

A survey of global radiation damage to 15 different protein crystal types at room temperature: a new decay model

Ricardo Miguel Ferraz Leal,^{a,b} Gleb Bourenkov,^{c*} Silvia Russi^a and Alexander N. Popov^{a*}

^aESRF, BP-220, 38043 Grenoble, France, ^bILL, BP 156, 38042 Grenoble, France, and ^cEMBL Hamburg Outstation, c/o DESY, Notkestrasse 85b, Hamburg 22607, Germany.
E-mail: gleb@embl-hamburg.de, apopov@esrf.fr

The radiation damage rates to crystals of 15 model macromolecular structures were studied using an automated radiation sensitivity characterization procedure. The diffracted intensity variation with dose is described by a two-parameter model. This model includes a strong resolution-independent decay specific to room-temperature measurements along with a linear increase in overall Debye–Waller factors. An equivalent representation of sensitivity *via* a single parameter, normalized half-dose, is introduced. This parameter varies by an order of magnitude between the different structures studied. The data show a correlation of crystal radiation sensitivity with crystal solvent content but no dose-rate dependency was detected in the range 0.05–300 kGy s⁻¹. The results of the crystal characterization are suitable for either optimal planning of room-temperature data collection or *in situ* crystallization plate screening experiments.

Keywords: radiation damage; room temperature; X-ray diffraction; data collection.

1. Introduction

The vast majority of macromolecular crystallographic data are collected at cryotemperatures. This is because the rate of overall decay in the diffraction intensity signal, also known as ‘global radiation damage’ (Holton, 2009), slows down by about two orders of magnitude at 100 K compared with at room temperature (RT) (Nave & Garman, 2005). Significant progress has been made in providing a quantitative phenomenological description of global radiation damage at cryotemperatures (reviewed by Holton, 2009; Garman, 2010). All macromolecular crystals that have been systematically examined to date show very similar radiation sensitivities. Therefore, the resolution-dependent intensity decay can be taken into account in the optimal planning of a diffraction experiment (Bourenkov & Popov, 2010). The remaining complications during cryogenic data collection arise mostly from the specific experimental conditions, in particular from a systematic mismatch of the beam size with crystal size at third-generation undulator beamlines (Krojer & von Delft, 2011). In contrast, the specific radiation-induced changes under cryo-conditions are often the source of severe difficulties in interpretation of structure and function (*e.g.* Dubnovitsky *et al.*, 2005; Borshchevskiy *et al.*, 2011).

Although infrequently used, RT data collection remains necessary for a number of studies. Cryogenic techniques

introduce artifacts, either directly *via* temperature effects or indirectly because of the cryoprotectants and temperature-induced pH changes (Juers & Matthews, 2001; Halle, 2004; Dunlop *et al.*, 2005). In recent years, a renaissance of interest in RT experiments has occurred in the context of *in situ* diffraction methods where the entire crystallization plate is mounted on the beamline. A number of developments aiming at automation pursue the idea of *in situ* data collection in order to bypass the crystal handling step that is notoriously difficult to automate. Maire *et al.* (2011), Bingel-Erlenmeyer *et al.* (2011), Hargreaves (2012) and Axford *et al.* (2012) have all used the ‘in-plate’ screening and/or data collection approach. The concept of microfluidic-based crystallization also relies on RT ‘on-chip’ data collection (Hansen *et al.*, 2006; Dhoubib *et al.*, 2009; Li & Ismagilov, 2010; Kisselman *et al.*, 2011). Recently, Axford *et al.* (2012) reported a series of case studies strongly supporting the *in situ* RT approach. Examples include the successful structure solution of bovine enterovirus-2, where successful cryoprotection of the crystals was not at all possible.

