Acta Cryst D
STRUCTURAL BIOLOGY

ISSN 2059-7983

SAD phasing of XFEL data depends critically on the error model

Aaron S. Brewster,[a]* Asmit Bhowmick,[a] Robert Bolotovsky,[a] Derek Mendez,[a] Petrus H. Zwart[a,b] and Nicholas K. Sauter[a]*

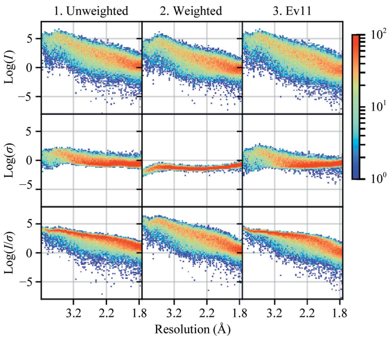
[a]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, and [b]Center for Advanced Mathematics for Energy Research Applications, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. *Correspondence e-mail: asbrewster@lbl.gov, nksauter@lbl.gov

Received 7 March 2019
Accepted 17 September 2019

Edited by R. J. Read, University of Cambridge, England

Keywords: XFELs; SAD phasing; error modeling; *cctbx.xfel*; serial crystallography.

A nonlinear least-squares method for refining a parametric expression describing the estimated errors of reflection intensities in serial crystallographic (SX) data is presented. This approach, which is similar to that used in the rotation method of crystallographic data collection at synchrotrons, propagates error estimates from photon-counting statistics to the merged data. Here, it is demonstrated that the application of this approach to SX data provides better SAD phasing ability, enabling the autobuilding of a protein structure that had previously failed to be built. Estimating the error in the merged reflection intensities requires the understanding and propagation of all of the sources of error arising from the measurements. One type of error, which is well understood, is the counting error introduced when the detector counts X-ray photons. Thus, if other types of random errors (such as readout noise) as well as uncertainties in systematic corrections (such as from X-ray attenuation) are completely understood, they can be propagated along with the counting error, as appropriate. In practice, most software packages propagate as much error as they know how to model and then include error-adjustment terms that scale the error estimates until they explain the variance among the measurements. If this is performed carefully, then during SAD phasing likelihood-based approaches can make optimal use of these error estimates, increasing the chance of a successful structure solution. In serial crystallography, SAD phasing has remained challenging, with the few examples of *de novo* protein structure solution each requiring many thousands of diffraction patterns. Here, the effects of different methods of treating the error estimates are estimated and it is shown that using a parametric approach that includes terms proportional to the known experimental uncertainty, the reflection intensity and the squared reflection intensity to improve the error estimates can allow SAD phasing even from weak zinc anomalous signal.



1. Introduction

Solving a novel protein structure using X-ray crystallography typically involves either a reliance on a similar structure from which molecular replacement (MR) can be used to derive phasing information, or the presence of heavy atoms that can provide anomalous differences for use in SAD (single-wavelength anomalous dispersion) or MAD (multiple-wavelength anomalous dispersion) phasing (among other methods). In SAD phasing, X-ray anomalous scattering by heavy atoms in the protein structure breaks inversion/Friedel symmetry in the diffraction pattern, with otherwise equivalent reflections typically exhibiting 3–4% differences in intensity. This information can be used to determine the heavy-atom substructure in the protein, which is then used to solve the phasing problem. This approach requires highly accurately measured intensities, and the analysis of such data has been shown to benefit from maximum-likelihood methods (de La Fortelle &

https://doi.org/10.1107/S2059798319012877

Bricogne, 1997; McCoy *et al.*, 2004), with the caveat that maximum-likelihood methods also require accurate estimates of the merged intensity errors.

In serial crystallography (SX), determining the reflection intensities with the required accuracy and estimating their error is challenging, which has made phasing new structures from SX data difficult. Typically, 10^2–10^7 crystals are exposed to either synchrotron or X-ray free-electron laser (XFEL) radiation. Each crystal is exposed once in a random orientation using liquid-stream injection, grid-based raster scanning or acoustic droplet injection (reviewed in Bergmann *et al.*, 2017). Individual diffraction patterns are indexed to determine the crystal orientation and unit-cell dimensions, and reflection locations are then predicted and integrated. Because the crystals are not rotated the reflections are only partially recorded, and therefore a post-refinement algorithm is used to apply a partiality correction factor in order to re-express the summed intensity in terms of the structure-factor equivalent. Finally, the redundantly measured reflections are merged together using either a simple average or a weighted average (White, 2014; Kabsch, 2014; Sauter, 2015; Uervirojnangkoorn *et al.*, 2015; Ginn *et al.*, 2015).

In crystallographic experiments, the error estimates from photon-counting statistics alone do not explain the variance observed in the measurements, always underestimating the variance owing to the presence of other sources of error. In 1985, an IUCr subcommittee on statistical descriptors was tasked to evaluate the validity of the statistical approaches used at the time to determine variances and provide recommendations (Schwarzenbach *et al.*, 1989). In their report, they suggested that if the multiplicity of the measurements was high enough then simply the spread of the measurements is sufficient to estimate the error. Otherwise, they recommended that crystallographic methods developers use error-propagation approaches to combine uncertainty in photon counting with random and systematic sources of error. Random error sources include readout noise and dark current. Systematic sources of error include X-ray attenuation from air, sample or water, detector misalignment and errors in estimating the wavelength or flux and, in the case of SX, the partiality. Because the reflections are only partially recorded, every measurement is reduced by anywhere from 0% to 100% of its full intensity, depending on the crystal orientation, mosaicity and the spectral characteristics of the beam. It is likely that partiality is the dominant source of error for SX data, where reflection tails touching the Ewald sphere introduce orders of magnitude more uncertainty than reflections that directly intersect the Ewald sphere.

The full list of sources of error is extensive and it is difficult to ensure that all sources of error have been accounted for. To this end, procedures have been developed to adjust error estimates, usually inflating them to larger values, using intensity-dependent and intensity-independent factors, after applying any other known corrections (Leslie, 1999, 2006; Otwinowski & Minor, 2001; Kabsch, 2010*a*,*b*; Evans, 2006, 2011). For a full set of references, see Rossmann & Arnold (2001).

In the present study, we have found that how the error estimates are obtained directly affects our ability to use SAD phasing to solve an XFEL structure *de novo*. We examined three methods for treating error and show that only some of them allowed us to find the Zn sites of a thermolysin data set using SAD and subsequently autobuild the structure. We also show that with better error treatment, interpretable maps can be obtained even with fewer measurements.

2. Methods

This work follows directly from the work reported in Brewster *et al.* (2018). The data set can be downloaded from cxi.db entry 81 (https://www.cxidb.org/id-81.html), and after indexing and integration consists of over 160 000 crystals from a thermolysin data set collected at the CXI endstation of LCLS on a CSPAD detector (Kern *et al.*, 2014; Hart *et al.*, 2012). After indexing, time-dependent ensemble refinement was applied, in which the data were grouped into batches of images and the detector models were then refined to account for the time-dependent shifts in sample position that are likely to arise from instability in the liquid-jetting system (Brewster *et al.*, 2018). The expected Bijvoet ratio for this system $(\langle|F_+ - F_-|\rangle/\langle F\rangle)$, comprising two Zn^{2+} and four Ca^{2+} atoms in a total of 2561 non-H atoms, is 2.1% (Terwilliger *et al.*, 2016; Hendrickson & Teeter, 1981).

