first scaled and merged for the assessment of their isomorphism. Non-isomorphous frame(s) or even partial data set(s) were excluded in the next round of scale and merge. Only data sets with diffraction intensities that were obtained under low-dose conditions (less than the Garman limit 2 MGy) were included in calculations. Pairwise partial data sets were analyzed for their consistency. To combine data from multiple crystals, re-indexing of reflections from one crystal or crystals was sometimes necessary. Once a combined data set reached satisfactory completeness, *SHELXD* was used for searching heavy-atom sites. A clear solution of the substructure is commonly an indicator of good data quality and an indicator of success in SAD phasing. *MLPHARE* (Otwinowski, 1991) was used for phasing, *DM* (Cowtan, 1994) for density modification and *ARP*/*wARP* (Langer *et al.*, 2008) as well as *Buccaneer* (Cowtan, 2006) for automated model building. The final models were obtained through alternating manual rebuilding in *COOT* (Emsley & Cowtan, 2004) and crystallographic refinement in *Phenix.refine* (Adams *et al.*, 2010). An identical protocol was used for structures determined at T = 100 K with the exception that complete data sets were collected from one crystal. The refinement statistics are shown in Table 1.

3. Results and discussion

3.1. *In situ* crystal diffraction and data collection

The majority of unscathed crystals within the crystallization plate diffracted well at T = 295 K, but decayed rapidly upon X-ray exposure allowing for a few degrees of data to be collected from a single spot on the crystal (Fig. 2). Data processed from initial images always showed remarkably low

mosaicities (<0.1°) of these crystals, indicating that *in situ* data collection avoids crystal lattice distortion introduced during crystal handling, cryo-protectant treatment and flash-cooling. Crystals grown under the same conditions (from the same or different droplet) show very similar diffraction properties and radiation damage rates. The most significant source of non-isomorphism in the *in situ* data originates from X-ray dose-dependent radiation damage. In some crystals, visible tracks of gas bubbles appear along the X-ray path (Fig. 3). The crystal diffraction pattern quickly vanishes, accompanied by an increase in mosaicity, $\chi^2(x,y)$ and other crystal parameters (Fig. 2). A preliminary test run generally provides valuable information on the radiation sensitivity of the crystal and helps design an effective strategy for subsequent data collection. As a result, high quality, nearly complete and redundant diffraction data sets could be collected *in situ* from crystals of both *Bf*GH20 and *Bc*GH3, and these structures were phased with an anomalous signal from selenium atoms (Table 1).

3.2. Phasing from a single data set obtained from a single crystal

Limited rotation range results in an incomplete data set from a single-crystal orientation. Nevertheless, the high quality of this partial data set still enables successful structure phasing, especially for properly oriented crystals with high symmetry. For example, the SeMet-labeled protein *Bf*GH20 crystallized in space group $P4_12_12$. Typical data-collection statistics are shown in Fig. 2. When only the first 30 frames (corresponding to 30°) of data were used for scaling and merging, a data set with 86.9% completeness was obtained (Table 1, column 1). By using this partial data set at a 2.45 Å resolution limit cutoff, the structure of the *Bf*GH20 was easily solved.