For the optimal design of the experiments described above, whether they involve screening or data collection, a prior knowledge of sample dose tolerance at RT is absolutely essential. Recent systematic studies (Southworth-Davies *et al.*, 2007; Barker *et al.*, 2009; Warkentin & Thorne, 2010; Warkentin *et al.*, 2011, 2012; Kmetko *et al.*, 2011; Rajendran *et al.*, 2011; Owen *et al.*, 2012) provided the global radiation

damage data for a small number of structures under conditions of varying temperatures, chemical additives, dose rates and detector read-out frequencies. However, despite a long history of RT data collection from macromolecular crystals, practically no earlier data exist in a form suitable for systematic analysis and for making predictions on likely crystal dose tolerances.

In this work we investigate the global radiation damage rates of crystals of 15 different model structures, at RT and under native crystallization conditions. The principle aim of the experiments was to obtain a parametric description of diffraction intensity decay behavior suitable for data collection strategy optimization taking radiation damage into account using *BEST* (Bourenkov & Popov, 2010). The approach employed involves the collection of multiple partial data sets as a function of dose, and the use of a data analysis method developed specifically to verify the radiation damage model, to extract the model parameters and to minimize systematic errors. The reproducibility of the results was tested systematically. The experiments made use of the procedure for automatic characterization of radiation sensitivities at cryotemperatures developed earlier (Leal *et al.*, 2011). The further development and the modifications of the method that were required to account for the features of global radiation damage specific to RT are described.

2. Sample preparation and data collection

A summary of the crystallographic parameters of the 15 crystal systems studied in this work is presented in Table 1. Different crystal forms of lysozyme, trypsin and insulin, as well as thaumatin, thermolysin and FAE, are standard test crystal systems. Crystal samples of bR, 6HLNO, LACV, TIM and TvNiR were kindly donated by the authors of the respective publications indicated in Table 1.

For data collection at RT, the open-flow humidity control apparatus HC1 was employed. The samples were prepared for measurements in the HC1 as described by Sanchez-Weatherby *et al.* (2009). The crystals were mounted on Mitegen micro-meshes and excess liquid was removed with filter paper in order to prevent the crystals from moving during the data collection.

The measurements were carried out on ESRF beamline ID23-1 (Nurizzo *et al.*, 2006) using an ADSC Q315 detector. The X-ray beam energy was kept constant at 12.75 keV for all measurements. The nominal beam size at the sample position was 35 μm vertically and 45 μm horizontally (full width at half-maxima). In all cases crystals were selected such that all of the beam cross section was intercepted by the crystal.

The absorbed doses and rates were estimated with the program *RADDOSE* (Murray *et al.*, 2004; Paithankar *et al.*, 2009) using the photon flux values estimated by measurements with standard calibrated intensity monitors and the chemical composition of the protein, the ligands and the crystallization solution according to the reference in Table 1.

The standard procedure for radiation sensitivity measurements (Leal *et al.*, 2011), as implemented in *MxCuBE/EDNA* (Gabadinho *et al.*, 2010; Incardona *et al.*, 2009), was used. In this procedure the collection of 11 successive wedges of data interleaved by X-ray exposures to 'burn' the crystal are performed in a narrow rotation range (usually 3°), thus excluding variation in the exposed crystal volume. The procedure involves optimization of the intensity decay measurements, and necessarily requires preliminary knowledge or an assumption of the radiation sensitivity of the sample under consideration. Assuming that both the absorbed dose rate and the crystal sensitivity are known approximately, the burning and data collection protocol is generated automatically by the program *BEST* (Bourenkov & Popov, 2010) on the basis of the data obtained from the initial sample characterization step. The protocol defines a complete set of the parameters required: exposure time, attenuator transmission, total rotation range, rotation range per frame, the resolution limit (d_{min}) for data collections, and the exposure time for irradiation. A detailed description of the methods used in the protocol generation is given by Leal *et al.* (2011). Using the model assumptions, the diffraction resolution limit and the dose for the burning cycles are selected in such a way that significant changes in the scattered intensities are induced, and the intensity measurements remain statistically significant up to the last cycle of data collection. At RT, the exposures between the burning cycles induced substantial radiation damage and often no burning cycles were required. In the method described by Leal *et al.* (2011) for cryotemperatures, the maximum dose per exposure cycle was 0.1 MGy. For RT data collection the maximum dose per exposure cycle was chosen such that the expected increase in the *B*-factor did not exceed 1 \AA^2 . This consideration, combined with the standard *BEST* calculation as described by Popov & Bourenkov (2003) and Bourenkov & Popov (2010), gives rise to a consistent choice of the resolution limit, exposure time and rotation width per frame.