Unlike in Brewster *et al.* (2018), the images were first converted from measured pixel values to photon units, dividing them by an estimated value of 25, as reported by the beamline staff. This experiment used an early-generation CSPAD with a non-uniform gain response; therefore, using a single gain-correction constant greatly oversimplifies the physics of the detector (Hart *et al.*, 2012). With these gain-corrected pixel values, we also needed to modify the merging protocol described in Brewster *et al.* (2018). We apply a per-image resolution filter during merging, in which the resolution cutoff of each image is determined by the point at which the signal-to-noise ratio (I/σ) falls below a given threshold. To compensate for the fact that I/σ decreases with the square root of the gain, we decreased the threshold from 0.5 to 0.1 [$0.1 = 0.5/(25)^{1/2}$].

We analyzed three methods for the treatment of error from SX data, as described in Sections 2.1–2.3. After the integrated intensity error estimates had been treated using one of these methods, we used them to create merged intensities I_h and merged error estimates σ_h according to the following procedure. Given a Miller index h with n measurements of the intensity of h, we define the jth measurement of h as $I^{\rm P}_{hj}$ and the associated photon-counting error as $\sigma^{\rm P}_{hj}$ [referred to as $\sigma_{\rm c}(I_{hj})$ in Brewster *et al.* (2018)]. The superscript P means that the reflection is only partially observed owing to the measurement being from a still image. The intensity and estimated error are both scaled to their full equivalent values, I_{hj} and σ_{hj}, using a per-image scale factor G_c, a Wilson B factor B_c and a per-reflection partiality correction P_{hj}, all of which were determined during scaling and post-refinement according to Sauter (2015),

$$I_{hj} = \frac{I^{\rm P}_{hj}}{K_{hj}}, \quad (1)$$

$$\sigma_{hj} = \frac{\sigma^{\rm P}_{hj}}{K_{hj}}, \quad (2)$$

$$K_{hj} = P_{hj} G_c \exp\left[-2B_c\left(\frac{\sin\theta_h}{\lambda_c}\right)^2\right], \quad (3)$$

where θ_h is the Bragg angle for Miller index h, λ_c is the incident wavelength and the subscript c denotes the crystal which gave rise to reflection hj. P_{hj} is the partiality-correction factor for this measurement [see equation (14) of Uervirojnangkoorn *et al.* (2015)], which depends on λ_c, mosaicity estimates and the unit-cell dimensions and orientation of crystal c. Importantly, the post-refinement of Sauter (2015) is similar to the post-refinement described in Winkler *et al.* (1979) and Rossmann *et al.* (1979) in that the target function refines the difference between the observed and predicted intensity values. However, the choice of the parameters being refined differs. Here, we refine the misorientation angles of the crystals, G_c, and B_c for each frame, but not the mosaicity itself, which is instead derived from empirically examining which reflections are observed on the image (Sauter *et al.*, 2014).

After frame-by-frame post-refinement, scaling and partiality correction, we merge the corrected intensities and error estimates according the three protocols detailed below and summarized in Table 1.

2.1. Protocol 1: unweighted mean

We begin with the suggestion of Schwarzenbach *et al.* (1989), in which we use the mean of the measurements to estimate the reflection intensity,

$$I_h = \frac{\sum_{j=1}^{n} I_{hj}}{n}, \quad (4)$$

and we use the observed spread of the measurements to determine the error estimates,

$$\sigma_{\rm res} = \left[\frac{\sum_{j=1}^{n}(I_{hj} - \langle I_h\rangle)^2}{n-1}\right]^{1/2}, \quad (5)$$

$$\sigma_h = \frac{\sigma_{\rm res}}{n^{1/2}}, \quad (6)$$

where $\sigma_{\rm res}$ refers to the residual differences between the measurements and their mean (*i.e.* the standard deviation), and σ_h refers to the merged error estimate for reflection h and is the standard error of the mean. Protocol 1 does not use the information in original error estimates from photon counting (σ_{hj}), and assumes that a large enough sample of the reflections is available to reliably estimate the uncertainty. This formulation is similar to that in Chapman *et al.* (2011) and White *et al.* (2012), differing slightly in the denominator of $\sigma_{\rm res}$ by using $n - 1$ instead of n.

Table 1
Summary of error-modeling methods.

Protocol	Weight†	Description
1	—	Unweighted error estimates
2	σ_{hj}	Photon-counting error estimates as weights
3	$\sigma_{\rm Ev11}$	Refine SDFAC terms to inflate photon-counting error estimates

† These are the weights used in (7) and (8), such that the weight $w = 1/\sigma^2$.

2.2. Protocol 2: weighted mean

The distribution of measurement intensities from still images does not follow a Gaussian distribution because every intensity is measured only partially. The reflection partiality is a function of crystal orientation, unit-cell dimensions, wavelength spectrum and crystal mosaicity. Difficulty in estimating these parameters results in integrating weak and highly partial reflections that skew the distribution towards zero. Because of the skewed distribution, the mean is not an ideal estimator of the structure-factor intensity, and so protocol 2 uses a weighted mean and a weighted standard error of the mean to estimate the reflection intensity and the uncertainty in that estimation,

$$I_h = \frac{\sum_{j=1}^{n} w_{hj} I_{hj}}{\sum_{j=1}^{n} w_{hj}}, \quad (7)$$

$$\sigma_h = \left(\frac{1}{\sum_{j=1}^{n} w_{hj}}\right)^{1/2}, \quad (8)$$

where the weights w_{hj} are variance weights derived from the photon-counting error estimates σ_{hj}, *i.e.* the estimated error derived from summing photons as described in Leslie (1999), which should follow a Poisson distribution:

$$w_{hj} = \frac{1}{\sigma^2_{hj}}. \quad (9)$$

2.3. Protocol 3: Ev11

Protocol 3 adjusts the error estimates using terms from Evans (2006) and Evans (2011): $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$.[1] In Brewster *et al.* (2018) we showed that applying these factors to the non-gain-corrected thermolysin data brings down the final merged I/σ estimate to around 30, which is more reasonable for protein crystallography (Diederichs, 2010). We also showed that applying these factors greatly increased the anomalous peak height of the Zn atom (from 44.6σ to 74.0σ). In Brewster *et al.* (2018), following the example of Evans (2011), our implementation used a simplex minimizer to refine these terms. In this work, we instead used a gradient-based nonlinear least-squares minimization procedure.

The equation to inflate the estimated error of the individual measurements is

[1] These terms are re-expressed from the Sdfac, SdB and Sdadd terms in Evans (2011), such that $s_{\rm fac}$ = Sdfac, $s_{\rm B}$ = (SdB)$^{1/2}$ and $s_{\rm add}$ = Sdadd.

$$\sigma^2_{\rm Ev11} = s^2_{\rm fac}[\sigma^2_{hj} + s^2_{\rm B}\langle I_h\rangle + s^2_{\rm add}\langle I_h\rangle^2], \qquad (10)$$

where $\langle I_h\rangle$ is the mean of the measurements of h after correcting by the factor K_{hj}. This equation is similar to error propagation, in which additional errors proportional to the intensity, likely derived from instrument instability ($s_{\rm add}$), are added in quadrature to the counting-error estimates σ_{hj}. In Evans (2011), the $s_{\rm fac}$ term is considered to account for effects such as errors in the gain, converting detector counts to photon counts. The $s_{\rm B}$ term was included to better fit the observed error estimates to a normal distribution, but in Evans (2011) the term was given no physical meaning. Here, we first show how we compute initial estimates of $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$ using normal probability analysis, following Evans (2006). We then use a limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS; Liu & Nocedal, 1989) minimizer to refine these parameters until the deviation of normalized error estimates best approaches 1.