3.3. Phasing using multiple data sets from a single crystal

When the sample is more susceptible to radiation damage, multiple data sets are needed to obtain sufficient completeness. If the crystal is large enough and belongs to a high-symmetry space group, the data can be collected from several different sites by translating the crystal in a grid format (Fig. 3). The separation of two spots should be no less than 25 µm to avoid overlap of the radiation-damaged zone from preceding exposures. The starting rotation angle is staggered to maximize the coverage of data within the limited rotation range. Such an approach was utilized for the SeMet-labeled protein *Bc*GH3 that crystallized in space group $P3_121$ (Table 1, column 4). Some of these crystals reached dimensions of ~0.25 mm × 0.25 mm × 0.52 mm. Typically, one large crystal enabled us to collect data from more than 16 sites using a 50 µm × 50 µm X-ray beam (Fig. 3). At each site, 15° of data were collected. A staggering rotation angle of 5° was used for starting data collection at the next spot (Fig. 3). When the maximum range of 45° is covered, the second cycle of staggered data collection is started. In the example given here, 16 data sets were collected from different sites on one crystal. After data processing and analysis, 11 out of the 16 data sets were used for final data merging, resulting in a partial data set with 85.3% completeness with a resolution limit to 2.85 Å. The resulting electron-density maps after phasing and density modification (DM) were easily interpreted at this moderate resolution limit and *Bc*GH3 model building was completed; results are shown in column 4 of Table 1.

3.4. Phasing from multiple data sets obtained from multiple crystals

Typically, a complete *in situ* data set requires multiple crystals in different orientations, particularly in the case of low crystal symmetry. Multiple crystals of different orientations could be obtained from the same crystallization well or different wells. In the case of *Bf*GH20, in addition to the data from the crystal described earlier, two more crystals from the same well were used for data collection. A combination of data from these three crystals completes a data set and greatly improved the anomalous signal for phasing (99.0% completeness to 2.45 Å resolution) (Table 1, column 2). In the case of *Bc*GH3, an example of a complete data set was obtained by combining data from seven data sets collected from two crystals (Table 1, column 5). The refinement of the *Bc*GH3 structure from the data set resulted in satisfactory statistics.

3.5. Comparison of *Bf*GH20 structures from *in situ* and cryo-conditions

Cryo-protectant treatment and flash-cooling commonly cause shrinking of the crystal lattice. Together with the mechanical forces applied during crystal harvesting, these common procedures often lead to increased crystal mosaicity. It has been reported that high mosaicity reduces the signal-to-noise ratio (Helliwell, 2005), which plays an important role in structure phasing, particularly in phasing with a weak anomalous signal, such as from sulfur atoms. On the other hand, flash-cooling has the potential to make protein and solvent molecules more ordered and thus helps to increase diffraction limits. Realistically, the diffraction quality of a protein crystal after cryo-cooling largely depends on a balance between these two-sided effects. Flash-cooling may also impact static and dynamic disorder. By comparing the *Bf*GH20 structures obtained using *in situ* data at T = 295 K and data obtained at cryo-temperature (T = 100 K), we can visualize these effects. It appears that cryo-cooling changed some local structures of the protein but at the same time it helped to reach a higher-resolution limit and reduce B-factors (Table 1, column 3).

The unit cell of *Bf*GH20 shrank by about 2% in volume after cryo-cooling, primarily along the a and b axes. There are two molecules in one asymmetric unit. A pairwise superposition of a *Bf*GH20 structure from *in situ* data and a *Bf*GH20 structure from data at T = 100 K results in RMSD values ranging between 0.51 and 0.61 Å, indicating no overall structural change caused by crystal handling and subsequent flash-cooling. There is also no relative movement between its N-terminal α/β domain (D1) and C-terminal $(\beta/\alpha)_8$ catalytic domain (D2) in either of the *Bf*GH20 molecules (Fig. 4).

