Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

De-icing: recovery of diffraction intensities in the presence of ice rings

Michael S. Chapman[a]* and Thayumanasamy Somasundaram[b]

[a]Department of Biochemistry and Molecular Biology, School of Medicine, Oregon Health and Science University, Portland, OR 97239-3098, USA, and [b]Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306-4380, USA

Correspondence e-mail: chapmami@ohsu.edu

Received 16 December 2009
Accepted 2 April 2010

PDB Reference: arginine kinase, 3m10.

Macromolecular structures are routinely determined at cryotemperatures using samples flash-cooled in the presence of cryoprotectants. However, sometimes the best diffraction is obtained under conditions where ice formation is not completely ablated, with the result that characteristic ice rings are superimposed on the macromolecular diffraction. In data processing, the reflections that are most affected by the ice rings are usually excluded. Here, an alternative approach of subtracting the ice diffraction is tested. High completeness can be retained with little adverse effect upon the quality of the integrated data. This offers an alternate strategy when high levels of cryoprotectant lead to loss of crystal quality.

1. Introduction

The collection of macromolecular diffraction data at cryogenic temperatures offers a way to mitigate radiation damage and has become the standard for high-resolution data sets (Garman & Owen, 2006). Samples are flash-cooled to avoid the formation of crystalline (hexagonal) ice which would disrupt the order of the macromolecular crystal lattice. One of several cryoprotective agents is added to inhibit the nucleation of ice crystals before a vitreous glass is formed (Rodgers, 2001). The addition of cryoprotectants is often deleterious, so procedures have been developed to search for the best reagent, to use concentrations only a few percent above minimal inhibitory levels and to change the conditions gradually to minimize osmotic and other stresses on the macromolecular lattice (Rodgers, 2001). Even small amounts of microcrystalline ice lead to characteristic rings of powder diffraction. These can usually be avoided by refinement of the cryoprotection scheme, but sometimes the highest resolution diffraction is only obtained in the presence of some ice diffraction.

This was the case with the current test case, arginine kinase, an enzyme belonging to the creatine kinase family that plays a role in maintaining ATP concentrations in cellular energy homeostasis (Ellington, 2001). Multiple attempts at cryo-data collection with various cryoprotectants had yielded data sets to ~3 Å resolution for the substrate-free form. In attempts to avoid cryoprotectants, new conditions were found in which crystals were grown by equilibration against 26% PEG 5000 MME. Preliminary experiments on a home source showed diffraction to 1.9 Å resolution without evidence of ice formation. Larger crystals were obtained by macroseeding into the same conditions and these crystals yielded diffraction to beyond 1.7 Å resolution on APS beamline 14-BM-C, but now with prominent ice rings. These data were truncated at 2.35 Å resolution, before the second ice ring, and used to solve the structure (PDB code 1m80; Yousef *et al.*, 2003). Several attempts at obtaining an improved ice-free data set were unsuccessful, motivating the current attempt to recover useful data by subtraction of the ice component.

2. Methods

A Python program, *DeIce*, has been developed to subtract ice powder rings from diffraction images in a pre-processing step prior to integration. While not ideal (see §3), pre-processing was a convenient means to test feasibility while avoiding the challenges in programming (and sometimes licensing) changes to existing integration

packages. *DeIce* also provides statistics on the radial distribution of intensity, which are helpful in choosing appropriate parameters.

The radial distribution of intensity (Fig. 1*b*) is strikingly similar to that attributed to cubic ice in Fig. 3 of Dowell & Rinfret (1960), with the three strongest maxima corresponding to the (111), (220) and (311) reflections at 3.67, 2.25 and 1.92 Å, respectively. Just as in the earlier study of ice, the presence of some hexagonal ice can be inferred from the shoulder in the cubic (111) peak that can be attributed to the (111) reflection of hexagonal ice at 3.90 Å resolution. The breadth of the peaks can also be attributed to the presence of a vitreous ice; the low-density amorphous (LDA) phase also has a scattering maximum at 3.67 Å (Kim *et al.*, 2008). Indeed, Dowell and Rinfret commented that the cubic phase was never observed in the complete absence of the vitreous phase and there is a precedent for seeing a similar mixture of phases in the context of a protein crystal in the case of glucose isomerase when pressure-frozen then warmed to 220 K (see Fig. 4*a* of Kim *et al.*, 2008). Although perhaps hexagonal ice is more common in macromolecular diffraction, the predominance of the cubic form here is consistent with a cryoprotectant level that is not quite sufficient to completely inhibit ice-crystal formation (Berejnov *et al.*, 2006).