After refinement of the $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$ terms, $1/\sigma^2_{\rm Ev11}$ is used as a weight in (7) and (8) to compute the weighted mean and weighted standard error of the mean of each reflection, as in protocol 2.

2.3.1. Initial parameter estimates. Estimates of error such as σ_{hj} represent the deviation of the measurements I_{hj} from the unknown population mean value. If these deviations from the mean are normally distributed then the normalized deviations will follow a standard normal distribution, *i.e.* a Gaussian distribution centered on zero with a standard deviation of 1. We choose initial values of $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$ that best adjust the original deviations such that the normalized deviations approach a standard normal distribution, according to the following procedure.

Normalized deviations. This formulation of the normalized deviations of a set of intensities and sigmas is similar to that described in Evans (2011), but includes the $(n-1)/n$ factor as currently implemented by *AIMLESS*. The normalized deviation $\delta_{hj{\rm norm}}$ for I_{hj} is

$$\delta_{hj{\rm norm}} = \left(\frac{n-1}{n}\right)^{1/2} \frac{I_{hj} - \langle I'_{hj}\rangle}{\sigma_{hj}}, \qquad (11)$$

where $\langle I'_{hj}\rangle$ is the mean of the measurements of h except for I_{hj}. In the special case where $n = 1$, $\langle I'_{hj}\rangle = 0$, and since in that case $n-1 = 0$, $\delta^2_{hj{\rm norm}} = 0$. These observations are not included in the normal probability analysis below for the initial parameter estimates.

Normal probability analysis. Using normalized deviations, we can initialize the $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$ parameters using a graphical technique called a 'normal probability plot', as suggested by Evans (2006) (see also Chambers *et al.*, 1983). A normal probability plot helps to determine how near a sampling of data approaches a normal distribution. Given a sampling of m observations, we sort them and then plot them versus a set of m theoretical or expected values. The theoretical values are perfectly distributed according to the normal distribution. If our observations are indeed normally distributed then the plot will be a straight line with slope 1 and offset 0. The 'perfect' theoretical values are normal order statistic medians, also referred to as rankits. In the simple case of $m = 5$ total observations, the second, third and fourth rankits are equal to the first quartile, median and third quartile of a normal distribution. We compute the rankits in the same way as *qqnorm* does in *R* (R Core Team, 2017). The rankit z_i for the ith value in m is

$$z_i = \Phi^{-1}\left(\frac{i-a}{m+1-2a}\right), \qquad (12)$$

where Φ^{-1} is the standard normal quantile function (the inverse of the cumulative distribution function) and where $a = 3/8$ if $m \le 10$ and 0.5 if $m > 10$. The expression $(i-a)/(m+1-2a)$ in (12) converts i to a number between 0 and 1; therefore, z_i is the expected value of the ranked ith sample from a normal distribution. Again, the normal probability plot, or the plot of the rankits versus $\delta_{hj{\rm norm}}$ (where all $\delta_{hj{\rm norm}}$ are first sorted by value), will have a slope of 1 with an offset of 0 if the error estimates are normally distributed. To determine an initial set of parameters, we determine the slope and offset of a line fitted to the central area of this plot (using the area between -0.5 and 0.5 to avoid fitting outliers). $s_{\rm fac}$ is initialized to the slope, as is performed in Evans (2006). In Evans (2006), $s_{\rm add}$ is set to 0.02. As we did not know whether this value was applicable to XFEL data, we experimented with initializing $s_{\rm add}$ to the normal probability plot offset and $s_{\rm B}$ to $s^{1/2}_{\rm add}$. This seemed to give reasonable results. As refinement proceeds, the normal probability plot becomes more linear and the slope approaches 1 as the parameters better correct the estimated errors to approach those derived from sampling a normal distribution (Fig. 1). Note that the normal probability analysis is only used to initialize the parameters; the refinement of the parameters is outlined below.

2.4. Parameter refinement

We refine the $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$ parameters using the LBFGS quasi-Newton minimizer requiring only first derivatives. For each step, we evaluate (10) for each σ_{hj} and then compute the normalized deviations using (11). The target function f_σ minimizes the deviation of the root-mean-squared deviation (r.m.s.d.) of the normalized deviations from 1, as determined over 100 intensity bins. We bin the intensities as follows. For each Miller index h, determine the mean intensity $\langle I_h\rangle$ of the measurements of h. The bin width will be the maximum of all $\langle I_h\rangle$ for all h minus the minimum $\langle I_h\rangle$ for all h divided by 100. For each h, all the measurements of h will be assigned to a single bin based on $\langle I_h\rangle$. There will be m_b measurements in intensity bin b. Call all the measurements in bin b I_{bk}, where k ranges from $k = 1$ to m_b. Each I_{bk} is associated with a normalized deviation, $\delta_{bk{\rm norm}}$, computed using the adjusted error estimate for that measurement of h,

$$\delta^2_{bk{\rm norm}} = \frac{n-1}{n}\frac{(I_{bk} - \langle I'_{hk}\rangle)^2}{\sigma^2_{\rm Ev11}}, \qquad (13)$$

where $\langle I'_{hk}\rangle$ is the mean of all measurements of Miller index h except for I_{bk}. Here, $\sigma_{\rm Ev11}$ is the corrected error estimate for measurement I_{bk} using (10) (note that the subscripts b and k

are suppressed in this reference to $\sigma_{\rm Ev11}$). The target function is then

$$f_\sigma = \sum_{b=1}^{100} w_b \left[1 - \left(\frac{\sum_{k=1}^{m_b} \delta^2_{bk{\rm norm}}}{m_b}\right)^{1/2}\right]^2, \qquad (14)$$

where b iterates over the 100 intensity bins. The term for each bin is weighted by $w_b = m_b^{1/2}$. After refinement of the $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$ parameters, we apply them to each σ_{hj} to compute the final estimated error for each measurement, $\sigma_{\rm Ev11}$.

The derivatives of the target function (14) with respect to the parameters are shown in Appendix *A*. The refinement of these terms using LBFGS is protocol 3.

3. Results

We reprocessed the data files from cxi.db entry 81 (Brewster *et al.*, 2018) comprising 160 000 lattices, including a gain correction (division of the pixel values by 25) prior to integration, and merged them using *cxi.merge*. In Brewster *et al.* (2018) the initial scale factors were derived from the known structure of thermolysin. Here, in contrast, we wanted to solve the structure *de novo*, so we used an alternate merging protocol. We first averaged all of the data without post-refinement using the *cxi.merge* default of weighted means and weighted standard errors of the mean (protocol 2). We then used this averaged data set as a scaling reference and merged again, applying post-refinement to each frame, refining the misorientation angles of the crystal, the scale factor and a Wilson *B* factor (Sauter, 2015), but again using the *cxi.merge* default of weighted means and weighted standard errors of the mean (protocol 2). We then re-merged a third time, using this post-refined data set as a reference for scaling. During this third merging, each of the three error models were applied. This bootstrapping approach to obtain a reference from the unscaled data is similar to how we have merged data before without a reference (Uervirojnangkoorn *et al.*, 2015). For protocol 3, the final values after refinement were $s_{\rm fac}$ = 1.32, $s_{\rm B}$ = 0.71 and $s_{\rm add}$ = 0.51.