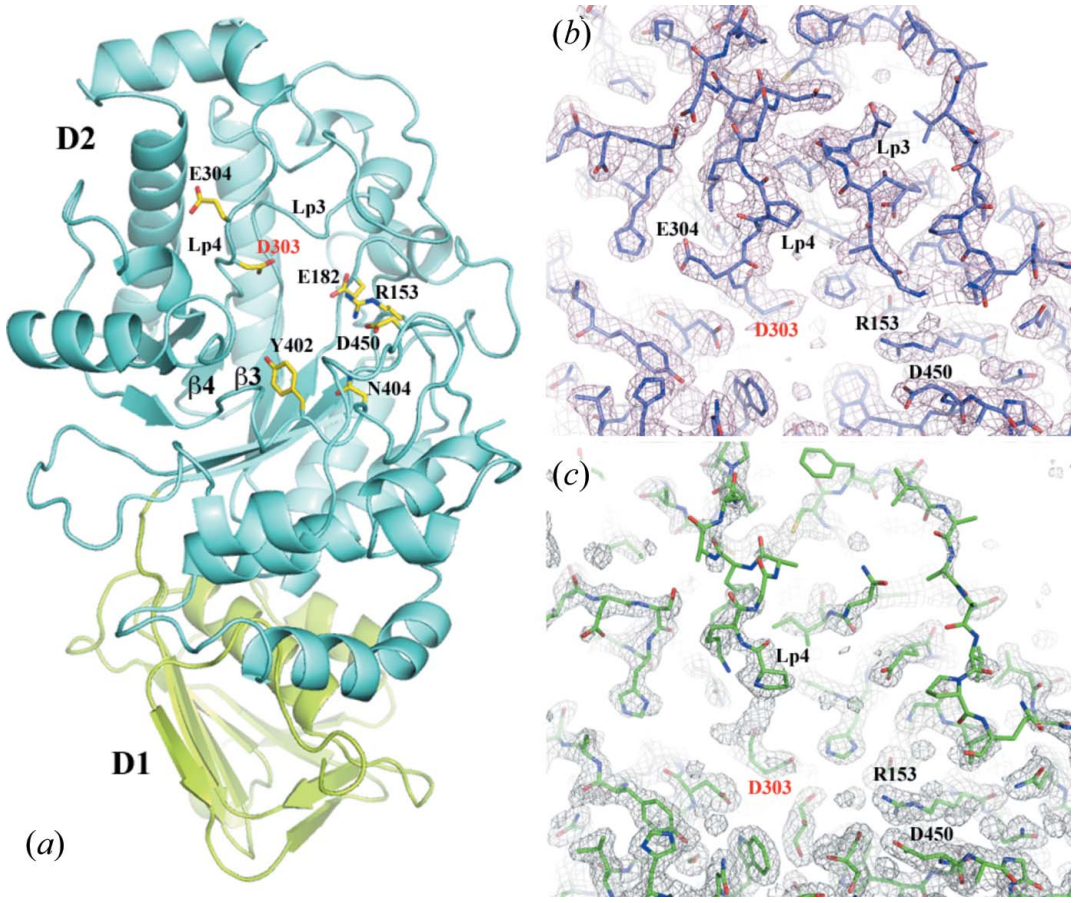


Figure 4
Comparison of the structures of *Bf*GH20 from *in situ* T = 295 K and cryo T = 100 K data. (*a*) A ribbon diagram of *Bf*GH20, which consists of a N-terminal α/β domain (D1) and a C-terminal $(\beta/\alpha)_8$ catalytic domain (D2). Residues forming the active site of the catalytic domain are drawn in stick format and labeled. The presumably catalytic residue D303 is highlighted in red. The two loops in discussion, Lp3 and Lp4, are the loop between β3 and α3 and the loop between β4 and α4 of the $(\beta/\alpha)_8$ barrel, respectively. (*b*) The active site of the *Bf*GH20 structure from *in situ* T = 295 K data. All residues in the view are drawn in stick format. The $2F_o - F_c$ electron-density map in purple mesh is contoured at 1σ. (*c*) The active site of the *Bf*GH20 structure from cryo T = 100 K data. The $2F_o - F_c$ electron-density map in gray mesh is also contoured at 1σ. However, parts of Lp3 and Lp4 of the structure are disordered.

However, in the structure determined at T = 100 K there is partial and complete disorder of both the main chain and the side chains forming the active site. In the *in situ* structure at T = 295 K, the corresponding region is well ordered despite relatively elevated temperature factors. Thus, the disordering of the part of the active site in the cryo-temperature structure might be a result of the alterations introduced during the crystal harvesting and/or flash-cooling.