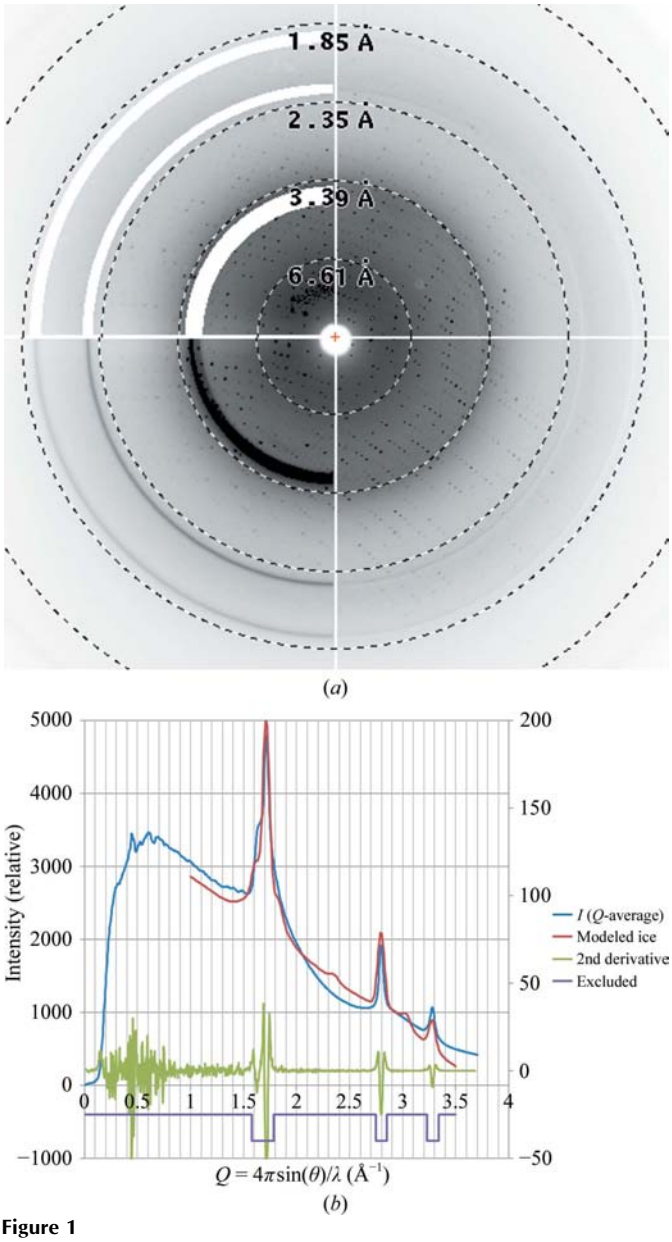

Figure 1
Effects of ice-ring subtraction upon diffraction images. (*a*) A representative diffraction image is divided into three areas. The bottom-left quartile shows the original image. Ice rings at ∼3.7, 2.2 and 1.9 Å are clearly visible. The top-left quartile shows the effect of zeroing the intensity in rings at these resolutions. The right half shows part of the same image following subtraction of the ice diffraction. (*b*) Variation in the intensity with scattering angle, binning pixels at a common distance from the direct beam. Blue, median intensity. Green, the second derivative of the median intensity, highlighting the inadequacy of a background plane fit during integration where the image is affected by ice rings. Brown, intensity calculated for a 5:1:5 mixture of cubic, hexagonal and amorphous ice that was fitted approximately to the observed intensity together with a linearly changing background. Purple, the result of subtracting the ice diffraction, assuming that the ice-removed background can be interpolated linearly through the ice ring. (*a*) was prepared using the output from *ADXV* (Andrew Arvai, Area Detector Systems Corp.).

It was possible to obtain semi-quantitative agreement between the observed spectrum and intensity modeled from the known *d*-spacings and the relative intensities for each reflection in each of the ice phases (Blackman & Lisgarten, 1957; Dowell & Rinfret, 1960). In empirically improving the fit (Fig. 1*b*), the widths of the Gaussian peaks were controlled by a single parameter for each phase and the weight for each phase was adjusted. The fitted model indicates an approximately 5:1:5 ratio of the cubic, hexagonal and vitreous forms, respectively. While the *d*-spacings could be fitted to the expected values with high internal consistency (±0.05%), quantitative agreement between modeled and observed intensity values might prove elusive with uncertainties in the underlying background and the effects of the distribution of microcrystalline domain sizes upon the line-shapes. This led us to numerical heuristics for background subtraction rather than the subtraction of fitted functional profiles. (It is possible that a profile-fitting approach might work with simpler hexagonal ice.)