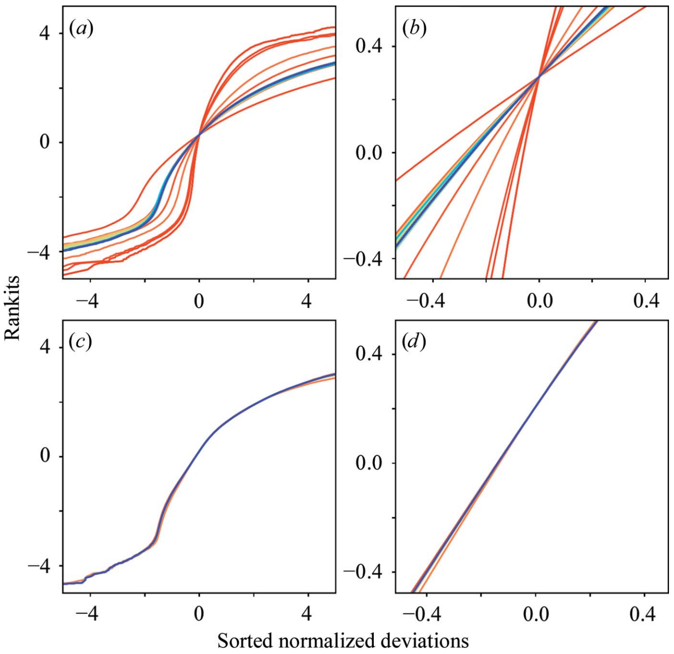


Figure 1
Normal probability plots for 5000 images. A 5000-image subset of the data was merged using protocol 3. During each step of the parameter refinement, a normal probability plot was generated (*a*). The rankits (equation 12) are plotted versus the sorted normalized deviations from the mean (equation 11). Each line represents one step during refinement and is colored using a rainbow color map from red (early steps) to blue (late steps). This is a non-gain-corrected data set. (*b*) Enlargement of the central area of (*a*) used to compute the slope and offset for initialization of the parameters. (*c*) As (*a*) but with a gain-corrected data set, in which each pixel was divided by 25. (*d*) As (*b*) for the central area of (*c*).

As mentioned above, we applied a gain correction to all images prior to integration, dividing the pixel values by 25 to convert to units of photons. As expected, correcting for gain also had a dramatic effect on the refinement of the SDFAC parameters for Ev11 (protocol 3). We processed a 5000-image subset without gain correction and found that refinement of the SDFAC parameters drove the functional (14) from 3306 to 122, driving the parameters from $s_{\rm fac}$ = 7.47, $s_{\rm B}$ = 0.72 and $s_{\rm add}$ = 0.52 to $s_{\rm fac}$ = 4.14, $s_{\rm B}$ = 0.00 and $s_{\rm add}$ = 0.52 over 66 steps. However, for the gain-corrected data, the refinement drove the functional from 156 to 149, driving the parameters from $s_{\rm fac}$ = 1.44, $s_{\rm B}$ = 0.67 and $s_{\rm add}$ = 0.45 to $s_{\rm fac}$ = 1.43, $s_{\rm B}$ = 0.96 and $s_{\rm add}$ = 0.45 over 14 steps. The difference between the two refinements can be seen in Fig. 1. Not only did a more substantial minimization need to be performed on the non-gain-corrected data, but the final $s_{\rm fac}$ parameter is quite a bit larger in magnitude, indicating a compensation for the absence of a gain correction. It is also worth noting that the difference between the two initial $s_{\rm fac}$ values is related to the gain ratio ($7.47^2/1.44^2 = 26.9$), again indicating the relationship between $s_{\rm fac}$ and the uncertainty in the gain estimate.

Properly scaled, partiality-corrected and merged intensities reported in units of photons from XFELs should be comparable to the full reflection intensities measured at synchrotrons if all systematic effects have been accounted for. One measure of comparison for the two techniques is the signal to noise, or the I/σ ratio. Fig. 2 shows I, σ and I/σ versus resolution plots for the merged data, which show that the error estimates for protocol 2 are orders of magnitude lower than for protocols 1 and 3. Fig. 3 shows I/σ versus I plots for all three data sets, as presented in Diederichs (2010) (note that these are for the unmerged data). While Diederichs (2010) was working with reflections that were much better measured and had much lower redundancy, I/σ in photons should be comparable (between 20–40), and indeed we see the overall I values are of the order that is expected (10^0–10^4). Protocols 1 and 3 show I/σ values of the order that would be expected from protein crystallography, while protocol 2 has I/σ values that are higher than expected. We also see that data sets 1 and 3 do not display the sigmoidal shape demonstrated in Diederichs (2010), indicating that the signal-to-noise ratio has not reached its limit for this system. This implies there is further work to be performed to remove systematic errors. Finally, note that the scattered data points in protocol 1 (upper left of Fig. 3) that have high I/σ but low I come from reflections with low redundancy (≤2–4). These error estimates, which for protocol 1 come only from the standard error of the mean of the observations, become unreliable with low redundancy. It is likely that a redundancy of at least 5 is required for SX data to be reliable using protocol 1.

We also examined the overall I/σ trends in the data set. In Hattne *et al.* (2014), we observed numerous intensities at large negative multiples of I/σ, and we used these negative measurements to compute an additional error-adjustment term to account for this extra uncertainty. To determine whether this approach (termed Ha14; see also Brewster *et al.*, 2018) was applicable to the data in this work, we examined the distribution of I/σ in a subset of images. We selected integration regions void of Bragg spots and compared the distribution of I/σ between these empty measurements and the measurements where signal is predicted. We found that the large negative intensity outliers seen in Hattne *et al.* (2014) are absent from our data and that the negative intensities have a similar distribution to the empty measurements (see Fig. 4). Therefore, Ha14 methods do not seem to apply.

Phasing and autobuilding was performed using *phenix.autosol* (Adams *et al.*, 2010), supplying the thermolysin amino-acid sequence from PDB entry 4tnl (Kern *et al.*, 2014), one NCS copy and using all defaults, except for specifying two Zn atoms as the search target, using a thorough *HySS* search to 4.0 Å and using a solvent fraction of 0.467 with extreme density modification. Phasing results are shown in Table 2.

