



Structure-factor amplitude reconstruction from serial femtosecond crystallography of two-dimensional membrane-protein crystals

Cecilia M. Casadei,^a Karol Nass,^a Anton Barty,^b Mark S. Hunter,^c Celestino Padeste,^a Ching-Ju Tsai,^a Sébastien Boutet,^d Marc Messerschmidt,^{d,e} Leonardo Sala,^a Garth J. Williams,^{d,f} Dmitry Ozerov,^a Matthew Coleman,^c Xiao-Dan Li,^a Matthias Frank^c and Bill Pedrini^{a,*}

Received 6 August 2018

Accepted 16 October 2018

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

Keywords: free-electron lasers; serial femtosecond crystallography; membrane proteins; two-dimensional crystals.

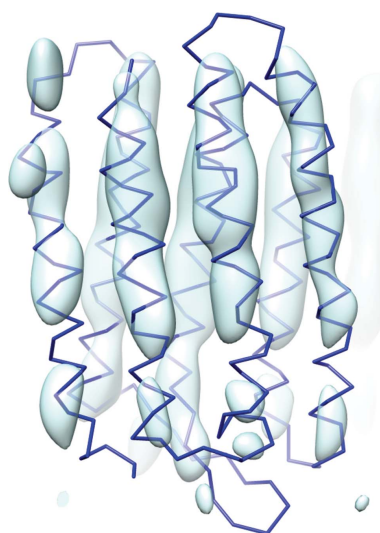
Supporting information: this article has supporting information at www.iucrj.org

^aPaul Scherrer Institute, 5232 Villigen PSI, Switzerland, ^bCenter for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany, ^cLawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA, ^dLinac Coherent Light Source, 2575 Sand Hill Road, Menlo Park, CA 94025, USA, ^eNational Science Foundation BioXFEL Science and Technology Center, 700 Ellicott Street, Buffalo, NY 14203, USA, and ^fNSLS-II, Brookhaven National Laboratory, PO Box 5000, Upton, NY 11973, USA. *Correspondence e-mail: bill.pedrini@psi.ch

Serial femtosecond crystallography of two-dimensional membrane-protein crystals at X-ray free-electron lasers has the potential to address the dynamics of functionally relevant large-scale motions, which can be sterically hindered in three-dimensional crystals and suppressed in cryocooled samples. In previous work, diffraction data limited to a two-dimensional reciprocal-space slice were evaluated and it was demonstrated that the low intensity of the diffraction signal can be overcome by collecting highly redundant data, thus enhancing the achievable resolution. Here, the application of a newly developed method to analyze diffraction data covering three reciprocal-space dimensions, extracting the reciprocal-space map of the structure-factor amplitudes, is presented. Despite the low resolution and completeness of the data set, it is shown by molecular replacement that the reconstructed amplitudes carry meaningful structural information. Therefore, it appears that these intrinsic limitations in resolution and completeness from two-dimensional crystal diffraction may be overcome by collecting highly redundant data along the three reciprocal-space axes, thus allowing the measurement of large-scale dynamics in pump–probe experiments.

1. Introduction

The ultrashort and ultrabright pulses provided by hard X-ray free-electron lasers (XFELs) have enabled innovative experimental investigation methods to address new scientific problems. In the field of macromolecular crystallography, which traditionally uses three-dimensional crystals as samples, the femtosecond pulse duration on one hand allows most radiation damage to be outrun (Lomb *et al.*, 2011; Barty *et al.*, 2011; Nass *et al.*, 2015), making measurements at room temperature and/or with smaller and smaller crystals (to the submicrometre range) possible (Gati *et al.*, 2017). On the other hand, structural dynamics at 1–3 Å resolution that can be triggered externally, such as side-chain movements or cofactor isomerizations, become accessible on the femtosecond time scale by pump–probe experiments (Kern *et al.*, 2014; Kupitz *et al.*, 2014; Tenboer *et al.*, 2014; Barends *et al.*, 2015; Nango *et al.*, 2016; Nogly *et al.*, 2016; Young *et al.*, 2016; Suga *et al.*, 2017) and on the millisecond timescale by mix-and-inject experiments (Stagno *et al.*, 2016; Olmos *et al.*, 2018). Because each crystal is destroyed by the interaction with the X-ray pulse, the



single-shot diffraction data have to be recorded from a large number of crystals to achieve sufficient completeness and redundancy. This data-acquisition strategy is called serial femtosecond crystallography (SFX). The crystals are delivered to the beam within a running liquid or viscous jet (Chapman *et al.*, 2011; Boutet *et al.*, 2012; Weierstall *et al.*, 2014) or by a solid support that is scanned through the beam (Hunter *et al.*, 2014; Cohen *et al.*, 2014; Roedig *et al.*, 2017). From the data-analysis point of view, the main challenge to be solved was the merging of diffraction patterns from crystals of different sizes and in random orientations illuminated by X-ray pulses of variable intensity and wavelength (White *et al.*, 2012; Sauter *et al.*, 2013; Neutze *et al.*, 2015; Schlichting, 2015).

Conformational changes at larger length scales of 3–6 Å are more challenging for investigation with three-dimensional crystals because the motions may be sterically hindered (Kühlbrandt, 2000). If available, two-dimensional crystals represent an opportunity because of the looser intermolecular contacts owing to the single layer of molecules. Membrane proteins are definitely the most relevant candidates (Stahlberg *et al.*, 2001) because their function typically involves such structural modifications (see, for example, Subramaniam & Henderson, 2000) and because the arrangement in two dimensions more closely mimics the environment on the cell membrane. When the relevant structural modifications take place on submillisecond timescales, investigations by electron microscopy and electron diffraction on samples whose dynamics have been frozen by flash-cooling are difficult (Subramaniam & Henderson, 1999). This opens a niche for SFX on two-dimensional crystals in pump–probe mode.

The diffraction signal of two-dimensional crystals is markedly lower than for analogous three-dimensional crystals with the same dimensions because the diffracting volume is orders of magnitude smaller and the reciprocal-space region that generates the diffraction is not concentrated in Bragg points but is diluted over one-dimensional Bragg rods. Because of the weak signal and the ill-effects of radiation damage, measuring high-resolution diffraction data from two-dimensional protein crystals at a continuous X-ray source is extremely challenging. Data collection at an XFEL represents a viable alternative (Frank *et al.*, 2014; Pedrini *et al.*, 2014). In recent work, we showed that the signal-to-noise ratio of the diffracted intensities is substantially enhanced by summing equivalent portions of images across the data set (Casadei *et al.*, 2018). In this way, the resolution of a highly redundant data set collected in November 2013 from two-dimensional bacteriorhodopsin crystals at zero tilt angle, *i.e.* with the incoming X-ray beam perpendicular to the crystal plane, could be extended from about 6 Å to the detector edge at 4 Å.

With zero-tilt data only a reciprocal-space slice is sampled, corresponding to two points on each Bragg rod. During the same November 2013 beamtime, we also collected data at a few different nonzero tilt angles, which cover three dimensions in reciprocal space and led to a genuine three-dimensional data set. We report here on the application of a novel method to merge the diffraction images and determine the structure-factor amplitudes along the Bragg rods. These were then

phased by molecular replacement. The composite OMIT maps (Terwilliger *et al.*, 2008) indicate that, despite their low completeness and the limited resolution of about 6 Å, the experimental data contain meaningful structural information. The measures required to improve the data quality, which are crucial to follow structural dynamics in future pump–probe experiments, are then discussed.

2. Results

2.1. Bragg rod intensity reconstruction

Sets of 1000 diffraction images of two-dimensional bacteriorhodopsin D96N mutant crystals were collected at the CXI experimental station of the LCLS free-electron laser at three different tilt angles $\eta = 5, 15$ and 20° . The diffraction images were analyzed assuming $p3$ symmetry (planar space group 13) of the crystal (Henderson *et al.*, 1990), with two unit-cell vectors of equal length a forming an angle γ of 120° . The corresponding two-dimensional reciprocal-space unit cell is spanned by two vectors of length $2\pi/a$ forming an angle of 60° . With the further assumption that Friedel symmetry of the diffraction intensity is valid, it follows that the point group of the structure-factor amplitudes, and thus of the diffracted intensity, is S_6 .

The data-analysis pipeline to calculate the structure-factor amplitudes consists of eight subsequent steps, which are schematized in Fig. 1. Because it presents a number of novel aspects with respect to that applied previously to untilted data (Casadei *et al.*, 2018), the pipeline is outlined below in some detail and for each step the obtained outcome is mentioned explicitly.

Step 1. The software *Cheetah* (Barty *et al.*, 2014) was first employed to apply dark-current and gain corrections to the raw diffraction data. Two examples are shown in Figs. 2(a) and 2(b), which clearly show patterns from tilted, $p3$ -symmetric two-dimensional crystals. The peaks are visible by eye down to a resolution of 7 Å. *Cheetah* was then used to generate a list of coordinates of the positions of the peaks, identified as clusters of pixels with high intensity.

Step 2. The peaks were grouped according to their mutual compatibility with a Bragg peak pattern from a single $p3$ -symmetric two-dimensional crystal with lattice constant fixed at $a = 62.45$ Å (Henderson *et al.*, 1990) positioned on a plane at a given tilt angle η . A Bragg peak originates from the intersection of the Ewald sphere with a Bragg rod labeled by two integers (h, k) (Figs. 2c and 2d), and its position on the detector $\mathbf{r}_{D,obs}$, given in terms of the angles (θ, φ_D) , can be calculated as a function of (h, k, φ, η) , where φ is the in-plane orientation angle parameter (see Section 4). Hence, imposing the best matching of the observed and the predicted peak positions on the detector gave a first estimate of φ . The grouping of the peaks and the optimization of φ were performed with an algorithm similar to that used for untilted data (Casadei *et al.*, 2018). If the peak subset associated with one two-dimensional crystal contained at least 18 *Cheetah* peaks, all potential Bragg peaks were searched directly in the

experimental diffraction images as connected regions of high-intensity pixels in proximity to the positions predicted

Table 1
Summarizing table.