At its simplest, the correction involves subtracting the difference between the median intensity (as a function of scattering angle) and the value that it would have in the absence of ice, estimated by linearly interpolating through each ice ring, as illustrated in Fig. 1(*b*). This relies on an accurate value for the beam position. *DeIce* calculates this as the center of an ice ring. In an iterative refinement, the peak intensities are located along radii emanating from the current beam position and a new position is calculated from the center of mass of many such peak determinations. The interference of Bragg reflections in this process is reduced by progressively limiting the peak search to the neighborhood of the expected ice ring and rejecting maxima that greatly exceed the median ring intensity.

With a precise beam location, median intensities are calculated at each radius at pixel resolution. Median values are preferred over mean values because they allow an estimation of the underlying background without being skewed by Bragg peaks or unobserved pixels, providing that these constitute a minor fraction of the detector area. At user-input ice-ring limits, *DeIce* calculates the median intensity at the end points of a linear baseline. For arginine kinase the following ice rings were used: 4.2–3.4, 2.33–2.16 and 1.98–1.87 Å, which were generous limits that included the tails of the ice rings and constituted ∼30% of reciprocal space to 1.75 Å resolution. Within each ice ring, a one-dimensional numerical correction function is calculated at pixel resolution as the difference between the radial median intensity and the sloping baseline across the ice ring. The diffraction image is then processed and for every pixel within an ice ring the correction appropriate for its scattering angle is obtained from the array by linear interpolation and subtracted from the intensity. Unmeasured (zero) and overflow intensities are not corrected. The quality of data were assessed following integration

Table 1
Scaling statistics for data with different ice-ring treatments.

The *HKL*-2000 package (Otwinowski & Minor, 2001) was used in all cases with the same near-default parameters. The left block entitled 'Using background rejection' relied on *DENZO*'s rejection of measurements with uneven background, without pre-processing of the images. The center block pre-processed the images to exclude pixels within the resolution ranges of the ice rings. The right block used *DeIce* pre-processing to subtract the ice diffraction. Resolution shells containing the ice rings, in which there is greatest impact, are shown in bold.

	Using background rejection					Processed with ice rings excluded					Processed with ice-ring subtraction				
Resolution (Å)	$\langle I \rangle$	$\langle I/\sigma(I) \rangle$	χ^2	$R_{\rm merge}$	Completeness (%)	$\langle I \rangle$	$\langle I/\sigma(I) \rangle$	χ^2	$R_{\rm merge}$	Completeness (%)	$\langle I \rangle$	$\langle I/\sigma(I) \rangle$	χ^2	$R_{\rm merge}$	Completeness (%)
∞–4.27	34036	22.9	0.94	0.032	85	33769	22.9	0.94	0.032	85	34541	22.9	0.92	0.032	85
4.27–3.39	**33383**	**19.1**	**0.99**	**0.036**	**57**	**33987**	**19.3**	**0.94**	**0.035**	**33**	**35182**	**19.2**	**1.19**	**0.037**	**90**
3.39–2.96	18028	16.9	1.04	0.046	94	17872	16.9	1.03	0.046	94	18058	16.9	1.02	0.046	94
2.96–2.69	9364	13.1	0.95	0.057	95	9307	13.1	0.95	0.057	95	9367	13.1	0.96	0.057	95
2.69–2.50	6431	10.6	0.94	0.070	96	6407	10.6	0.94	0.070	96	6409	10.6	0.95	0.070	96
2.50–2.35	4996	9.3	1.01	0.084	96	4996	9.3	1.01	0.084	96	4960	9.4	1.00	0.084	96
2.35–2.23	**3814**	**5.2**	**0.88**	**0.100**	**46**	**3891**	**5.3**	**0.90**	**0.106**	**40**	**4928**	**5.3**	**1.97**	**0.164**	**89**
2.23–2.14	**3088**	**4.5**	**0.85**	**0.121**	**71**	**3347**	**4.8**	**0.83**	**0.121**	**60**	**3613**	**5.1**	**1.13**	**0.128**	**96**
2.14–2.05	2831	6.4	1.16	0.144	97	2838	6.4	1.16	0.144	97	2796	6.4	1.17	0.146	97
2.05–1.98	2216	5.5	1.19	0.172	97	2201	5.4	1.18	0.173	97	2217	5.5	1.20	0.172	97
1.98–1.92	**870**	**2.4**	**1.22**	**0.312**	**63**	**1093**	**2.9**	**1.24**	**0.283**	**43**	**1986**	**4.6**	**1.96**	**0.261**	**95**
1.92–1.87	**318**	**0.7**	**1.04**	**0.666**	**38**	**831**	**1.5**	**0.94**	**0.371**	**20**	**1881**	**2.3**	**1.46**	**0.348**	**94**
1.87–1.82	713	2.6	1.23	0.402	98	709	2.6	1.25	0.406	98	736	2.7	1.26	0.391	98
1.82–1.77	662	2.5	1.24	0.431	98	663	2.5	1.24	0.430	98	658	2.5	1.25	0.434	98
1.77–1.73	560	2.1	1.27	0.504	98	558	2.1	1.28	0.510	98	547	2.1	1.28	0.516	98
∞–1.73	7891	12.3	1.07	0.064	82	7671	12.2	1.08	0.065	77	8244	11.7	1.23	0.069	95