3.6. Comparison of *Bc*GH3 structures from *in situ* and cryo-conditions

The protein *Bc*GH3 crystallized in the space group $P3_121$ at T = 295 K *in situ* and remains in the same space group after cryo-cooling (T = 100 K). The cell parameters, however, shrank by 0.91% and 4.4% along the *a*- and *c*-axes, respectively (Table 1, column 6). This was accompanied by a small relative rotation between the two molecules in one asymmetric unit of the crystal (data not shown). *Bc*GH3 consists only of one $(\beta/\alpha)_8$ barrel domain. There is no main chain break in either the *in situ* or the cryo structural model. Pairwise superposition of any two molecules between *in situ* and cryo structures or between two molecules within the same structure gives no clues of any significant conformational changes of *Bc*GH3 upon cryo-cooling. However, changes in solvent structure and conformation of side chains are observed.

Fig. 5(*a*) shows the model and associated electron density in the *Bc*GH3 structure from *in situ* T = 295 K data. In the asymmetric unit, the two *Bc*GH3 molecules have significantly different average *B*-factors, 66.9 Å^2 (molecule A) *versus* 42.2 Å^2 (molecule B). The region shown in Fig. 5 is the C-terminal end of the eighth β-strand of the $(\beta/\alpha)_8$ barrel of molecule B, where the active site of a GH3 family member is located. The protein model fits the electron-density map very well at this moderate resolution, 2.85 Å. Only one water molecule can be assigned, which forms two hydrogen bonds to protein atoms, including one to D336. In the T = 100 K cryo structure refined to 2.25 Å resolution, the *B*-factors of two corresponding *Bc*GH3 molecules dropped to 38.2 Å^2 and 29.9 Å^2, respectively. A number of ordered water molecules and an electron-density peak interpreted as a glycerol molecule with a double conformation in the middle of the $(\beta/\alpha)_8$