The results of data processing including success rates of indexing-ambiguity solution and scale-factor determination as explained in the text.

η (data-set tilt angle)	5°	15°	20°
No. of collected images	968	1232	968
Percentage of images with lattices	58.2%	63.4%	58.2%
No. of lattices	992 (100%)	1383 (100%)	799 (100%)
No. of re-indexed lattices (first)	902 (90.9%)	1180 (85.3%)	507 (63.4%)
No. of re-indexed and rescaled lattices (first)	771 (77.7%)	1027 (74.3%)	419 (52.4%)
No. of re-indexed and rescaled lattices (second)	513 (51.7%)	1020 (73.8%)	651 (81.5%)

assuming the orientation φ . If at least 45% of the predicted peaks with in-plane resolution down to 6 Å were identified, the peak set was called ‘lattice’ and kept for the following evaluation steps. Up to four lattices were extracted from single images. The total number of lattices is reported in the third row of Table 1.

For each lattice, the unit-cell parameter a , the in-plane crystal orientation φ and the direct beam coordinates x_b , y_b were further refined simultaneously by minimizing the expression

$$f(a, \varphi, x_b, y_b) = \sum \frac{|\mathbf{r}_{D,obs} - \mathbf{r}_{D,calc}(a, \varphi, x_b, y_b)|}{r_{D,calc}}, \quad (1)$$

where the sum is over all identified peaks and $\mathbf{r}_{D,calc}$ is the predicted Bragg peak position on the detector. The minimum

was determined by using multiple iterations of either a systematic grid search in the four-dimensional parameter space with progressively finer grid spacing or the Powell algorithm (Powell, 1964). The average refined unit-cell vector length was 62.62 Å, with a standard deviation of 0.06 Å. The shift of the beam position $(\Delta x_b^2 + \Delta y_b^2)^{1/2}$ was typically below one pixel. We refrained from optimizing additional parameters such as the X-ray wavelength (energy) and the lattice tilt direction because of the large overhead in computational time.

For each predicted peak position on a Bragg rod (h, k) , the out-of-plane momentum transfer $q_{rod}(h, k, \varphi, \eta)$ was calculated (see Section 4), and the intensity $I(h, k, q_{rod})$ was determined by integration in a circular area of radius corresponding to five detector pixels, after having subtracted a background modeled by an affine function. To account for the Lorentz and polarization effects, each intensity was multiplied by the correction factor

$$\frac{\cos(\eta_L)}{1 - \sin^2(2\theta) \cos^2 \varphi_D}, \quad (2)$$

where η_L is the angle between the diffracted beam direction and the Bragg rod.

Step 3. Any transformation of the crystal inside the two-dimensional crystal plane which leaves the Bragg peak positions unal-

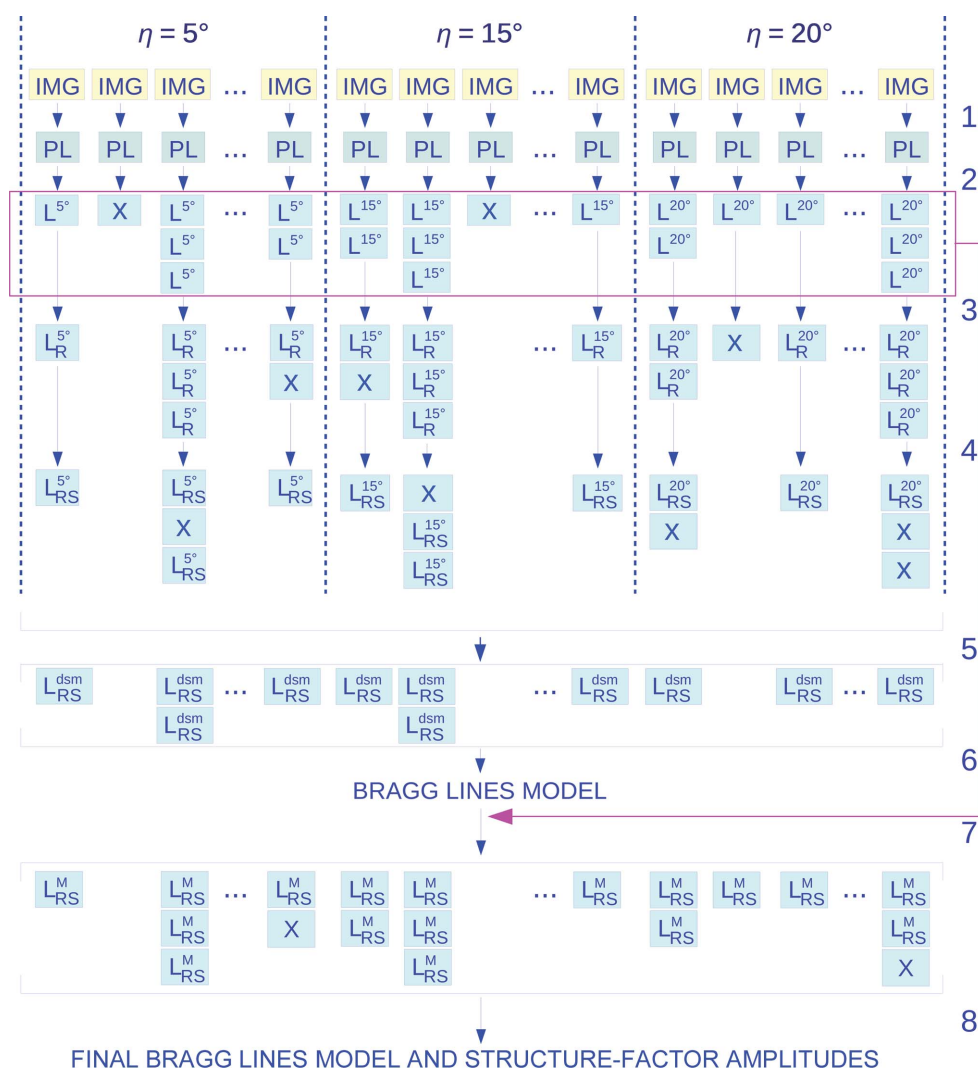


Figure 1

Data-analysis protocol. Abbreviations: IMG, diffraction image; PL, peak list; L, lattice; X, discarded; R, re-indexed; S, scaled; dsm, after data-set merging; M, model. Step 1, extraction of a list of high-intensity peaks from each diffraction image. Step 2, lattice identification and refinement, spot search on the images, refinement of parameters and spot integration. Step 3, indexing-ambiguity solution within each data set. Step 4, scaling of intensities within each data set. Step 5, indexing-ambiguity solution and scaling between different data sets. Step 6, data merging and modeling of intensities along Bragg lines. Step 7, indexing-ambiguity solution and scaling of lattices determined in Step 2 using the intensity model as a reference. Step 8, data merging and fitting to obtain the final intensity model and structure-factor amplitude extraction.

tered but is not a symmetry of the structure-factor amplitude leads to an ambiguity in indexing the lattice. Rotation of the crystal by 180° around the crystal-plane normal is always such a transformation, the ambiguity being in the assignment of a peak to either the reciprocal-space point h, k, q_{rod} or the point $h, k, -q_{\text{rod}}$. Furthermore, in the considered S_6 case, rotations of a crystal by multiples of 60° preserve the peak positions, but only rotations by multiples of 120° are symmetries of the structure-factor amplitude. Finally, the face-flip of the two-dimensional crystal also preserves the peak position but not the structure-factor amplitude, with the resulting ambiguity being between (h, k, q_{rod}) and $(k, h, -q_{\text{rod}})$. Because of the S_6 symmetry, each reflection could be mapped for simplicity to an equivalent reflection on a Bragg rod (h, k) with $h \geq 0$ and $k > 0$. With this simplification, the re-indexing of the lattice consists of transforming the indices assigned to each peak with one of the following operations, in order to achieve the best mutual correlation of intensities of equivalent peaks from pairs of lattices:

$$\begin{aligned} T_0 &: (h, k, q_{\text{rod}}) \rightarrow (h, k, q_{\text{rod}}), \\ T_1 &: (h, k, q_{\text{rod}}) \rightarrow (-h, -k, q_{\text{rod}}) \cong (h, k, -q_{\text{rod}}), \\ T_2 &: (h, k, q_{\text{rod}}) \rightarrow (k, h, -q_{\text{rod}}), \\ T_3 &: (h, k, q_{\text{rod}}) \rightarrow (-k, -h, -q_{\text{rod}}) \cong (k, h, q_{\text{rod}}). \end{aligned}$$

For the last two transformations the equivalence $(0, k, q_{\text{rod}}) \equiv (k, 0, -q_{\text{rod}})$ is used to maintain the condition $k > 0$. The determination of the per-lattice transformations was performed similar to the case of untilted data (Casadei *et al.*, 2018; see Section 4), in which peaks with an in-plane resolution of $7 \text{ \AA}$ were considered. The success of the re-indexing was between 60% and 90% depending on the data set, as reported in the fourth row of Table 1.