The data collection protocols were initially generated under the assumption of 70-fold higher radiation sensitivity at RT as compared with cryogenic conditions; the factor was chosen according to the previous studies at RT (Nave & Garman, 2005; Kmetko *et al.*, 2011). This corresponds to a total exposure time of approximately 1–2 s with the unattenuated X-ray beam, depending on the crystal absorbance and incident flux (varying with the ESRF filling mode). The shortest exposure time permitted on the ID23-1 diffractometer is 0.1 s per frame; thus for most of the measurements attenuation of the beam intensity was necessary.

Preliminary on-line data analysis indicated a strong variation in the radiation sensitivity between the samples studied. For sensitive crystals, the data processing typically failed on the last wedges whereas, for less sensitive samples, significant reduction of the intensity (about 50%) was not reached. For all systems studied (excluding TIM), after adjusting the sensitivity parameters according to the initial estimates of the decay rates, the measurements were repeated several times, either on the same crystal, after translating unexposed parts of

radiation damage

Table 1

Crystallographic parameters, experimental conditions and radiation damage parameter statistics.

Ref = Reference. No. = Number of crystals or crystal centerings. Res = Resolution. Ave = average; SD = standard deviation. Dose rate units: kGy s⁻¹. β units: Å² MGy⁻¹. γ units: MGy⁻¹. $D_{1/2}^N$ units: MGy.