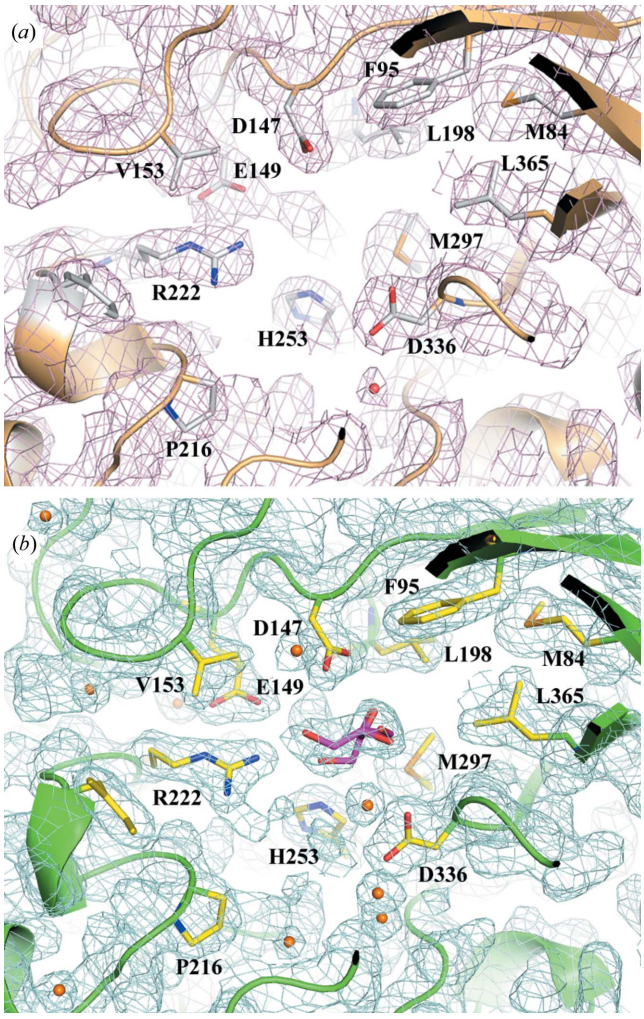


Figure 5
Comparison of the structures of *Bc*GH3 from *in situ* and cryo-cooling data. (*a*) A ribbon diagram of the presumably active site of the *Bc*GH3 structure from *in situ* data. Most residues lining the active site are drawn in stick format. Magenta mesh shows $2F_o - F_c$ electron-density map contoured at 1σ. Only one water molecule drawn as a small sphere is assigned, which forms two hydrogen bonds to protein atoms including an interaction to D336. (*b*) A ribbon diagram of the same active site of the *Bc*GH3 structure from cryo-cooling data. Green mesh shows $2F_o - F_c$ electron-density map contoured at 1σ. A number of water molecules mediate a hydrogen-bond network. A new peak in the electron density in the active site is interpreted as a glycerol molecule with a double conformation. The glycerol molecule is colored with its C atoms in magenta.

barrel are visible. Reduced motion of solvent molecules in the active site is also reflected in decreased *B*-factors. In the structure obtained at $T = 100$ K these water molecules form a hydrogen-bonding network, often observed in protein structures at cryo-temperatures. The question remains as to whether this well defined, water-mediated contact is identical with the protein–solvent interaction at $T = 295$ K. In the *in situ* structure, water or other molecules present in the solvent seem to be more mobile or dynamic. Further, the trapping of solvent molecules into the active site by cryo-cooling changed the conformation of active site residues. For example, F95 and D147 apparently show rotation upon binding the glycerol molecule at the active site (Fig. 5*b*). It is not clear whether the conformational changes represent potential conformational change upon substrate binding or whether it is simply an artifact introduced by cryo-cooling. Clearly, more comparative studies are needed.

3.7. Lysozyme crystal diffraction

As a most commonly used standard protein crystal for evaluation of an X-ray diffraction system, crystalline lysozyme was also used for assessment of our plate scanner performance. A high-quality atomic-resolution data set, complete out to 1.23 Å resolution, was obtained by merging eight data sets collected from three crystals (Table 1, column 7).

4. Conclusions

Our results show that *in situ* diffraction experiments from crystallization plates can be carried out at SBC 19-ID for crystal screening. In some cases, complete and redundant data sets can be collected and successful SAD phasing experiments performed on crystals examined during screening experiments. However, there are still some technical challenges to overcome in both hardware and software, especially with crystal alignment to the X-ray beam. In principle, this requires diffraction-based centering evaluation to overcome optical refraction effects from the plate and crystallization droplet. Development of software for automated drop and crystal detection and integration with *SBCcollect* and *SBCserver* is in progress. Development of robust software for the combination of data from different crystals or even data from different spots of the same crystal is also in progress. We also note that having a separate end-station dedicated to crystal screening would also be highly beneficial to facilitate smooth integration of the cryo-temperatures data collection with the $T = 295$ K *in situ* option. The SBC will continue exploring these possibilities to incorporate scanner-based experiments into the user program. In the future, we also anticipate better integration of the beamline with the adjoining Advanced Protein Characterization Facility, which hosts resources for high-throughput protein production and crystallization.

Acknowledgements

The authors would like to thank members of the MCSG and SBC for their support. Specifically we would like to thank Michael Endres and Shonda Clancy for help with gene cloning, Dr Stephen Ginell for initial testing of X-ray absorption and scattering of commercial crystallization plates and Dr Krzysztof Lazarski for help with data collection. This research has been funded in part by a grant from the National Institutes of Health GM094585 (AJ), and by the US Department of Energy, Office of Biological and Environmental Research, under Contract DE-AC02-06CH11357.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Aller, P. *et al.* (2015). *Methods Mol. Biol.* **1261**, 233–253.

Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
Axford, D. *et al.* (2012). *Acta Cryst.* D**68**, 592–600.
Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X., Ebner, S. G., Isenegger, A., Schneider, R., Schneider, J., Glettig, W., Pradervand, C., Panepucci, E. H., Tomizaki, T., Wang, M. & Schulze-Briese, C. (2011). *Cryst. Growth Des.* **11**, 916–923.
Bulek, A. M., Madura, F., Fuller, A., Holland, C. J., Schauenburg, A. J., Sewell, A. K., Rizkallah, P. J. & Cole, D. K. (2012). *J. Immunol. Methods*, **382**, 203–210.
Campobasso, N., Begun, J., Costello, C. A., Begley, T. P. & Ealick, S. E. (1998). *Acta Cryst.* D**54**, 448–450.
Carter, C. W. Jr & Carter, C. W. (1979). *J. Biol. Chem.* **254**, 12219–12223.
Coquelle, N., Brewster, A. S., Kapp, U., Shilova, A., Weinhausen, B., Burghammer, M. & Colletier, J.-P. (2015). *Acta Cryst.* D**71**, 1184–1196.
Cowtan, K. (1994). *Jnt CCP4 ESF-EACBM Newslett. Protein Crystallogr.* **31**, 34–38.
Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.
Delucas, L. J., Hamrick, D., Cosenza, L., Nagy, L., McCombs, D., Bray, T., Chait, A., Stoops, B., Belgovskiy, A., Wilson, W. W., Parham, M. & Chernov, N. (2005). *Prog. Biophys. Mol. Biol.* **88**, 285–309.
Donnelly, M. I., Zhou, M., Millard, C. S., Clancy, S., Stols, L., Eschenfeldt, W. H., Collart, F. R. & Joachimiak, A. (2006). *Protein Expr. Purif.* **47**, 446–454.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Eschenfeldt, W. H., Stols, L., Millard, C. S., Joachimiak, A. & Mark, I. D. (2009). *Methods Mol. Biol.* **498**, 105–115.
Fraser, J. S., van den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N. & Alber, T. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 16247–16252.
Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.
Ginn, H. M., Messerschmidt, M., Ji, X., Zhang, H., Axford, D., Gildea, R. J., Winter, G., Brewster, A. S., Hattne, J., Wagner, A., Grimes, J. M., Evans, G., Sauter, N. K., Sutton, G. & Stuart, D. I. (2015). *Nat. Commun.* **6**, 6435.
Grimm, C., Chari, A., Reuter, K. & Fischer, U. (2010). *Acta Cryst.* D**66**, 685–697.
Helliwell, J. R. (2005). *Acta Cryst.* D**61**, 793–798.
Heymann, M., Opthalage, A., Wierman, J. L., Akella, S., Szebenyi, D. M. E., Gruner, S. M. & Fraden, S. (2014). *IUCrJ*, **1**, 349–360.
Hoggan, D. B., Chao, J. A., Prasad, G. S., Stout, C. D. & Williamson, J. R. (2003). *Acta Cryst.* D**59**, 466–473.
Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Acta Cryst.* D**60**, 1670–1673.
Juers, D. H., Lovelace, J., Bellamy, H. D., Snell, E. H., Matthews, B. W. & Borgstahl, G. E. O. (2007). *Acta Cryst.* D**63**, 1139–1153.
Juers, D. H. & Matthews, B. W. (2001). *J. Mol. Biol.* **311**, 851–862.
Juers, D. H. & Matthews, B. W. (2004). *Q. Rev. Biophys.* **37**, 105–119.
Keedy, D. A., van den Bedem, H., Sivak, D. A., Petsko, G. A., Ringe, D., Wilson, M. A. & Fraser, J. S. (2014). *Structure*, **22**, 899–910.