While all protocols were able to find the six heavy-atom sites, protocol 2 essentially failed during SAD phasing and autobuilding, while the unweighted protocol 1 partially succeeded. Over two thirds of the structure was built with

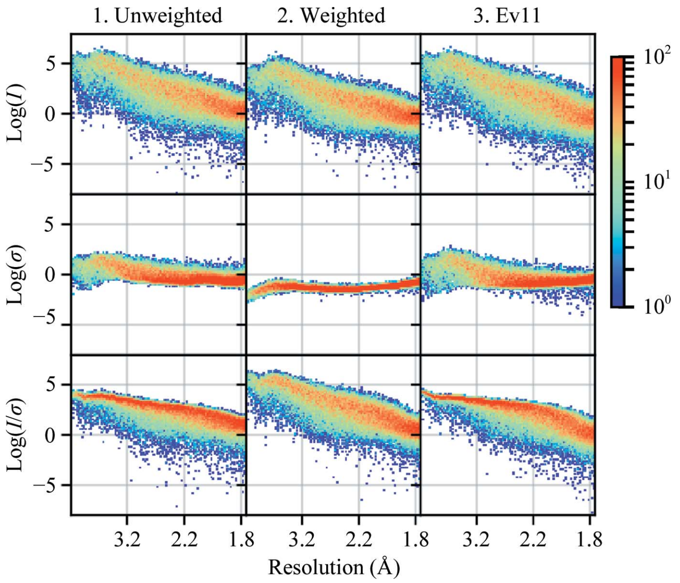


Figure 2
Intensity and σ versus resolution. 2D histograms of I, σ and I/σ (top, middle and bottom) versus resolution for the three error models. Data are for merged values. Note that the y axes and the color are on a logarithmic scale.

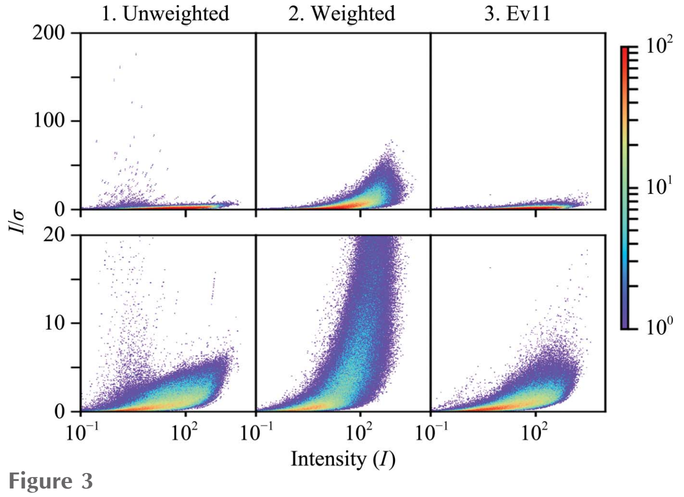


Figure 3
I/σ versus I plots with different error models. 2D histograms of I/σ versus I for the three error models. Unmerged intensities and error estimates are shown. In the top and bottom plots the same data are presented but with different scales for the y axis. Note that the color is on a logarithmic scale.

Table 2
SAD phasing results for different error models.

Values in parentheses are for the highest resolution bin.

Protocol	1	2	3
Weight†	—	σ_{hj}	$\sigma_{\rm Ev11}$
Resolution (Å)	80.78–1.80 (1.86–1.80)		
I/σ‡	13.8 (2.7)	59.7 (2.0)	14.0 (1.4)
$CC_{1/2}$ (%)	99.9 (73.8)	99.8 (63.3)	99.9 (81.4)
Zn^{2+} peak height (σ)	53.1	50.5	67.0
No. of sites found by *HySS*§	6 ± 0	6 ± 0	6 ± 0
No. of residues built (of 316)§	252.4 ± 15.2	104.1 ± 1.4	297.2 ± 6.5
Model–map CC§¶ (%)	71.0 ± 0.4	30.3 ± 0.2	80.0 ± 0.1
$R_{\rm work}$§ (%)	27.4 ± 1.8	54.8 ± 0.3	21.2 ± 1.3
$R_{\rm free}$§ (%)	29.9 ± 2.0	57.3 ± 0.8	23.7 ± 1.7

† As in Table 1, for a given weight w where $w = 1/\sigma^2$. ‡ These are higher than in Fig. 3 because the higher intensity observations are given a higher weight during merging. § Numbers are mean ± standard deviation over ten trials with differing random-number seeds. ¶ Phased map correlation to the known structure.

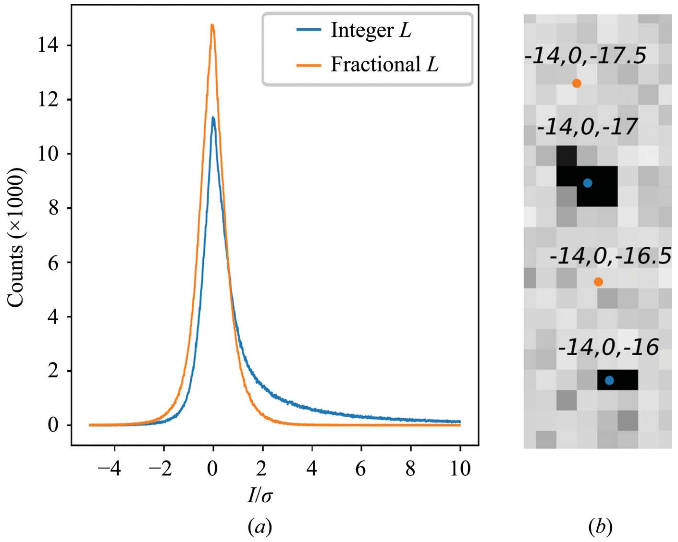


Figure 4
Histogram of I/σ for signal versus noise. (*a*) A random subset of 3800 images from one processing run of thermolysin was re-integrated, including the prediction of non-existent reflections at the halfway positions along the c^* axis. These predictions, which are halfway between observed reflections, are composed of only noise. (*b*) Example of reflections labeled with integer L and fractional L indices.

protocol 1 and it is likely that the model could be finished manually. Phasing and autobuilding were successful using error estimates inflated by SDFAC parameters (protocol 3). This protocol also showed an improved ability to phase and autobuild the structure compared with using the unweighted variance (protocol 1). The LBFGS version of SDFAC refinement shows nearly the same results as a simplex minimizer (not shown), but importantly LBFGS is deterministic, does not rely on the randomness in the initialization inherent to simplex minimization and converges in less time and in fewer steps than the simplex minimizer (see below).

To determine whether these algorithms improve the number of images needed for phasing, for each of the three protocols we re-merged the data using increasing numbers of images from 1000 images to the full data set (160 000+ images; Fig. 5). In addition, since we used random sampling to create these subsets, we repeated this sampling ten times for each subset. For the full data set, which could not be sub-sampled, we instead ran the auto-solver ten times with random seeds, as recommended by Bunkóczi *et al.* (2015).