Step 4. To account for the fluctuations of the crystal area illuminated by X-rays and of the intensity of the X-ray pulse itself, a lattice-dependent multiplicative factor was calculated to scale the intensities from diffraction patterns recorded at the same tilt angle. The rescaling factors were determined with

a procedure generalizing that applied previously for untilted data (Casadei *et al.*, 2018; see Section 4), comparing intensities of equivalent peaks from different lattices. Two peaks from different lattices on a same rod $\{(h, k)\}$ were considered to be comparable if the difference in q_{rod} was smaller than a threshold value of $0.003 \text{ \AA}^{-1}$. Bragg rods down to a three-dimensional resolution of $7 \text{ \AA}$ were considered. As shown in the fifth row of Table 1, the rescaling factor could be determined for 50–75% of the lattices, depending on the data set.

Step 5. The three data sets collected at different tilt angles were merged into a single set by data-set-specific re-indexing, which was performed similarly to the procedure in Step 3 to maximize intensity correlations between the sets, and overall rescaling by data-set-specific multiplicative factors, which was performed in analogy to Step 4 to obtain the best correspondence of the intensities.

Step 6. The intensities of each Bragg rod (h, k) were modeled using the Levenberg–Marquardt algorithm for least-squares minimization with the function

$$I_{(h,k)}^M(q_{\text{rod}}) = \sum_{i=-n}^{+n} A_i \frac{\sin[d(q_{\text{rod}} - i\delta q_{\text{rod}})]}{d(q_{\text{rod}} - i\delta q_{\text{rod}})}, \quad (3)$$

with fit parameters A_i , which was proposed to handle analog electron-diffraction data from two-dimensional

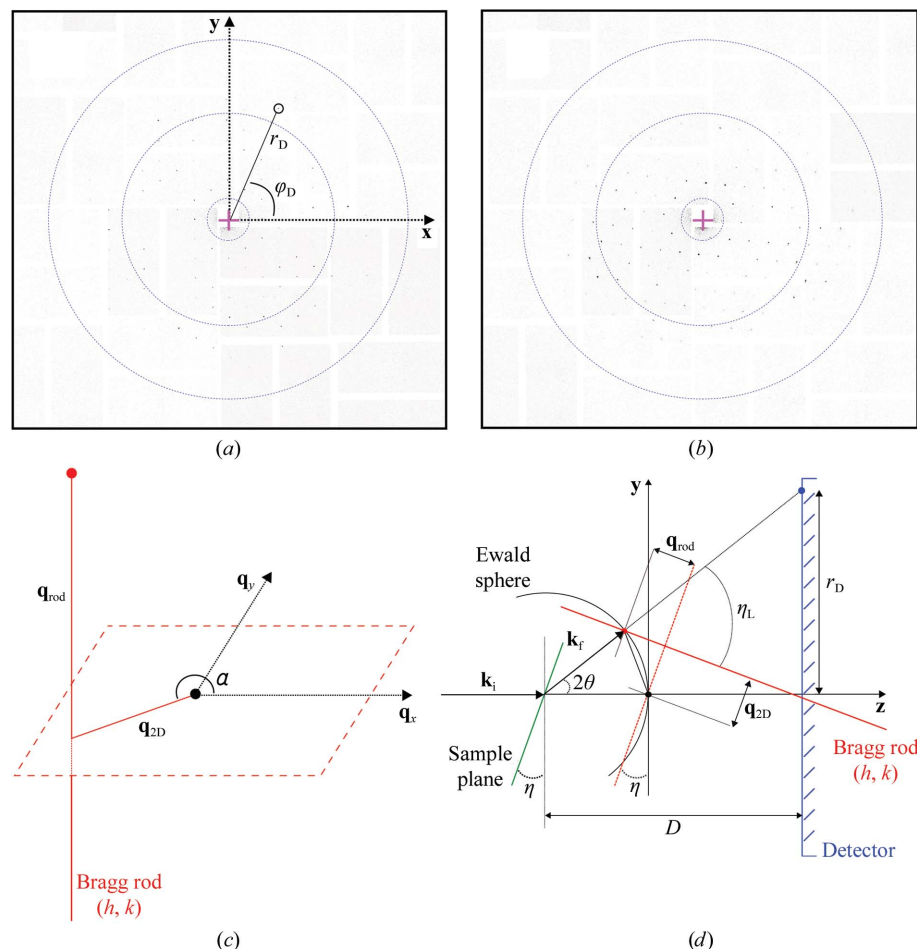


Figure 2
Diffraction from two-dimensional crystals. (a, b) Example diffraction images at 5° and 20° tilt angle, respectively. The circles are drawn at 50, 10 and $6 \text{ \AA}$ resolution. In (a), the detector-plane position in polar coordinates (r_D, φ_D) of a peak is superimposed. (c) Three-dimensional reciprocal space with a Bragg rod $h\mathbf{a}^* + k\mathbf{b}^* + q_{\text{rod}}\mathbf{q}_{\text{rod}}$ represented as a thick red line. (d) Ewald sphere construction for the diffraction process. The incoming and diffracted wavevectors are labeled $\mathbf{k}_i$ and $\mathbf{k}_f$, respectively. The diffracted beam is generated by the red dot on the Bragg rod. The diffraction angle of the Bragg peak on the detector (blue dot) has scattering angle 2θ . The sample plane with tilt angle η is represented in green.

crystals (Crowther *et al.*, 1970; Leifer & Henderson, 1983; Baldwin & Henderson, 1984). It consists of a sum of sinc functions centered at equidistant base points with spacing $\delta q_{\text{rod}} = 2\pi/2d$, defined by Shannon's sampling theorem (Shannon, 1949) for a two-dimensional crystal of thickness d in the out-of-plane direction (see Appendix A).

Step 7. All lattices, including those rejected previously, were again re-indexed and rescaled but now using the model $I_{\{(h,k)\}}^{\text{M}}(q_{\text{rod}})$ obtained in Step 6 as a reference. The re-indexing transformation of each lattice was set as that giving the best correlation for the linear regression between the lattice intensities and the corresponding intensities of the model. Only those lattices for which the correlation coefficient exceeded 0.92 were retained, and the rescaling factor was then set as the regression coefficient. With the mentioned threshold, about 50–80% of the initial lattices were retained (depending on the data set), as shown in row 6 of Table 1. The rationale behind this additional iteration was to cross-validate the model and retain as many lattices as possible for the next steps. At this point the data set consisted of Bragg rods $\{(h, k)\}$, each comprising about 9000 intensity observations labeled $I_{\{(h,k)\}}^{\text{M}}(q_{\text{rod}})$ and measured at different out-of-plane momenta q_{rod} . Each intensity measurement is shown as a yellow dot in the example Bragg rods in Fig. 3 and Supplementary Fig. S1.

Step 8. For each Bragg rod, the intensities $I_{\{(h,k)\}}$ were fitted again with the model function of (3) to obtain an updated intensity model $I_{\{(h,k)\}}^{\text{M}'}(q_{\text{rod}})$, represented by the magenta line in the examples in Fig. 3 and Supplementary Fig. S1. The new model is almost identical to the first model, which is shown in blue in the figure.

In view of molecular replacement, the intensities $I_{\{(h,k)\}}(l)$ were extracted at discrete, equally spaced points $q_{\text{rod}} = l\delta q_{\text{rod}}$, with the spacing $\delta q_{\text{rod}} = (1/2)(2\pi/2d)$ corresponding to over-sampling by a factor of two with respect to the minimal Shannon sampling interval. The error associated with each intensity value was calculated by propagating the errors on the fit parameters A_i in (3). Finally, the intensities and the errors were corrected using the procedure devised by French & Wilson (1978), which is routinely used in three-dimensional protein crystallography to handle meaningless negative intensities derived from the experimental data. The corrected intensities $I_{\{(h,k)\}}^{\text{FW}}(l)$ are shown as black points in Fig. 3 with the corrected error bars. As expected, only the lower intensities are subject to a relevant correction. The French and Wilson procedure also provides corrected values of the structure-factor amplitudes $F_{\{(h,k)\}}^{\text{O}}(l)$ and their error estimates, the calculated values of which are reported in the Supporting Information. Table 2 summarizes the data-set statistics as typically reported in traditional three-dimensional crystallography. The quality indicators in the last five columns were calculated as described in Section 4, with some modifications with respect to the three-dimensional case to account for the different structure of the two-dimensional SFX data. Their values indicate that data to the detector edge, corresponding to 6 Å in-plane resolution, can be retained for further analysis. Obviously, the data suffer from the same missing cone problem as in electron diffraction (Unwin & Henderson,