Acronym	Protein	Source	Ref†	Unit cell (Å, °)	Space group	Solvent content (%)	No.	Res (Å)	Dose rate		β		γ		$D_{1/2}^N$	
									Min	Max	Ave	SD	Ave	SD	Ave	SD
LYZM	Lysozyme	Hen egg-white	^a	$a = 28$ $b = 63$ $c = 60$ $\beta = 90.2$	$P2_1$	34.0	7	1.9	0.4	225	15	0.5	1.0	0.2	0.56	0.07
LYZT	Lysozyme	Hen egg-white	^b	$a = 79$ $b = 79$ $c = 38$ $\beta = 90.2$	$P4_32_12$	41.4	8	1.9	15	294	19	4	2.0	0.3	0.32	0.03
INSR	Insulin	Bovine pancreas	^c	$a = 83$ $b = 83$ $c = 109$ $\beta = 90.2$	$H3$	38.3	4	2.0	10	13	24	3	1.0	0.3	0.47	0.06
INSC	Insulin	Bovine pancreas	^d	$a = 79$ $b = 79$ $c = 79$ $\beta = 90.2$	$I2_13$	67.0	11	2.0	0.4	87	32	3	3.6	0.3	0.18	0.01
BPTTR	Trypsin	Bovine pancreas	^e	$a = 559$ $b = 559$ $c = 109$ $\beta = 90.2$	$P3_121$	36.5	8	2.0	8	40	13	1	1.1	0.3	0.58	0.10
BPTOH	Trypsin	Bovine pancreas	^f	$a = 55$ $b = 59$ $c = 68$ $\beta = 90.2$	$P2_12_12_1$	45.0	3	1.7	9.6	44	14	3	1.7	0.2	0.40	0.04
BPTOL	Trypsin	Bovine pancreas	^g	$a = 63$ $b = 64$ $c = 69$ $\beta = 90.2$	$P2_12_12_1$	56.9	10	2.2	11	14.3	20	4	2.8	0.4	0.24	0.03
bR	Bacteriorhodopsin	<i>Halobacterium salinarium</i>	^h	$a = 62$ $b = 62$ $c = 110$ $\beta = 90.2$	$P6_3$	51.1	3	3.0	1.1	124	29	6	1.0	0.06	0.42	0.05
TIM	Triosephosphate isomerase	<i>Leishmania mexicana</i>	ⁱ	$a = 99$ $b = 53$ $c = 61$ $\beta = 118$	$C2$	52.1	1	2.5	111		33	–	1.6	–	0.3	–
THER	Thermolysin	<i>Bacillus thermoproteolyticus</i>	^j	$a = 95$ $b = 95$ $c = 143$ $\beta = 118$	$P6_122$	53.8	4	2.8	17.4	20.4	28	5	0.9	0.2	0.44	0.02
THAU	Thaumatococin	<i>Thaumatococcus daniellii</i>	^k	$a = 59$ $b = 59$ $c = 152$ $\beta = 118$	$P4_12_11$	56.8	4	2.5	11	16	20	2	1.3	0.2	0.43	0.03
FAE	SeMet-FAE‡	<i>Clostridium thermocellum</i>	^l	$a = 66$ $b = 110$ $c = 114$ $\beta = 118$	$P2_12_12_1$	63.3	5	2.5	54	79	34	5	1.2	0.2	0.36	0.06
6hIno	6-hydroxy-L-nicotinic oxidase	<i>Arthrobacter nicotinovorans</i>	^m	$a = 167$ $b = 167$ $c = 167$ $\beta = 118$	$P432$	69.7	3	3.5	26.5	28	215	30	3.2	0.3	0.07	0.01
LACV	L-protein polymerase N-terminal domain	La Crosse orthobunyavirus	ⁿ	$a = 125$ $b = 125$ $c = 167$ $\beta = 118$	$P6_122$	69.9	5	3.9	27	113	142	20	5.7	1.0	0.08	0.003
TvNiR	Cytochrome c nitrite reductase	<i>Thioalkalivibrio nitratireducens</i>	^o	$a = 197$ $b = 197$ $c = 197$ $\beta = 118$	$P2_13$	77.7	3	2.6	30		11	1	0.6	0.1	0.91	0.09

† ^aHogle *et al.* (1981). ^bBlake *et al.* (1965). ^cSmith *et al.* (2005). ^dNanao *et al.* (2005). ^eBode & Huber (1978). ^fMarquart *et al.* (1983). ^gBartunik *et al.* (1989). ^hBorshchevskiy *et al.* (2011). ⁱAlahuhta & Wierenga (2010). ^jMueller-Dieckmann *et al.* (2007). ^kCharron *et al.* (2002). ^lPrates *et al.* (2001). ^mKachalova *et al.* (2010). ⁿReguera *et al.* (2010). ^oBoyko *et al.* (2006). ‡ Selenomethionine labeled feruloyl esterase module of xylanase 10B.

the crystal into the beam, or by using different crystals. Finally all the data, including those obtained during the initial cycles, were included in the analysis.

3. Data analysis and results

3.1. Scaling

The data analysis performed in these experiments aimed for a parametric statistical description of the scattering power

of the crystal as a function of absorbed dose. The scaling procedure differs conceptually from the standard scaling that minimizes intensity differences between (predominantly strong) equivalent reflections, as is commonly carried out during data reduction (Kabsch, 2010; Evans, 2011).