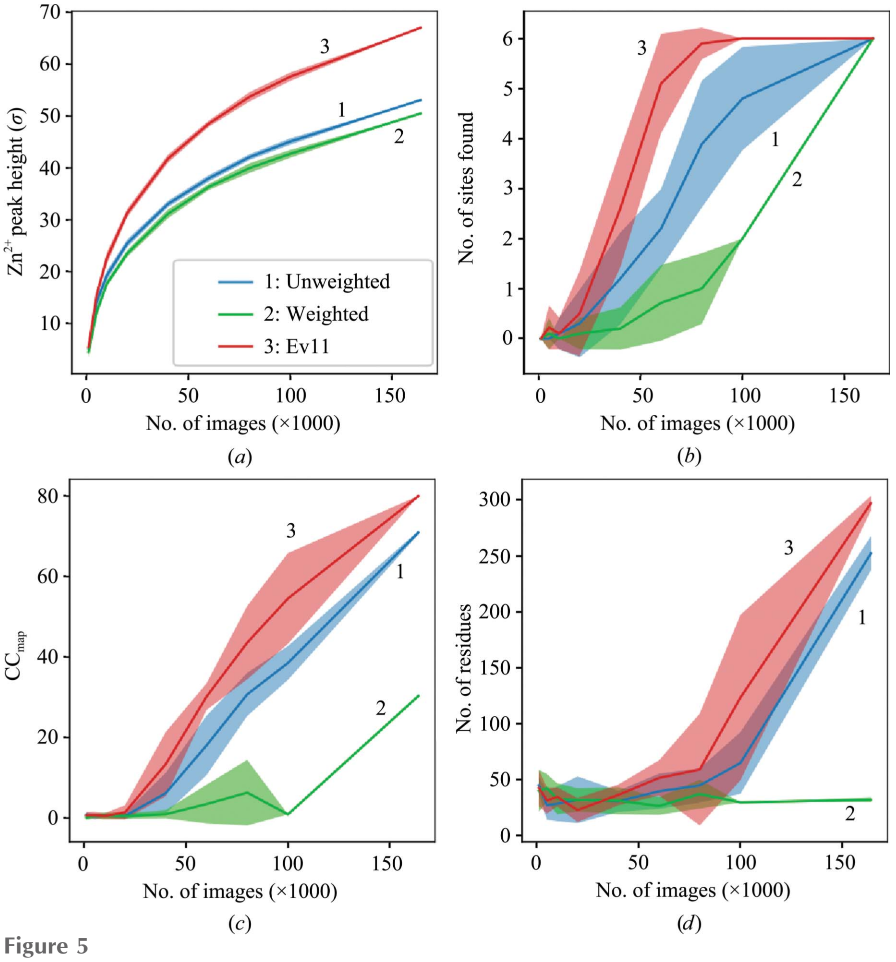


Figure 5
Effect of image count on autobuilding success. For each of the three protocols, increasing numbers of images were processed. The anomalous peak height for the Zn^{2+} atom (*a*), the number of heavy-atom sites found (out of six) (*b*), the known model-to-map CC (*c*) and the number of residues built (*d*) are shown versus the number of images in the data set. In each case, shaded areas indicate the standard deviation of either the ten subsamples (data sets 1000–100 000) or the ten random seeds (full data set, 164 063 images). Note that for (*b*) certain data points have the same number of sites found in all trials and hence have no standard deviation.

We found that for this zinc SAD phasing experiment we still needed nearly all of the images to autobuild the structure. Autobuilding built about half of the structure with 100 000 images with protocol 3 (Fig. 5*d*), but failed with fewer images and with the other protocols. However, we can still examine the phasing ability of the data by examining the Zn^{2+} anomalous peak height (Fig. 5*a*), the CC_{map} statistic, which is the correlation of the phased map with the known structure with PDB code 1lnd (Holland *et al.*, 1995) (Fig. 5*c*), and the number of sites found by *phenix.hyss* out of the six possible, as determined by using *phenix.emma* to match sites between the known structure and the SAD-determined sites (Fig. 5*b*). SDFAC treatment improves the result over the residual-only treatment (compare protocols 1 and 3). Protocol 2 consistently underperformed.

Finally, a note on the performance of simplex versus LBFGS. Using derivative-based minimization drives the optimization to a similar solution using fewer steps. During one trial (not shown) with 10 000 images, the simplex refiner took 88 steps over 932.8 s. However, the LBFGS minimizer took 51 steps over 444.5 s. Both implementations are in Python with C++ sections for the computing-intensive portions. The further addition of OpenMP multiprocessing during the C++ computation of the normalized deviations and derivatives reduced the LBFGS runtime to 322 s (64 cores, accelerating equations 10, 11 and 15).

4. Discussion

The phasing of serial crystallographic data has been notoriously difficult. Unlike in the rotation method, where the integration, scaling, error-treatment and merging protocols have been well studied, in SX the algorithms continue to be refined to account for the sparseness of the data recorded from each crystal. Most of the few *de novo* XFEL structures in the PDB required tens to hundreds of thousands of images to solve (Barends *et al.*, 2014; Nakane *et al.*, 2015, 2016; Colletier *et al.*, 2016; Nass *et al.*, 2016; Hunter *et al.*, 2016; Yamashita *et al.*, 2015; Gorel *et al.*, 2017). In this work, we have shown that some of the difficulty arises from how the error estimates are treated and used when merging SX data. In addition to affecting the final merged intensity by being used in the weighted sum, the merged error estimates themselves are used extensively in the maximum-likelihood techniques used by phasing algorithms. For example, in McCoy *et al.*

(2004), equation (2) describes the probability of the magnitudes of a set of unphased structure factors given a set of phased structure-factor vectors. While this equation uses only the intensities as inputs, a set of adjustments propagated from the estimated experimental error is presented in Appendix *B* as initial values used in maximum-likelihood refinement. With this, it is not surprising that accurate estimates of the errors are useful, but it is notable here how striking the difference improving the estimates of errors in the measured reflections makes in the ability to phase XFEL data.

It also interesting to note how important using a good set of weights is when merging the data using a weighted mean. Protocol 1 (unweighted mean) performed consistently better than protocol 2 (weighted mean), even though a weighted mean should be a better estimator of a population mean (especially for the left-skewed intensity distributions seen in XFEL data). Stated differently, using the photon-counting error estimates alone as weights is not optimal, at least in terms of this anomalous phasing exercise. It is only after adjusting the individual measurement error estimates such that they better explain the observed variance through the Ev11 approach that using the error estimates as weights improves the results over the unweighted mean (protocol 3, Ev11).

By no means do we assert that the methods presented here are an exhaustive list of possible ways of treating the errors in XFEL data collection. While we want to note that using the initial estimates of error from integration and applying adjustments to bring them closer to explaining the observed error is important for *de novo* phasing success, at least when using a weighted sum for averaging intensities together, none of the methods here propagate the error from the partiality correction itself. From (1)–(3) we see that the intensities are corrected using a partiality term, a scale factor and a Wilson *B* factor. (2) propagates the error in inflating I_{hj} by K_{hj}, assuming that K_{hj} is a constant, but in reality the parameters comprising K_{hj} are refined quantities. The partiality is dependent on the crystal orientation, unit-cell dimensions, wavelength spectrum and estimated mosaicity (Sauter, 2015), and while the true errors in these terms are unknown, they could be estimated from the population of crystals used for merging and then propagated to the factor K_{hj} (see Appendix *B*). Likewise, the estimated error estimates of these terms could be refined. Initial efforts in this direction have been made and can be accessed using experimental parameters in *cctbx.xfel*. While this still will not account for the full set of unknown random and systematic errors present in SX data collection, any error propagation in this manner should reduce the reliance on inflationary terms to account for the observed variance in the sample.

5. Software availability

Instructions for downloading and using *cctbx.xfel* are available at the cctbx.xfel wiki at http://cci.lbl.gov/xfel. See also Brewster *et al.* (2019) for instructions on using the *cctbx.xfel* graphical user interface (GUI).