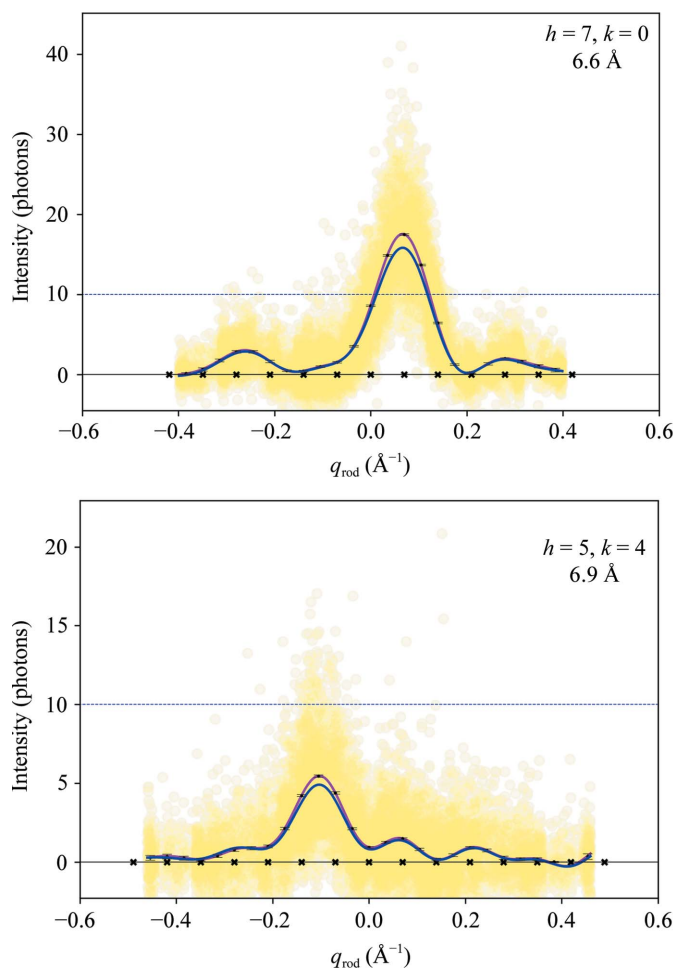


Figure 3
Bragg rod intensities. Intensity observations as a function of q_{rod} for Bragg rods $\{(7, 0)\}$ and $\{(5, 4)\}$. Yellow dots, experimental observations. Blue and magenta lines, preliminary and final intensity model from (3); the black crosses along the horizontal axis mark the base points of the sinc functions of the model. Black dots with error bars, Bayesian estimates of intensities at discrete q_{rod} sampling points. The Miller indices h, k and the in-plane resolution of each Bragg rod are indicated. The dashed blue line represents the intensity level corresponding to ten photons.

1975), which gives a very low completeness of 38.2% inside the resolution sphere at 5.3 Å.

2.2. Density maps from molecular replacement

Although the data set was of low completeness, the experimental structure-factor amplitudes $F_{\{(h,k)\}}^{\text{O}}(l)$ and the corresponding error estimates were used as input for molecular replacement and subsequent rigid-body refinement (see Section 4). The bacteriorhodopsin structure obtained by electron microscopy and diffraction, available as entry 1fbb in the Protein Data Bank (Berman *et al.*, 2003), was used as a starting model.

Owing to the bias introduced by the use of model phases, molecular-replacement maps are not to be considered representative of the information content of the data. To assess this

content, we calculated the composite OMIT map (see Section 4), which is shown in Fig. 4 for two different views of the molecule. The map shows that the data contain information about the position and orientation of the α -helices in the structure. Although the presence of a missing cone of data in the q_{rod} direction leads to real-space features which are elongated along the z axis, at sufficiently high contour levels only density overlapping with the expected positions of helices is present in the maps. These conclusions also emerge from the 'single-helices' OMIT maps (see Section 4) shown in Supplementary Fig. S3. Furthermore, the findings are reproduced by the molecular-replacement procedure from the intensities

sampled at $\delta q_{\text{rod}} = 2\pi/2d$ and $\delta q_{\text{rod}} = (1/4)2\pi/2d$ (see Supplementary Figs. S4 and S5).

3. Discussion

We present a protocol which allowed a three-dimensional X-ray diffraction data set from two-dimensional protein crystals to be analysed. The three-dimensionality of the data set in reciprocal space is a consequence of the tilting of the membrane supporting the samples with respect to the X-ray beam. These data differ from electron diffraction data in that the Ewald sphere cannot be considered to be flat, even at low

resolution, and in that each lattice diffraction pattern is a snapshot from one crystal, independent of any other pattern and not, for example, a representative in a tilt series from the same crystal. These differences triggered the development of a novel method which combines approaches from traditional X-ray crystallography (lattice identification, lattice-parameter refinement and Bayesian estimates of unique intensities, amplitudes and their error), three-dimensional SFX (merging of images from individual crystals affected by indexing ambiguity and intensity scaling) and two-dimensional electron diffraction (intensity modeling along Bragg rods). Applying this method, we reconstructed the diffraction intensities along Bragg rods in reciprocal space, from which the structure-factor amplitudes were extracted and their phases were determined by molecular replacement. The electron-density composite OMIT maps show that despite their low completeness and resolution, the data are meaningful.

The completeness is enhanced by recording data sets at higher tilt angles, which in general increases the q_{rod} coverage of the Bragg rods (see Supplementary Fig. S2). As a concrete example, with a tilt of 40° the completeness for the same crystal structure down to the same resolution range increases to 68.9%. Increasing the tilt angle unfortunately leads to an increased background because of the longer

Table 2

Data-processing statistics in three-dimensional resolution bins.

Columns 1 and 2, resolution range; column 3, number of observations; column 4, number of unique reflections; column 5, ratio of the two previous columns; columns 6–10, merging R value, half-data-set correlation coefficient ($CC_{1/2}$), CC^* , signal-to-noise ratio (S/N) and completeness as defined in Section 4.

Low ($\text{\AA}$)	High ($\text{\AA}$)	N_{obs}	N_{unique}	$N_{\text{obs}}/N_{\text{unique}}$	R	$CC_{1/2}$	CC^*	S/N	Completeness
54.09	25.20	32141	13	2472.38	0.18	0.9999	1.0000	138.69	0.459
25.20	15.61	32619	24	1359.12	0.38	0.9991	0.9998	71.93	0.310
15.61	12.71	31643	39	811.36	0.26	0.9995	0.9999	77.77	0.398
12.71	11.63	32228	34	947.88	0.27	0.9994	0.9998	93.58	0.513
11.63	10.18	32579	44	740.43	0.22	0.9997	0.9999	77.11	0.347
10.18	9.40	31688	54	586.81	0.29	0.9989	0.9997	58.76	0.463
9.40	8.64	31915	55	580.27	0.22	0.9996	0.9999	89.70	0.378
8.64	7.93	32254	86	375.05	0.47	0.9968	0.9992	27.38	0.451
7.93	7.52	32138	58	554.10	0.43	0.9985	0.9996	48.39	0.396
7.52	7.13	32036	76	421.53	0.56	0.9977	0.9994	27.80	0.421
7.13	6.79	33074	86	384.58	0.76	0.9921	0.9980	17.79	0.477
6.79	6.54	30574	75	407.65	0.81	0.9898	0.9974	20.09	0.448
6.54	6.20	33199	109	304.58	1.13	0.9778	0.9944	12.39	0.423
6.20	6.02	31883	86	370.73	1.16	0.9679	0.9918	14.37	0.494
6.02	5.30	31985	151	211.82	2.55	0.8233	0.9503	4.09	0.238
54.09	5.30	481956	990	486.82	0.30	0.9657	0.9906	35.11	0.382

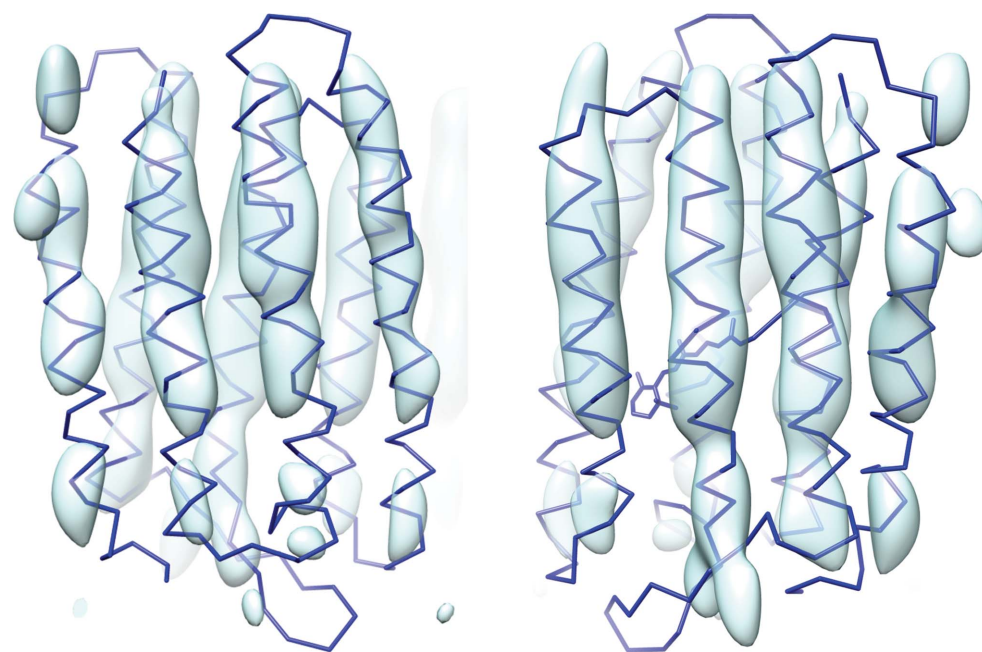


Figure 4

Composite OMIT map. Blue trace, backbone of atomic model 1fbb rigid-body refined using the experimental data with sampling $\delta q_{\text{rod}} = (1/2)(\pi/d)$. Cyan surface, the composite OMIT electron-density map with a contour level of 3.0σ . The structure is shown from two different views.

path of the X-ray beam inside the sample support. The image-summing approach presented in previous work (Casadei *et al.*, 2018), which aims to extend the achievable resolution by enhancing the signal-to-noise ratio, can be generalized straightforwardly. Obviously, to achieve the same improvement, for each q_{rod} bin of a tilted data set the same redundancy as an untilted data set has to be achieved, boosting the amount of required sample and the data-collection time by orders of magnitude. In this regard, the new high-speed scanning stage that has recently been commissioned at the CXI station opens new perspectives (Roedig *et al.*, 2017).