Table 2
Summary scaling statistics for three data sets, comparing the conventional approach of background rejection with the new approach of ice-ring subtraction.

The first two data sets were collected from *C*222 crystals of human acidic fibroblast growth factor mutants (Lee *et al.*, 2008). 'FGF-d' was used for the structure determination of mutant L26N/D28A (PDB code 3ba7). 'FGF-e' was an ice-affected data set for mutant K112N/N114A that was replaced prior to structure determination (PDB code 3bag). 'Sxa63' is a data set from a *P*222 crystal of the *Pyrococcus furiosus* box C/D ribonucleoprotein particle that was also replaced prior to structure determination (Li *et al.*, submitted work). Resolution shells affected by ice rings where the new procedure improves the completeness are shown in bold.

FGF-d					FGF-e					Sxa63				
	Background rejection		Ice-ring subtraction			Background rejection		Ice-ring subtraction			Background rejection		Ice-ring subtraction	
Resolution (Å)	$R_{\rm merge}$	Completeness (%)	$R_{\rm merge}$	Completeness (%)	Resolution (Å)	$R_{\rm merge}$	Completeness (%)	$R_{\rm merge}$	Completeness (%)	Resolution (Å)	$R_{\rm merge}$	Completeness (%)	$R_{\rm merge}$	Completeness (%)
∞–4.00	0.049	94	0.050	95	∞–5.12	0.048	100	0.047	100	∞–6.79	0.070	100	0.076	100
4.00–3.18	**0.075**	**70**	**0.071**	**99**	5.12–4.08	0.062	100	0.062	100	6.79–5.42	0.088	100	0.088	100
3.18–2.78	0.083	96	0.083	97	**4.08–3.57**	**0.117**	**37**	**0.147**	**98**	5.42–4.75	0.097	100	0.099	100
2.78–2.52	0.091	87	0.091	87	**3.57–3.24**	**0.130**	**65**	**0.125**	**100**	4.75–4.32	0.102	100	0.102	100
2.52–2.34	0.096	76	0.097	76	3.24–3.01	0.143	99	0.120	98	4.32–4.01	0.111	100	0.116	100
2.34–2.20	**0.128**	**51**	**0.142**	**76**	3.01–2.83	0.171	95	0.135	95	**4.01–3.78**	**0.170**	**72**	**0.165**	**100**
2.20–2.09	0.120	74	0.113	76	2.83–2.69	0.201	92	0.153	92	**3.78–3.59**	**0.404**	**47**	**0.346**	**100**
2.09–2.00	0.159	77	0.159	77	2.69–2.57	0.277	92	0.214	91	3.59–3.43	0.233	100	0.229	100
2.00–1.93	0.190	77	0.174	78	2.57–2.48	0.274	91	0.205	90	3.43–3.30	0.243	100	0.258	100
1.93–1.86	0.204	69	0.226	78	2.48–2.39	0.325	89	0.233	89	3.30–3.19	0.273	100	0.327	100
1.86–1.80	0.244	77	0.236	77	2.39–2.32	0.418	89	0.244	88	3.19–3.09	0.339	100	0.398	100
1.80–1.75	0.280	77	0.283	77	**2.32–2.25**	**0.595**	**33**	**0.645**	**87**	3.09–3.00	0.387	100	0.439	99
∞–1.75	0.081	77	0.081	83	∞–2.25	0.124	82	0.122	94	∞–3.00	0.119	93	0.125	100

and scaling using *DENZO* and *SCALEPACK* from the *HKL*-2000 suite (Otwinowski & Minor, 2001).