APPENDIX *A*
Derivatives of the target function

As part of least-squares minimization we take the partial derivative of (14) with respect to each of the $s_{\rm fac}$, $s_{\rm B}$ and $s_{\rm add}$ parameters, which we refer collectively here as the parameters p. We can perform this via (10), (13), (14) and the chain rule. We first take derivatives with respect to the square of each parameter:

$$\begin{aligned}\frac{\partial f_\sigma}{\partial p^2} = 2\sum_{b=1}^{100} w_b \Biggl\{ \Biggl[1 - \left(\frac{\sum_{k=1}^{m_b} \delta^2_{bk{\rm norm}}}{m_b} \right)^{1/2} \Biggr] \\ \times \Biggl[-\frac{1}{2} \left(\frac{\sum_{k=1}^{m_b} \delta^2_{bk{\rm norm}}}{m_b} \right)^{-1/2} \Biggr] \frac{\sum_{k=1}^{m_b} \frac{\partial \delta^2_{bk{\rm norm}}}{\partial p^2}}{m_b} \Biggr\}. \qquad (15)\end{aligned}$$

Since the intensity value as used in the computation of $\delta_{bk{\rm norm}}$ does not depend on the parameters being refined,

$$\frac{\partial \delta^2_{bk{\rm norm}}}{\partial p^2} = \frac{\partial \delta^2_{bk{\rm norm}}}{\partial \sigma^2_{\rm Ev11}} \frac{\partial \sigma^2_{\rm Ev11}}{\partial p^2} = -\frac{n-1}{n} \frac{(I_{bk} - \langle I'_{hk} \rangle)^2}{(\sigma^2_{\rm Ev11})^2} \frac{\partial \sigma^2_{\rm Ev11}}{\partial p^2}. \qquad (16)$$

We can now compute the partial derivatives of (10) with respect to the parameters p. Note that the minimizer refines the terms themselves instead of the squares of the terms, and that $(\partial p^2/\partial p) = 2p$. Therefore,

$$\frac{\partial \sigma^2_{\rm Ev11}}{\partial s_{\rm fac}} = \frac{\partial \sigma^2_{\rm Ev11}}{\partial s^2_{\rm fac}} \frac{\partial s^2_{\rm fac}}{\partial s_{\rm fac}} = (\sigma^2_{hj} + s^2_{\rm B}\langle I_h \rangle + s^2_{\rm add}\langle I_h \rangle^2) 2 s_{\rm fac}, \qquad (17)$$

$$\frac{\partial \sigma^2_{\rm Ev11}}{\partial s_{\rm B}} = \frac{\partial \sigma^2_{\rm Ev11}}{\partial s^2_{\rm B}} \frac{\partial s^2_{\rm B}}{\partial s_{\rm B}} = s^2_{\rm fac}\langle I_h \rangle 2 s_{\rm B}, \qquad (18)$$

$$\frac{\partial \sigma^2_{\rm Ev11}}{\partial s_{\rm add}} = \frac{\partial \sigma^2_{\rm Ev11}}{\partial s^2_{\rm add}} \frac{\partial s^2_{\rm fac}}{\partial s_{\rm add}} = s^2_{\rm fac}\langle I_h \rangle^2 2 s_{\rm add}. \qquad (19)$$

APPENDIX *B*
Error propagation for partial reflections

In this work, for each reflection we compute a scaling term K_{hj} that includes the partiality correction, the Wilson *B* factor and the scaling factor *G*. K_{hj} depends on the crystal orientation, unit-cell parameters, wavelength, mosaicity and so forth (equations 1 and 3). A simple error propagation including the photon-counting error $\sigma^{\rm P}_{hj}$ and assuming no error in K_{hj} is shown in (2). However, if estimates of the errors in terms comprising K_{hj} were available they could be propagated, and the first few steps of this process are shown here. Given parameters p (p_1, p_2, ...) that contribute to K_{hj}, the propagated error is

$$\sigma^2_{hj} = (\sigma^{\rm P}_{hj})^2 \left(\frac{\partial \sigma^{\rm P}_{hj}}{\partial p_1} \right)^2 + (\sigma^{\rm P}_{hj})^2 \left(\frac{\partial \sigma^{\rm P}_{hj}}{\partial p_2} \right)^2 + \ldots, \qquad (20)$$

where again $\sigma^{\rm P}_{hj}$ is the photon-counting error and σ^2_{hj} is the propagated error. By the chain rule,

$$\begin{aligned}\sigma_{hj}^2 &= (\sigma_{hj}^{\rm P})^2\left(\frac{\partial\sigma_{hj}^{\rm P}}{\partial K_{hj}}\right)^2\left(\frac{\partial K_{hj}}{\partial p_1}\right)^2 + (\sigma_{hj}^{\rm P})^2\left(\frac{\partial\sigma_{hj}^{\rm P}}{\partial K_{hj}}\right)^2\left(\frac{\partial K_{hj}}{\partial p_2}\right)^2 + \ldots \\ &= \frac{(\sigma_{hj}^{\rm P})^2}{K_{hj}^2}\left(\frac{\partial K_{hj}}{\partial p_1}\right)^2 + \frac{(\sigma_{hj}^{\rm P})^2}{K_{hj}^2}\left(\frac{\partial K_{hj}}{\partial p_2}\right)^2 + \ldots, \qquad (21)\end{aligned}$$

which reduces to (2) if the errors in the parameters p are ignored.

Initial implementation of these and further derivatives of K_{hj} with respect to the parameters p as well as refinement of the associated error terms is available through experimental options in *cctbx.xfel*.

Acknowledgements

We thank Phil Evans, James Holton and Jan Kern for productive conversations regarding error modeling.

Funding information

This research was supported by NIH grant GM117126. NKS acknowledges support from the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the US Department of Energy (DOE) Office of Science and the National Nuclear Security Administration. Further support originates from the National Institute of General Medical Sciences of the National Institutes of Health under Award R01GM109019 and the Advanced Scientific Computing Research and the Basic Energy Sciences programs, which are supported by the Office of Science of the US Department of Energy (DOE) under Contract DE-AC02-05CH11231. Portions of this research were carried out at LCLS at the SLAC National Accelerator Laboratory, supported by the DOE Office of Science, OBES under Contract No. DE-AC02-76SF00515. Data processing was performed in part at the National Energy Research Scientific Computing Center, supported by the DOE Office of Science, Contract No. DEAC02-05CH11231.

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.* D**66**, 213–221.

Barends, T. R. M., Foucar, L., Botha, S., Doak, R. B., Shoeman, R. L., Nass, K., Koglin, J. E., Williams, G. J., Boutet, S., Messerschmidt, M. & Schlichting, I. (2014). *Nature (London)*, **505**, 244–247.

Bergmann, U., Yachandra, V. & Yano, J. (2017). *X-ray Free Electron Lasers.* Cambridge: The Royal Society of Chemistry.

Brewster, A. S., Waterman, D. G., Parkhurst, J. M., Gildea, R. J., Young, I. D., O'Riordan, L. J., Yano, J., Winter, G., Evans, G. & Sauter, N. K. (2018). *Acta Cryst.* D**74**, 877–894.

Brewster, A. S., Young, I. D., Lyubimov, A., Bhowmick, A. & Sauter, N. K. (2019). *Comput. Crystallogr. Newslett.* **10**, 22–39.

Bunkóczi, G., McCoy, A. J., Echols, N., Grosse-Kunstleve, R. W., Adams, P. D., Holton, J. M., Read, R. J. & Terwilliger, T. C. (2015). *Nature Methods*, **12**, 127–130.

Chambers, J. M., Cleveland, W. S., Kleiner, B. & Tukey, P. A. (1983). *Graphical Methods for Data Analysis*, ch. 6. Belmont: Wadsworth.