Kim, C. U., Wierman, J. L., Gillilan, R., Lima, E. & Gruner, S. M. (2013). *J. Appl. Cryst.* **46**, 234–241.
Kim, Y., Babnigg, G., Jedrzejczak, R., Eschenfeldt, W. H., Li, H., Maltseva, N., Hatzos-Skintges, C., Gu, M., Makowska-Grzyska, M., Wu, R., An, H., Chhor, G. & Joachimiak, A. (2011). *Methods*, **55**, 12–28.
Kisselman, G., Qiu, W., Romanov, V., Thompson, C. M., Lam, R., Battaile, K. P., Pai, E. F. & Chirgadze, N. Y. (2011). *Acta Cryst.* D**67**, 533–539.
Kuller, A., Fleri, W., Bluhm, W. F., Smith, J. L., Westbrook, J. & Bourne, P. E. (2002). *Trends Biochem. Sci.* **27**, 213–215.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nat. Protoc.* **3**, 1171–1179.
Lasala, R., Coudray, N., Abdine, A., Zhang, Z., Lopez-Redondo, M., Kirshenbaum, R., Alexopoulos, J., Zolnai, Z., Stokes, D. L. & Ubarretxena-Belandia, I. (2015). *J. Struct. Biol.* **189**, 123–134.
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.
Litzinger, S., Fischer, S., Polzer, P., Diederichs, K., Welte, W. & Mayer, C. (2010). *J. Biol. Chem.* **285**, 35675–35684.
Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. (2014). *Nucleic Acids Res.* **42**, D490–495.
Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). *Nature (London)*, **389**, 251–260.
Maeki, M., Yoshizuka, S., Yamaguchi, H., Kawamoto, M., Yamashita, K., Nakamura, H., Miyazaki, M. & Maeda, H. (2012). *Anal. Sci.* **28**, 65.
Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* D**62**, 859–866.
Otwinowski, Z. (1991). *Proceedings of the Daresbury Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Perry, S. L., Guha, S., Pawate, A. S., Bhaskarla, A., Agarwal, V., Nair, S. K. & Kenis, P. J. (2013). *Lab Chip*, **13**, 3183–3187.
Petrova, T., Ginell, S., Mitschler, A., Kim, Y., Lunin, V. Y., Joachimiak, G., Cousido-Siah, A., Hazemann, I., Podjarny, A., Lazarski, K. & Joachimiak, A. (2010). *Acta Cryst.* D**66**, 1075–1091.
Petrova, T., Lunin, V. Y., Ginell, S., Hazemann, I., Lazarski, K., Mitschler, A., Podjarny, A. & Joachimiak, A. (2009). *J. Mol. Biol.* **387**, 1092–1105.
Rosenbaum, G. *et al.* (2006). *J. Synchrotron Rad.* **13**, 30–45.
Sawaya, M. R. *et al.* (2014). *Proc. Natl Acad. Sci. USA*, **111**, 12769–12774.
Skrzypczak-Jankun, E., Bianchet, M. A., Amzel, L. M. & Funk, M. O. Jr (1996). *Acta Cryst.* D**52**, 959–965.
Spence, J. C. H. (2015). *IUCrJ*, **2**, 305–306.
Wikoff, W. R., Duda, R. L., Hendrix, R. W. & Johnson, J. E. (1998). *Virology*, **243**, 113–118.
Yadav, M. K., Gerdts, C. J., Sanishvili, R., Smith, W. W., Roach, L. S., Ismagilov, R. F., Kuhn, P. & Stevens, R. C. (2005). *J. Appl. Cryst.* **38**, 900–905.
Yin, X., Scalia, A., Leroy, L., Cuttitta, C. M., Polizzo, G. M., Ericson, D. L., Roessler, C. G., Campos, O., Ma, M. Y., Agarwal, R., Jackimowicz, R., Allaire, M., Orville, A. M., Sweet, R. M. & Soares, A. S. (2014). *Acta Cryst.* D**70**, 1177–1189.
Zhang, H. *et al.* (2015). *Cell*, **161**, 833–844.