3. Results and discussion

Visual inspection shows the ice rings to have been successfully removed (Fig. 1*a*). The corrected image is largely, but not completely, free from artifacts. Near the ice rings, particularly at 2.25 and 1.92 Å, arcs of lighter or darker intensity can be discerned in the horizontal and vertical directions, respectively. Similar but more prominent artifacts had been seen near all three ice rings prior to refinement of the direct-beam location. Refinement to the center of the inner (3.7 Å) ice ring improved the position by two pixels (0.16 mm) relative to that obtained during integration by *DENZO*, achieving a standard error of 0.15 pixels (12 µm) as determined by refinements against different images. As the ice diffraction intensity changes sharply (Fig. 1), it is important to determine the background correction from pixels at the same radial distance, so sensitivity to the exact beam location is not unexpected.

The remaining variation in background intensity near the 2.25 and 1.92 Å ice rings is likely to result from a failure to account for detector tilt (with reference to the beam) and for uncorrected geometrical distortions in the raw image. The latter could come, for example, from optical tapers in CCD detectors (Otwinowsi & Minor, 2001) and result in imperfect correlation between pixel coordinates in the raw image and scattering angle. (Errors in our ice-ring subtraction could also result from directionally dependent variation in the background, although this is not evident in our images.) There are ways that tilt and distortions could be better handled, especially if ice-ring subtraction were to be fully integrated into data-processing packages where detector corrections are already applied, but this was a task well beyond the current scope. Even with non-optimal handling of these effects, the radial curvature of the image intensity is decreased tenfold by ice-ring subtraction so that most of the Bragg reflections can now be integrated.

The quality of the integrated and scaled data set following our ice-ring correction is compared in Table 1 with two methods of excluding the worst-afflicted reflections. Recent versions of programs such as

Table 3
Refinement of the atomic model of arginine kinase against the de-iced data set.

Resolution shells from which ice rings have been subtracted are shown in bold.

Resolution (Å)	R_{work}	R_{free}
26.62–4.06	0.158	0.193
4.06–3.22	**0.157**	**0.222**
3.22–2.82	0.178	0.229
2.82–2.56	0.173	0.238
2.56–2.38	0.169	0.200
2.38–2.24	**0.212**	**0.310**
2.24–2.12	0.187	0.240
2.12–2.03	0.191	0.238
2.03–1.95	0.209	0.257
1.95–1.89	**0.318**	**0.365**
1.89–1.83	0.243	0.286
1.83–1.77	0.249	0.315
1.77–1.73	0.257	0.291
Overall	0.190	0.242

MOSFLM (Leslie, 2006) and *XDS* (Kabsch, 2010) offer options to exclude pixels within user-defined resolution ranges. A similar effect was obtained using an option of our new pre-processing program, *DeIce*, by setting the intensities of these pixels to zero. This approach yields an R_{merge} value of 0.065 to the 2σ limit at 1.73 Å, but at 77% completeness it represents a further loss of 17% of the data (middle columns of Table 1), even though the exclusion ranges were trimmed to the tails of the ice rings (3.98–3.52, 2.29–2.21 and 1.95–1.88 Å). *DENZO* offers a different approach of excluding reflections whose background varies excessively (Otwinowski & Minor, 2001). The default parameters did a good job of rejecting the reflections closest to the peaks of the ice diffraction rings. Fewer data were lost (13% more than the new program; left-hand columns of Table 1). Data near the 1.92 Å ring have systematically lowered intensity, suggesting that the ice ring is elevating the background estimation for some reflections, but overall the data are of similar quality (R_{merge} = 0.064). For the data where the ice diffraction was subtracted (right-hand columns in Table 1), an R_{merge} that was only marginally worse (0.069) was achieved while retaining essentially complete data. The pre-processing correction is not perfect. There is a modest increase in χ^2 near the ice rings and ∼5% of reflections in the affected shells are still rejected owing to varying background. For the scalings compared in Table 1, identical error parameters were used that varied smoothly over the entire resolution range and were not tailored to ice-ring shells. Furthermore, identical rejection criteria were used. Thus, the higher χ^2 values after de-icing result from the inclusion of reflections in the statistics whose background estimates remain worse than average but are now good enough to avoid the rejection that occurs in conventional processing. Overall, subtraction of the ice diffraction appears to be the most appropriate strategy, with a substantial improvement in the completeness of the data set and little impact upon the quality.