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., Gorke, 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., Andritschke, 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., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature (London)*, **470**, 73–77.

Colletier, J.-P., Sawaya, M. R., Gingery, M., Rodriguez, J. A., Cascio, D., Brewster, A. S., Michels-Clark, T., Hice, R. H., Coquelle, N., Boutet, S., Williams, G. J., Messerschmidt, M., DePonte, D. P., Sierra, R. G., Laksmono, H., Koglin, J. E., Hunter, M. S., Park, H.-W., Uervirojnangkoorn, M., Bideshi, D. K., Brunger, A. T., Federici, B. A., Sauter, N. K. & Eisenberg, D. S. (2016). *Nature (London)*, **539**, 43–47.

Diederichs, K. (2010). *Acta Cryst.* D**66**, 733–740.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.

Ginn, H. M., Brewster, A. S., Hattne, J., Evans, G., Wagner, A., Grimes, J. M., Sauter, N. K., Sutton, G. & Stuart, D. I. (2015). *Acta Cryst.* D**71**, 1400–1410.

Gorel, A., Motomura, K., Fukuzawa, H., Doak, R. B., Grünbein, M. L., Hilpert, M., Inoue, I., Kloos, M., Kovácsová, G., Nango, E., Nass, K., Roome, C. M., Shoeman, R. L., Tanaka, R., Tono, K., Joti, Y., Yabashi, M., Iwata, S., Foucar, L., Ueda, K., Barends, T. R. M. & Schlichting, I. (2017). *Nature Commun.* **8**, 1170.

Hart, P., Boutet, S., Carini, G., Dubrovin, M., Duda, B., Fritz, D., Haller, G., Herbst, R., Herrmann, S., Kenney, C., Kurita, N., Lemke, H., Messerschmidt, M., Nordby, M., Pines, J., Schafer, D., Swift, M., Weaver, M., Williams, G., Zhu, D., Van Bakel, N. & Morse, J. (2012). *Proc. SPIE*, **8504**, 85040C.

Hattne, J., Echols, N., Tran, R., Kern, J., Gildea, R. J., Brewster, A. S., Alonso-Mori, R., Glöckner, C., Hellmich, J., Laksmono, H., Sierra, R. G., Lassalle-Kaiser, B., Lampe, A., Han, G., Gul, S., DiFiore, D., Milathianaki, D., Fry, A. R., Miahnahri, A., White, W. E., Schafer, D. W., Seibert, M. M., Koglin, J. E., Sokaras, D., Weng, T. C., Sellberg, J., Latimer, M. J., Glatzel, P., Zwart, P. H., Grosse-Kunstleve, R. W., Bogan, M. J., Messerschmidt, M., Williams, G. J., Boutet, S., Messinger, J., Zouni, A., Yano, J., Bergmann, U., Yachandra, V. K., Adams, P. D. & Sauter, N. K. (2014). *Nature Methods*, **11**, 545–548.

Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.

Holland, D. R., Hausrath, A. C., Juers, D. & Matthews, B. W. (1995). *Protein Sci.* **4**, 1955–1965.

Hunter, M. S., Yoon, C. H., DeMirci, H., Sierra, R. G., Dao, E. H., Ahmadi, R., Aksit, F., Aquila, A. L., Ciftci, H., Guillet, S., Hayes, M. J., Lane, T. J., Liang, M., Lundström, U., Koglin, J. E., Mgbam, P., Rao, Y., Zhang, L., Wakatsuki, S., Holton, J. M. & Boutet, S. (2016). *Nature Commun.* **7**, 13388.

Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.

Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.

Kabsch, W. (2014). *Acta Cryst.* D**70**, 2204–2216.

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). *Nature Commun.* **5**, 4371.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.

Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.

Liu, D. C. & Nocedal, J. (1989). *Math. Program.* **45**, 503–528.

McCoy, A. J., Storoni, L. C. & Read, R. J. (2004). *Acta Cryst.* D**60**, 1220–1228.

Nakane, T., Hanashima, S., Suzuki, M., Saiki, H., Hayashi, T., Kakinouchi, K., Sugiyama, S., Kawatake, S., Matsuoka, S., Matsumori, N., Nango, E., Kobayashi, J., Shimamura, T., Kimura, K., Mori, C., Kunishima, N., Sugahara, M., Takakyu, Y., Inoue, S., Masuda, T., Hosaka, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Inoue, T., Nureki, O., Iwata, S., Murata, M. & Mizohata, E. (2016). *Proc. Natl Acad. Sci. USA*, **113**, 13039–13044.

Nakane, T., Song, C., Suzuki, M., Nango, E., Kobayashi, J., Masuda, T., Inoue, S., Mizohata, E., Nakatsu, T., Tanaka, T., Tanaka, R., Shimamura, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Iwata, S. & Sugahara, M. (2015). *Acta Cryst.* D**71**, 2519–2525.

Nass, K., Meinhart, A., Barends, T. R. M., Foucar, L., Gorel, A., Aquila, A., Botha, S., Doak, R. B., Koglin, J., Liang, M., Shoeman, R. L., Williams, G., Boutet, S. & Schlichting, I. (2016). *IUCrJ*, **3**, 180–191.

Otwinowski, Z. & Minor, W. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 226–235. Dordrecht: Kluwer Academic Publishers.

R Core Team, (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Rossmann, M. G. & Arnold, E. (2001). Editors. *International Tables for Crystallography*, Vol. *F*, ch. 11. Dordrecht: Kluwer Academic Publishers.

Rossmann, M. G., Leslie, A. G. W., Abdel-Meguid, S. S. & Tsukihara, T. (1979). *J. Appl. Cryst.* **12**, 570–581.

Sauter, N. K. (2015). *J. Synchrotron Rad.* **22**, 239–248.

Sauter, N. K., Hattne, J., Brewster, A. S., Echols, N., Zwart, P. H. & Adams, P. D. (2014). *Acta Cryst.* D**70**, 3299–3309.

Schwarzenbach, D., Abrahams, S. C., Flack, H. D., Gonschorek, W., Hahn, T., Huml, K., Marsh, R. E., Prince, E., Robertson, B. E., Rollett, J. S. & Wilson, A. J. C. (1989). *Acta Cryst.* A**45**, 63–75.

Terwilliger, T. C., Bunkóczi, G., Hung, L.-W., Zwart, P. H., Smith, J. L., Akey, D. L. & Adams, P. D. (2016). *Acta Cryst.* D**72**, 346–358.

Uervirojnangkoorn, M., Zeldin, O. B., Lyubimov, A. Y., Hattne, J., Brewster, A. S., Sauter, N. K., Brunger, A. T. & Weis, W. I. (2015). *Elife*, **4**, e05421.

White, T. A. (2014). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130330.

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.

Winkler, F. K., Schutt, C. E. & Harrison, S. C. (1979). *Acta Cryst.* A**35**, 901–911.

Yamashita, K., Pan, D., Okuda, T., Sugahara, M., Kodan, A., Yamaguchi, T., Murai, T., Gomi, K., Kajiyama, N., Mizohata, E., Suzuki, M., Nango, E., Tono, K., Joti, Y., Kameshima, T., Park, J., Song, C., Hatsui, T., Yabashi, M., Iwata, S., Kato, H., Ago, H., Yamamoto, M. & Nakatsu, T. (2015). *Sci. Rep.* **5**, 14017.