In agreement with the results from pioneering two-dimensional electron diffraction work (Unwin & Henderson, 1975), we observe that the intensities decay in an anisotropic fashion with increasing resolution. We quantify this effect by modeling the ratio $|F^{\text{o}}|^2/|F^{\text{m,iso}}|^2$, with model amplitudes $F^{\text{m,iso}}$ calculated in the same way as above from the 1fbb model but without anisotropic B factors (see Section 4), with a two-dimensional Gaussian function:

$$A \exp(-\delta B_{2D} q_{2D}^2 - \delta B_{\text{rod}} q_{\text{rod}}^2).$$

The fit is shown in Fig. 5, and the values of the obtained fit parameters $\delta B_{2D} = -0.27 \text{ \AA}^2$ and $\delta B_{\text{rod}} = 6.70 \text{ \AA}^2$ indicated that the experimental intensities decay remarkably faster in the q_{rod} direction than in the in-plane direction. We carried out the same treatment using observed and model structure factors from PDB entry 5b6v (Nango *et al.*, 2016), a structure of bR from three-dimensional SFX. In this case the fit parameters were refined to $\delta B_{2D} = 0.39 \text{ \AA}^2$ and $\delta B_{\text{rod}} = 0.51 \text{ \AA}^2$, showing that anisotropic effects are negligible with three-dimensional crystals. The large decay rate of two-dimensional crystal intensities along q_{rod} hints at increased disorder in the real-space out-of-plane direction as expected for a single-layer arrangement. Such an increase can be quantified by observing that the difference between experimental and model mean-square out-of-plane displacements amounts to approximately $6.7 \text{ \AA}^2$.

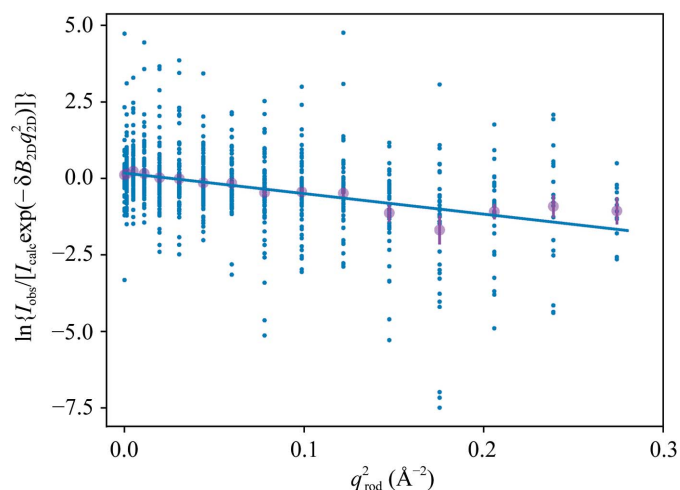


Figure 5
Anisotropic intensity decay. Blue dots, $\ln\{I_{\text{obs}}/[I_{\text{calc}}\exp(-\delta B_{2D}q_{2D}^2)]\}$, with $\delta B_{2D} = -0.27 \text{ \AA}^2$, as a function of q_{rod}^2 . Magenta dots, bin averages with error bars. Blue line, curve $\ln A - \delta B_{\text{rod}}q_{\text{rod}}^2$ with $\delta B_{\text{rod}} = 6.70 \text{ \AA}^2$.

In conclusion, we have shown that the structure-factor amplitudes derived from the two-dimensional SFX data contain meaningful and structural information, and have made the point that the completeness and resolution limitations are overcome by enhancing the redundancy in the data collection. It therefore appears that with the present status of XFELs, three-dimensional difference electron-density maps at a few ångströms resolution can be determined between protein molecules with different configurations in two-dimensional crystals. Of particular interest are large-scale configuration changes on this length scale that are sterically hindered in three-dimensional crystals. If these movements are triggered by optical stimuli, two-dimensional SFX data sets can be measured at different delays between the exciting laser pulse and the X-ray probing pulse, in a fashion that is nowadays standard in three-dimensional SFX (Standfuss & Spence, 2017).

4. Methods

4.1. Sample preparation

Purple membrane was isolated from *Halobacterium salinarum* expressing the gene for the D96N bacteriorhodopsin mutant (bR-D96N) and detergent-stabilized two-dimensional crystal suspensions were prepared using previously described procedures (Frank *et al.*, 2014; Pedrini *et al.*, 2014). The two-dimensional crystals were washed with 6 mM octylglucoside, suspended in 0.5% (w/v) glucose to a final protein concentration of 0.4 mg ml^{-1} and subsequently applied onto the sample carrier for X-ray diffraction data collection.

Silicon chips with areas of 25×25 and $12.5 \times 25 \text{ mm}^2$ with $200 \text{ }\mu\text{m}$ thickness, produced by Silson Inc., were used as sample carriers. The chips had a 44×44 or 22×44 array of $100 \times 100 \text{ }\mu\text{m}$ windows of 20 nm thick Si_3N_4 . A total of about $20 \text{ }\mu\text{l}$ bR-D96N two-dimensional crystal suspension was deposited onto the silicon chip and allowed to dry in air. The resulting glucose layer served to protect the protein sample from dehydration.

4.2. Experimental setup and data collection

The X-ray diffraction measurements were carried out using the $0.1 \text{ }\mu\text{m}$ focus setup of the CXI experimental station (Liang *et al.*, 2015) at the Linac Coherent Light Source. The beam size was estimated to be below 200 nm full width at half maximum (FWHM). The photon energy was set to 8.5 keV ($1.5 \text{ \AA}$), the pulse energy was approximately 2 mJ and the pulse length was approximately 35 fs FWHM.

The chips covered with two-dimensional bR-D96N crystals were mounted on a metallic frame that was fixed to the sample stages inside the vacuum experimental chamber. The sample stages were scanned in steps at a rate of about 1.5 s^{-1} . The silicon frames were kept in a nonperpendicular configuration with respect to the X-ray beam, with tilt angles of 5 , 15 and 20° about the x axis (Fig. 2). Diffraction patterns were recorded using a 2.3 megapixel Cornell–SLAC pixel-array detector,

which was positioned 285 mm downstream of the sample in the same vacuum chamber (Blaj *et al.*, 2015).

4.3. Software

Unless specified otherwise, the processing was performed using dedicated algorithms written in the Python 2.7 language, which are available on request.

4.4. Peak indexing

The geometry of the diffraction experiment using two-dimensional crystals is schematized in Figs. 2(c) and 2(d), where z denotes the direction of the incoming X-ray beam and η denotes the sample-support tilt angle about the x axis. The reciprocal-space plane spanned by the reciprocal basis vectors $\mathbf{a}^*$ and $\mathbf{b}^*$ can alternatively be described using the orthonormal vectors $\hat{\mathbf{q}}_x$ and $\hat{\mathbf{q}}_y$, with

$$\mathbf{a}^* = a_{q_x}^* \hat{\mathbf{q}}_x + a_{q_y}^* \hat{\mathbf{q}}_y, \quad (4)$$

$$\mathbf{b}^* = b_{q_x}^* \hat{\mathbf{q}}_x + b_{q_y}^* \hat{\mathbf{q}}_y. \quad (5)$$

The in-plane component of the momentum-transfer vector $\mathbf{q} = \mathbf{k}_f - \mathbf{k}_i$ is

$$\mathbf{q}_{2D} = h\mathbf{a}^* + k\mathbf{b}^* \quad (6)$$

and forms an azimuthal angle α with the $\hat{\mathbf{q}}_x$ axis given by

$$\begin{aligned} \alpha(\varphi, h, k) &= \arcsin\left(\frac{ha_{q_y}^* + kb_{q_y}^*}{q_{2D}}\right) \\ &= \varphi + \arcsin\left(\frac{ha_{q_{y,0}}^* + kb_{q_{y,0}}^*}{q_{2D}}\right), \end{aligned} \quad (7)$$

where φ is the random in-plane orientation of the two-dimensional crystal and $a_{q_{y,0}}^*$ and $b_{q_{y,0}}^*$ are the components of $\mathbf{a}^*$ and $\mathbf{b}^*$ along $\hat{\mathbf{q}}_y$ when the two-dimensional crystal is in the reference in-plane orientation. The transferred wavevector is

$$\mathbf{k}_f - \mathbf{k}_i = \begin{pmatrix} q_{2D} \cos \alpha \\ q_{2D} \cos \eta \sin \alpha + q_{\text{rod}} \sin \eta \\ q_{2D} \sin \eta \sin \alpha - q_{\text{rod}} \cos \eta \end{pmatrix} \quad (8)$$

with

$$\mathbf{k}_i = \begin{pmatrix} 0 \\ 0 \\ q \end{pmatrix}, \quad (9)$$

where $q = 2\pi/\lambda$. By squaring $\mathbf{k}_f$ and solving for q_{rod} one obtains

$$\begin{aligned} q_{\text{rod}}(\eta, h, k, \varphi) &= q \cos \eta \\ &\pm (q^2 \cos^2 \eta - q_{2D}^2 - 2qq_{2D} \sin \eta \sin \alpha)^{1/2}. \end{aligned} \quad (10)$$