Similar improvements have been obtained with several crystals diffracting to various resolution limits (Table 2). In all cases, our de-icing procedure was an improvement upon background-based rejection of affected reflections and substantially increased the completeness of the processed diffraction data. This can be achieved without significant loss of quality, as measured by the overall R_{merge}, even though the reflections from ice-affected shells are still not integrated quite as accurately as those from other parts of the diffraction pattern.

Another measure of improvement is cross-validation following atomic refinement, which has been repeated for the re-processed arginine kinase data set. In two of the three shells containing ice rings R_{free} is elevated (Table 3), but not enormously. It is best if ice rings can be avoided, but if not, the data salvaged by de-icing can contribute productively to refinement. The prior deposition of this structure (PDB code 1m80) had been refined in 2002 against the same data without de-icing, truncated at 2.35 Å resolution and yielded R_{free} = 0.237. Following a new refinement against the de-iced 1.73 Å data set, the structure (PDB code 3m10) has R_{free} = 0.221 when statistics are calculated to the prior 2.35 Å limit. The reduction in R_{free} of 0.02 may underrepresent the real improvement, because the 2002 R_{free} was lowered by the absence of high-resolution reflections near the ice rings. These reflections are now included, but with errors that are above average. If the ice-affected data were used for refinement but excluded from the statistics, as in the prior refinement, the model would yield R_{free} = 0.21 or 0.22 for 2.35 and 1.73 Å resolution limits, respectively. These 'adjusted' statistics suggest that the real model improvement from inclusion of the ice-affected data corresponds to a ΔR_{free} of 0.03.

The presence of ice rings usually encourages a more extensive search for a more suitable cryoprotection protocol. Occasionally crystals are intransigent, with all perturbations detracting from diffraction quality. Providing that bulk ice does not destroy the crystal lattice, it may be that the best diffraction is only obtained with some ice diffraction overlaid. Here, the feasibility of subtracting the ice component of the diffraction is demonstrated. Should there be sufficient call for this approach, then it would be worthwhile to modify existing integration packages with algorithms for background estimation that can tolerate large variation in the radial component, or to subtract the ice component in a pre-processing step as in this work but with the added benefit of corrections for detector tilt and distortions.

Arginine kinase crystals were prepared by Shawn A. Clark, who also helped in data collection. Arginine kinase data were collected on the BioCARS beamline 14-BM-C at the Advanced Photon Source, while the other data sets were collected on the SER-CAT beamline 22 at the Advanced Photon Source and were generously made available by Mike Blaber and Hong Li, the latter prior to publication. Source code and documentation is available at http://xtal.ohsu.edu/. This work was supported by the National Institutes of Health grant R01 GM077643 to MSC.

References

Berejnov, V., Husseini, N. S., Alsaied, O. A. & Thorne, R. E. (2006). *J. Appl. Cryst.* **39**, 244–251.
Blackman, M. & Lisgarten, N. D. (1957). *Proc. R. Soc. Lond. A*, **239**, 93–107.
Dowell, L. G. & Rinfret, A. P. (1960). *Nature (London)*, **188**, 1144–1148.
Ellington, W. R. (2001). *Annu. Rev. Physiol.* **63**, 289–325.
Garman, E. F. & Owen, R. L. (2006). *Acta Cryst.* D**62**, 32–47.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kim, C. U., Chen, Y.-F., Tate, M. W. & Gruner, S. M. (2008). *J. Appl. Cryst.* **41**, 1–7.
Lee, J., Dubey, V. K., Longo, L. M. & Blaber, M. (2008). *J. Mol. Biol.* **377**, 1251–1264.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
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.
Rodgers, D. W. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 202–208. Dordrecht: Kluwer Academic Publishers.
Yousef, M. S., Clark, S. A., Pruett, P. K., Somasundaram, T., Ellington, W. R. & Chapman, M. S. (2003). *Protein Sci.* **12**, 103–111.