By considering

$$q_{xy}^2 = k_{i,x}^2 + k_{i,y}^2$$

and replacing with the values from (8), the following expression for the azimuth detector coordinate of the diffraction spot is obtained,

$$\varphi_D = \arcsin\left(\frac{k_{f,y}}{q_{xy}}\right), \quad (11)$$

where

$$\begin{aligned} \frac{k_{f,y}}{q_{xy}} &= \frac{q_{2D} \cos \eta \sin \alpha + q_{\text{rod}} \sin \eta}{[q_{2D}^2 - (q_{2D}^2 \sin^2 \alpha - q_{\text{rod}}^2) \sin^2 \eta + 2q_{2D}q_{\text{rod}} \sin \alpha \sin \eta \cos \eta]^{1/2}}. \end{aligned} \quad (12)$$

The radial coordinate of the diffraction spot is

$$r_D = D \tan 2\theta, \quad (13)$$

where D is the detector distance and 2θ is the scattering angle given by

$$2\theta = 2 \arcsin\left[\frac{(q_{2D}^2 + q_{\text{rod}}^2)^{1/2}}{2q}\right]. \quad (14)$$

4.5. Re-indexing

The task is determining lattice-specific re-indexing transformations $\mathbf{T}_i$ (with $i = 0, 1, 2, 3$) that make the assignment of reciprocal-space indices coherent across the data set. The set of possible transformations is

$$\begin{aligned} \mathbf{T}_0 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & \mathbf{T}_1 &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ \mathbf{T}_2 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, & \mathbf{T}_3 &= \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

A lattice is randomly extracted from the data set and used as a reference R . The transformation $\mathbf{T}_{LR}$ required to re-index lattice L and make it compatible with R is established based on the calculation of intensity correlation coefficients. This determination of $\mathbf{T}_{LR}$ is accepted if the expression

$$\mathbf{T}_{LR} \cdot \mathbf{T}_{RL'} \cdot \mathbf{T}_{L'L} \quad (15)$$

equals the identity for at least 70% of a large number (~ 100) of randomly selected lattices L' . The procedure is repeated using different reference lattices and the consistency of the results is checked.

The transformation $\mathbf{T}_{L_1L_2}$ required to re-index L_1 and make it compatible with L_2 is determined by calculating a correlation coefficient between intensities from the two lattices in each of the four different indexing scenarios. The largest coefficient is considered to be representative of the correct transformation. The intensity correlation coefficient CC_i related to the transformation $\mathbf{T}_i$ is calculated by matching every spot (h, k, q_{rod}) in L_1 to any of the following $p3$ -symmetry equivalent spots in L_2 ,

$$\begin{bmatrix} \left(\begin{matrix} T_i^{11} & T_i^{12} & T_i^{13} \\ T_i^{21} & T_i^{22} & T_i^{23} \\ T_i^{31} & T_i^{32} & T_i^{33} \end{matrix} \right) \cdot \begin{pmatrix} h \\ k \\ \tilde{q}_{\text{rod}} \end{pmatrix} \end{bmatrix}^T \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\begin{bmatrix} \left(\begin{matrix} T_i^{11} & T_i^{12} & T_i^{13} \\ T_i^{21} & T_i^{22} & T_i^{23} \\ T_i^{31} & T_i^{32} & T_i^{33} \end{matrix} \right) \cdot \begin{pmatrix} h \\ k \\ \tilde{q}_{\text{rod}} \end{pmatrix} \end{bmatrix}^T \cdot \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\begin{bmatrix} \left(\begin{matrix} T_i^{11} & T_i^{12} & T_i^{13} \\ T_i^{21} & T_i^{22} & T_i^{23} \\ T_i^{31} & T_i^{32} & T_i^{33} \end{matrix} \right) \cdot \begin{pmatrix} h \\ k \\ \tilde{q}_{\text{rod}} \end{pmatrix} \end{bmatrix}^T \cdot \begin{pmatrix} 0 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where T indicates the transpose, and their Friedel mates. Importantly, in the general tilt case it is necessary to allow $\tilde{q}_{\text{rod}} \simeq q_{\text{rod}}$ and to set an upper limit on the absolute q_{rod} difference of matched spots.

4.6. Evaluation of data quality

A merging residual was calculated in three-dimensional resolution bins,

$$R = \frac{\sum_{\{(h,k)\}} \sum_i |I_{\{(h,k)\}}^{i'}(q_{\text{rod}}) - I_{\{(h,k)\}}^{M'}(q_{\text{rod}})|}{\sum_{\{(h,k)\}} \sum_i |I_{\{(h,k)\}}^{M'}(q_{\text{rod}})|}, \quad (16)$$

where the sum extends over all Bragg lines $\{(h, k)\}$ and observations i within the resolution bin (Baldwin & Henderson, 1984). A half-data-set correlation coefficient $CC_{1/2}$ was calculated per resolution bin as follows. Observations in each q_{rod} bin of width $\pi/(2d)$ were split randomly into two groups of (approximately) equal size and the linear correlation coefficient between half-data-set averages was calculated. The random splitting was repeated ten times and the average correlation coefficient was considered. The value of CC^* was calculated according to the definition in Karplus & Diederichs (2012). The global $CC_{1/2}$ and CC^* values were calculated as the weighted averages of individual bin values, with weights based on the number of unique reflections. Signal-to-noise ratios were calculated as three-dimensional resolution-bin averages of ratios between Bayesian estimates of intensities and their standard deviations for unique reflections. Completeness values in three-dimensional resolution bins were calculated as the ratio between the number of reciprocal-space points sampled by two-dimensional crystal diffraction with tilt angle 20° , considering $\delta q_{\text{rod}} = \pi/2d$, and the number of points within the corresponding spherical shell.

4.7. Molecular replacement

Bayesian estimates of unique reflection structure-factor amplitudes F^o and their errors were converted to MTZ format using the CCP4 program *F2MTZ* (Winn *et al.*, 2011). The data were phased by molecular replacement in *Phaser* (McCoy *et al.*, 2007) using the structural model 1fbb (Subramaniam & Henderson, 2000) from the Protein Data Bank (Berman *et al.*, 2003). The solution was rigid-body refined in *PHENIX* (Adams *et al.*, 2010) to obtain model structure-factor ampli-

tudes and phases (F^c , φ^c). To verify that the model was not biased by the phases from molecular replacement, we calculated composite OMIT electron-density maps (Terwilliger *et al.*, 2008) using the *PHENIX* software suite. We also carried out the standard procedure of removing a portion of the model, in this case a sequence of 20 amino acids, and using structure-factor amplitudes and phases ($F^{c,\text{OMIT}}$, $\varphi^{c,\text{OMIT}}$) determined from the truncation of the complete model. The Fourier coefficients of the OMIT maps ($mF^o - DF^{c,\text{OMIT}}$) $\exp(i\varphi^{c,\text{OMIT}})$, where m are the Sim weights and D are the Luzzati factors (Read, 1986), are conceived so that any feature accounted for in the data, but absent in the model, is represented by a region of positive electron density in the map.

4.8. Anisotropy modeling

To estimate anisotropic decay parameters, model structure factors $F^{\text{m,iso}}$ were calculated using *PHENIX* (Adams *et al.*, 2010) with disabled anisotropic scaling and including bulk-solvent corrections.

APPENDIX A

Derivation of the intensity model

We show that the summation in (3) represents an appropriate model for intensities in the q_{rod} direction because of the finite thickness d of the molecule monolayer. The diffracted intensity is the square amplitude of the Fourier transform of the charge density,

$$I_{\{(h,k)\}}(q_z) \propto \left| \int dx \int dy \int_{-d/2}^{+d/2} dz \rho(x, y, z) \exp[i2\pi(hx + ky + q_z z)] \right|^2. \quad (17)$$

This is equivalent to the expression

$$I_{\{(h,k)\}}(q_z) \propto \int dx \int dy \int_{-d}^{+d} dz P(x, y, z) \exp[i2\pi(hx + ky + q_z z)]. \quad (18)$$

P is the charge-density autocorrelation, q_z is the reciprocal-space coordinate along the direction of the Bragg lines, x and y are fractional coordinates in terms of the real-space crystallographic vectors **a** and **b**, and z is the real-space coordinate in the direction perpendicular to the plane spanned by **a** and **b**. It should be emphasized that since the monolayer has a finite thickness d , the extent of the autocorrelation function is limited to $|z| \leq d$. The autocorrelation function can be expressed by means of a discrete sum

$$P(x, y, z) \propto \sum_h \sum_k \sum_{\{q_{z,i}\}} I_{\{(h,k)\}}(q_{z,i}) \exp[-i2\pi(hx + ky + q_{z,i}z)], \quad (19)$$

where the sampling points $\{q_{z,i}\}$ along Bragg lines are equally spaced by $1/(2d)$ as defined by Shannon's sampling theorem. Including (19) in (18) gives

$$I_{\{(h,k)\}}(q_z) \propto \int dx \int dy \int_{-d}^{+d} dz \left\{ \sum_{h'} \sum_{k'} \sum_{\{q_{z,i}\}} I_{\{(h',k')\}}(q_{z,i}) \right. \\ \left. \times \exp[-i2\pi(h'x + k'y + q_{z,i}z)] \right\} \exp[i2\pi(hx + ky + q_z z)]. \quad (20)$$

Considering that

$$\int dx \exp[i2\pi(h - h')x] \propto \delta(h - h') \quad (21)$$

and

$$\int dy \exp[i2\pi(k - k')y] \propto \delta(k - k'), \quad (22)$$

the following expression for intensity is obtained,

$$I_{\{(h,k)\}}(q_z) \propto \int_{-d}^{+d} dz \sum_{h'} \sum_{k'} \sum_{\{q_{z,i}\}} I_{\{(h',k')\}}(q_{z,i}) \delta(h - h') \delta(k - k') \\ \times \exp[i2\pi(q_z - q_{z,i})z], \quad (23)$$

leading to

$$I_{\{(h,k)\}}(q_z) \propto \sum_{q_{z,i}} I_{\{(h,k)\}}(q_{z,i}) \int_{-d}^{+d} dz \exp[i2\pi(q_z - q_{z,i})z]. \quad (24)$$

The integral term in (24) is the Fourier transform of a step function extending from $-d$ to $+d$, namely a sinc function

$$I_{\{(h,k)\}}(q_z) \propto \sum_{q_{z,i}} I_{\{(h,k)\}}(q_{z,i}) \frac{\sin[2\pi d(q_z - q_{z,i})]}{2\pi d(q_z - q_{z,i})}. \quad (25)$$

(25) corresponds to (3), where the factor 2π is included in the definition of q_{rod} .

Acknowledgements

We thank Rafael Abela, Gebhard Schertler, John Spence, Geoffrey Feld and Richard Kirian for discussions and support of this work. CMC thanks Tim Gruene for fruitful discussion.

Funding information

This work was performed in part under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344 and supported by LLNL Lab-directed Research and Development (LDRD) project 12-ERD-031 and NIH grant 1R01GM117342-01. This work was supported in part by National Science Foundation grant 1231306 to M. Messerschmidt. Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. This research was supported in part by resources of the National Synchrotron Light Source II, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704.

References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* **D66**, 213–221.
- Baldwin, J. & Henderson, R. (1984). *Ultramicroscopy*, **14**, 319–335.
- Barends, T. R. M., Foucar, L., Ardevol, A., Nass, K., Aquila, A., Botha, S., Doak, R. B., Falahati, K., Hartmann, E., Hilpert, M., Heinz, M., Hoffmann, M. C., Köfinger, J., Koglin, J. E., Kovacsova, G., Liang, M., Milathianaki, D., Lemke, H., Reinstein, J., Roome, C. M., Shoeman, R. L., Williams, G. J., Burghardt, I., Hummer, G., Boutet, S. & Schlichting, I. (2015). *Science*, **350**, 445–450.
- Barty, A., Caleman, C., Aquila, A., Timneanu, N., Lomb, L., White, T. A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T. R. M., Barthelmess, M., Bogan, M. J., Bostedt, C., Bozek, J. D., Coffee, R., Coppola, N., Davidsson, J., DePonte, D. P., Doak, R. B., Ekeberg, T., Elser, V., Epp, S. W., Erk, B., Fleckenstein, H., Foucar, L., Fromme, P., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C. Y., Hartmann, R., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Hunter, M. S., Johansson, L., Kassemeyer, S., Kimmel, N., Kirian, R. A., Liang, M., Maia, F. R. N. C., Malmerberg, E., Marchesini, S., Martin, A. V., Nass, K., Neutze, R., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Scott, H., Schlichting, I., Schulz, J., Seibert, M. M., Shoeman, R. L., Sierra, R. G., Soltau, H., Spence, J. C. H., Stellato, F., Stern, S., Strüder, L., Ullrich, J., Wang, X., Weidenspointner, G., Weierstall, U., Wunderer, C. B. & Chapman, H. N. (2011). *Nat. Photonics*, **6**, 35–40.
- Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A. & Chapman, H. (2014). *J. Appl. Cryst.* **47**, 1118–1131.
- Berman, H., Henrick, K. & Nakamura, H. (2003). *Nat. Struct. Mol. Biol.* **10**, 980.
- Blaj, G., Caragiulo, P., Carini, G., Carron, S., Dragone, A., Freytag, D., Haller, G., Hart, P., Hasi, J., Herbst, R., Herrmann, S., Kenney, C., Markovic, B., Nishimura, K., Osier, S., Pines, J., Reese, B., Segal, J., Tomada, A. & Weaver, M. (2015). *J. Synchrotron Rad.* **22**, 577–583.
- Boutet, S., Lomb, L., Williams, G. J., Barends, T. R. M., Aquila, A., Doak, R. B., Weierstall, U., DePonte, D. P., Steinbrener, J., Shoeman, R. L., Messerschmidt, M., Barty, A., White, T. A., Kassemeyer, S., Kirian, R. A., Seibert, M. M., Montanez, P. A., Kenney, C., Herbst, R., Hart, P., Pines, J., Haller, G., Gruner, S. M., Philipp, H. T., Tate, M. W., Hromalik, M., Koerner, L. J., van Bakel, N., Morse, J., Ghonsalves, W., Arnlund, D., Bogan, M. J., Caleman, C., Fromme, R., Hampton, C. Y., Hunter, M. S., Johansson, L. C., Katona, G., Kupitz, C., Liang, M., Martin, A. V., Nass, K., Rededge, L., Stellato, F., Timneanu, N., Wang, D., Zatsepin, N. A., Schafer, D., Defever, J., Neutze, R., Fromme, P., Spence, J. C. H., Chapman, H. N. & Schlichting, I. (2012). *Science*, **337**, 362–364.
- Casadei, C. M., Tsai, C.-J., Barty, A., Hunter, M. S., Zatsepin, N. A., Padeste, C., Capitani, G., Benner, W. H., Boutet, S., Hau-Riege, S. P., Kupitz, C., Messerschmidt, M., Ogren, J. I., Pardini, T., Rothschild, K. J., Sala, L., Segelke, B., Williams, G. J., Evans, J. E., Li, X.-D., Coleman, M., Pedrini, B. & Frank, M. (2018). *IUCrJ*, **5**, 103–117.
- Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gork, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.-U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson,

- O., Svenda, M., Stern, S., Nass, K., Andrichke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R. M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhove, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature (London)*, **470**, 73–77.
- Cohen, A. E., Soltis, S. M., González, A., Aguila, L., Alonso-Mori, R., Barnes, C. O., Baxter, E. L., Brehmer, W., Brewster, A. S., Brunger, A. T., Calero, G., Chang, J. F., Chollet, M., Ehrensberger, P., Eriksson, T. L., Feng, Y., Hattne, J., Hedman, B., Hollenbeck, M., Holton, J. M., Keable, S., Kobilka, B. K., Kovaleva, E. G., Kruse, A. C., Lemke, H. T., Lin, G., Lyubimov, A. Y., Manglik, A., Mathews, I. L., McPhillips, S. E., Nelson, S., Peters, J. W., Sauter, N. K., Smith, C. A., Song, J., Stevenson, H. P., Tsai, Y., Uervirojnangkorn, M., Vinetsky, V., Wakatsuki, S., Weiss, W. I., Zadvornyy, O. A., Zeldin, O. B., Zhu, D. & Hodgson, K. O. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 17122–17127.
- Crowther, R. A., DeRosier, D. J. & Klug, A. (1970). *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* **317**, 319–340.
- Frank, M., Carlson, D. B., Hunter, M. S., Williams, G. J., Messerschmidt, M., Zatspein, N. A., Barty, A., Benner, W. H., Chu, K., Graf, A. T., Hau-Riege, S. P., Kirian, R. A., Padeste, C., Pardini, T., Pedrini, B., Segelke, B., Seibert, M. M., Spence, J. C. H., Tsai, C.-J., Lane, S. M., Li, X.-D., Schertler, G., Boutet, S., Coleman, M. & Evans, J. E. (2014). *IUCrJ*, **1**, 95–100.
- French, S. & Wilson, K. (1978). *Acta Cryst.* **A34**, 517–525.
- 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., Zatspein, 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.
- Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E. & Downing, K. H. (1990). *J. Mol. Biol.* **213**, 899–929.
- Hunter, M. S., Segelke, B., Messerschmidt, M., Williams, G. J., Zatspein, N. A., Barty, A., Benner, W. H., Carlson, D. B., Coleman, M., Graf, A., Hau-Riege, S. P., Pardini, T., Seibert, M. M., Evans, J., Boutet, S. & Frank, M. (2014). *Sci. Rep.* **4**, 6026.
- Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
- Kern, J., Tran, R., Alonso-Mori, R., Koroidov, S., Echols, N., Hattne, J., Ibrahim, M., Gul, S., Laksmono, H., Sierra, R. G., Gildea, R. J., Han, G., Hellmich, J., Lassalle-Kaiser, B., Chatterjee, R., Brewster, A. S., Stan, C. A., Glöckner, C., Lampe, A., DiFiore, D., Milathianaki, D., Fry, A. R., Seibert, M. M., Koglin, J. E., Gallo, E., Uhlig, J., Sokaras, D., Weng, T.-C., Zwart, P. H., Skinner, D. E., Bogan, M. J., Messerschmidt, M., Glatzel, P., Williams, G. J., Boutet, S., Adams, P. D., Zouni, A., Messinger, J., Sauter, N. K., Bergmann, U., Yano, J. & Yachandra, V. K. (2014). *Nat. Commun.* **5**, 4371.
- Kühlbrandt, W. (2000). *Nature (London)*, **406**, 569–570.
- Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatspein, N. A., Rendek, K. N., Hunter, M. S., Shoeman, R. L., White, T. A., Wang, D., James, D., Yang, J.-H., Cobb, D. E., Reeder, B., Sierra, R. G., Liu, H., Barty, A., Aquila, A. L., Deponte, D., Kirian, R. A., Bari, S., Bergkamp, J. J., Beyerlein, K. R., Bogan, M. J., Coleman, C., Chao, T.-C., Conrad, C. E., Davis, K. M., Fleckenstein, H., Galli, L., Hau-Riege, S. P., Kassemeyer, S., Laksmono, H., Liang, M., Lomb, L., Marchesini, S., Martin, A. V., Messerschmidt, M., Milathianaki, D., Nass, K., Ros, A., Roy-Chowdhury, S., Schmidt, K., Seibert, M., Steinbrener, J., Stellato, F., Yan, L., Yoon, C., Moore, T. A., Moore, A. L., Pushkar, Y., Williams, G. J., Boutet, S., Doak, R. B., Weierstall, U., Frank, M., Chapman, H. N., Spence, J. C. H. & Fromme, P. (2014). *Nature (London)*, **513**, 261–265.
- Leifer, D. & Henderson, R. (1983). *J. Mol. Biol.* **163**, 451–466.
- Liang, M., Williams, G. J., Messerschmidt, M., Seibert, M. M., Montanez, P. A., Hayes, M., Milathianaki, D., Aquila, A., Hunter, M. S., Koglin, J. E., Schafer, D. W., Guillet, S., Busse, A., Bergan, R., Olson, W., Fox, K., Stewart, N., Curtis, R., Miahnahri, A. A. & Boutet, S. (2015). *J. Synchrotron Rad.* **22**, 514–519.
- Lomb, L., Barends, T. R. M., Kassemeyer, S., Aquila, A., Epp, S. W., Erk, B., Foucar, L., Hartmann, R., Rudek, B., Rolles, D., Rudenko, A., Shoeman, R. L., Andreasson, J., Bajt, S., Barthelmess, M., Barty, A., Bogan, M. J., Bostedt, C., Bozek, J. D., Caleman, C., Coffee, R., Coppola, N., DePonte, D. P., Doak, R. B., Ekeberg, T., Fleckenstein, H., Fromme, P., Gebhardt, M., Graafsma, H., Gumprecht, L., Hampton, C. Y., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Holton, J. M., Hunter, M. S., Kabsch, W., Kimmel, N., Kirian, R. A., Liang, M., Maia, F. R. N. C., Meinhart, A., Marchesini, S., Martin, A. V., Nass, K., Reich, C., Schulz, J., Seibert, M. M., Sierra, R., Soltat, H., Spence, J. C. H., Steinbrener, J., Stellato, F., Stern, S., Timneanu, N., Wang, X., Weidenspointner, G., Weierstall, U., White, T. A., Wunderer, C., Chapman, H. N., Ullrich, J., Strüder, L. & Schlichting, I. (2011). *Phys. Rev. B*, **84**, 214111.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A.-N., Standfuss, J., Neutze, R. & Iwata, S. (2016). *Science*, **354**, 1552–1557.
- Nass, K., Foucar, L., Barends, T. R. M., Hartmann, E., Botha, S., Shoeman, R. L., Doak, R. B., Alonso-Mori, R., Aquila, A., Bajt, S., Barty, A., Bean, R., Beyerlein, K. R., Bubltz, M., Drachmann, N., Gregersen, J., Jönsson, H. O., Kabsch, W., Kassemeyer, S., Koglin, J. E., Krumrey, M., Mattle, D., Messerschmidt, M., Nissen, P., Reinhard, L., Sitsel, O., Sokaras, D., Williams, G. J., Hau-Riege, S., Timneanu, N., Caleman, C., Chapman, H. N., Boutet, S. & Schlichting, I. (2015). *J. Synchrotron Rad.* **22**, 225–238.
- Neutze, R., Brändén, G. & Schertler, G. F. (2015). *Curr. Opin. Struct. Biol.* **33**, 115–125.
- 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.
- Olmos, J. L., Pandey, S., Martin-Garcia, J. M., Calvey, G., Katz, A., Knoska, J., Kupitz, C., Hunter, M. S., Liang, M., Oberthuer, D., Yefanov, O., Wiedorn, M., Heyman, M., Holl, M., Pande, K., Barty, A., Miller, M. D., Stern, S., Roy-Chowdhury, S., Coe, J., Nagarathnam, N., Zook, J., Verburg, J., Norwood, T., Poudyal, I., Xu, D., Koglin, J., Seaberg, M. H., Zhao, Y., Bajt, S., Grant, T., Mariani, V., Nelson, G., Subramanian, G., Bae, E., Fromme, R., Fung, R., Schwander, P., Frank, M., White, T. A., Weierstall, U., Zatspein, N., Spence, J., Fromme, P., Chapman, H. N., Pollack, L., Tremblay, L., Ourmazd, A., Phillips, G. N. & Schmidt, M. (2018). *BMC Biol.* **16**, 59.
- Pedrini, B., Tsai, C.-J., Capitani, G., Padeste, C., Hunter, M. S., Zatspein, N. A., Barty, A., Benner, W. H., Boutet, S., Feld, G. K., Hau-Riege, S. P., Kirian, R. A., Kupitz, C., Messerschmidt, M., Ogren, J. I., Pardini, T., Segelke, B., Williams, G. J., Spence, J. C. H., Abela, R., Coleman, M., Evans, J. E., Schertler, G. F. X., Frank, M. & Li, X.-D. (2014). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130500.
- Powell, M. J. D. (1964). *Comput. J.* **7**, 155–162.
- Read, R. J. (1986). *Acta Cryst.* **A42**, 140–149.
- 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.
- Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. (2013). *Acta Cryst. D* **69**, 1274–1282.
- Schlichting, I. (2015). *IUCrJ*, **2**, 246–255.
- Shannon, C. E. (1949). *Proc. IRE*, **37**, 10–21.
- 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., Knoska, 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. (2016). *Nature (London)*, **541**, 242–246.
- Stahlberg, H., Fotiadis, D., Scheuring, S., Rémy, H., Braun, T., Mitsunaka, K., Fujiyoshi, Y. & Engel, A. (2001). *FEBS Lett.* **504**, 166–172.
- Standfuss, J. & Spence, J. (2017). *IUCrJ*, **4**, 100–101.
- Subramaniam, S. & Henderson, R. (1999). *J. Struct. Biol.* **128**, 19–25.
- Subramaniam, S. & Henderson, R. (2000). *Nature (London)*, **406**, 653–657.
- Suga, M., Akita, F., Sugahara, M., Kubo, M., Nakajima, Y., Nakane, T., Yamashita, K., Umena, Y., Nakabayashi, M., Yamane, T., Nakano, T., Suzuki, M., Masuda, T., Inoue, S., Kimura, T., Nomura, T., Yonekura, S., Yu, L.-J., Sakamoto, T., Motomura, T., Chen, J.-H., Kato, Y., Noguchi, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Nango, E., Tanaka, R., Naitow, H., Matsuura, Y., Yamashita, A., Yamamoto, M., Nureki, O., Yabashi, M., Ishikawa, T., Iwata, S. & Shen, J.-R. (2017). *Nature (London)*, **543**, 131–135.
- Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G. J., Koglin, J. E., Oberthuer, D., Heymann, M., Kupitz, C., Conrad, C., Coe, J., Roy-Chowdhury, S., Weierstall, U., James, D., Wang, D., Grant, T., Barty, A., Yefanov, O., Scales, J., Gati, C., Seuring, C., Srajer, V., Henning, R., Schwander, P., Fromme, R., Ourmazd, A., Moffat, K., Van Thor, J. J., Spence, J. C. H., Fromme, P., Chapman, H. N. & Schmidt, M. (2014). *Science*, **346**, 1242–1246.
- Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Adams, P. D., Read, R. J., Zwart, P. H. & Hung, L.-W. (2008). *Acta Cryst. D* **64**, 515–524.
- Unwin, P. N. T. & Henderson, R. (1975). *Nature (London)*, **257**, 28–32.
- Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C. H., Doak, R. B., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Han, G. W., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G. J., Koglin, J. E., Seibert, M. M., Klinker, M., Gati, C., Shoeman, R. L., Barty, A., Chapman, H. N., Kirian, R. A., Beyerlein, K. R., Stevens, R. C., Li, D., Shah, S. T. A., Howe, N., Caffrey, M. & Cherezov, V. (2014). *Nat. Commun.* **5**, 3309.
- White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A. & Chapman, H. N. (2012). *J. Appl. Cryst.* **45**, 335–341.
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst. D* **67**, 235–242.
- Young, I. D., Ibrahim, M., Chatterjee, R., Gul, S., Fuller, F. D., Koroidov, S., Brewster, A. S., Tran, R., Alonso-Mori, R., Kroll, T., Michels-Clark, T., Laksmono, H., Sierra, R. G., Stan, C. A., Hussein, R., Zhang, M., Douthett, L., Kubin, M., de Lichtenberg, C., Vo Pham, L., Nilsson, H., Cheah, M. H., Shevela, D., Saracini, C., Bean, M. A., Seuffert, I., Sokaras, D., Weng, T.-C., Pastor, E., Weninger, C., Fransson, T., Lassalle, L., Bräuer, P., Aller, P., Docker, P. T., Andi, B., Orville, A. M., Glowacki, J. M., Nelson, S., Sikorski, M., Zhu, D., Hunter, M. S., Lane, T. J., Aquila, A., Koglin, J. E., Robinson, J., Liang, M., Boutet, S., Lyubimov, A. Y., Uervirojnangkoorn, M., Moriarty, N. W., Liebschner, D., Afonine, P. V., Waterman, D. G., Evans, G., Wernet, P., Dobbek, H., Weis, W. I., Brunger, A. T., Zwart, P. H., Adams, P. D., Zouni, A., Messinger, J., Bergmann, U., Sauter, N. K., Kern, J., Yachandra, V. K. & Yano, J. (2016). *Nature (London)*, **540**, 453–457.