Here the scattering power is described by an expectation value of the reflection intensity $\widehat{J}(\mathbf{h})$ at a reciprocal lattice vector \mathbf{h} . It can be expressed *via* a product of an empirical curve defining the radial shape $\widehat{J}(h = |\mathbf{h}|)$ of the function, the

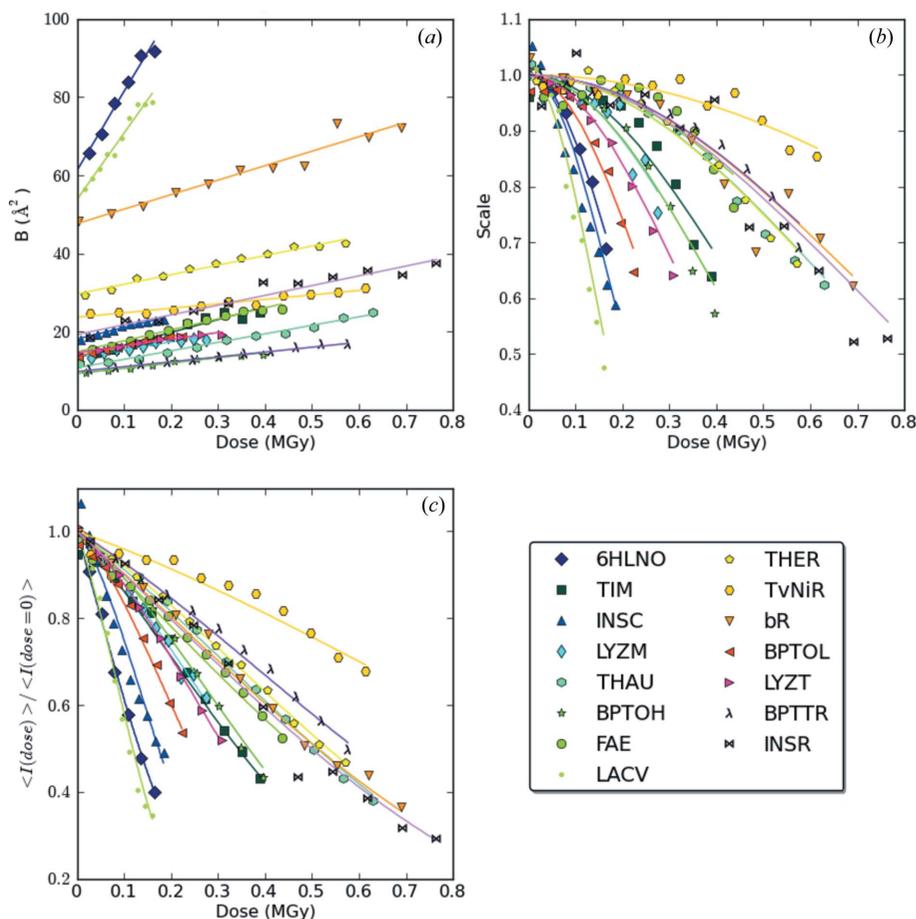


Figure 1
Variation in the scaling parameters and total scattering intensity with dose. For each of the systems studied, one measurement is shown. (a) *B*-factors: the solid line represents an approximation according to equation (2) using the best fit of B_0 and β . (b) *scale* factors divided by the constant [equation (3)]: the solid line represents $\exp(-\gamma^2 D^2)$ using the best fit value of γ . (c) Total scattered intensity, estimated by intensity summation (dots) and calculated according to equation (4) (solid lines), using the best fit B_0 , β and γ as in (a) and (b).

scaling factor (*scale*) and the overall Debye–Waller factor (Popov & Bourenkov, 2003),

$$\hat{J}(\mathbf{h}) = \hat{J}(h) \text{scale} \exp(-Bh^2/2). \quad (1)$$

The function $\hat{J}(h)$ is defined by the interatomic distance distribution in macromolecules. The representation described above is widely used in current crystallographic methodology (e.g. Morris *et al.*, 2004; Adams *et al.*, 2010). Radiation damage does not lead to measurable changes in $\hat{J}(h)$. An empirical $\hat{J}(h)$ curve was tabulated as described (Popov & Bourenkov, 2003) and is on an arbitrarily chosen fixed scale. In the work presented here, use is made of the isotropic *B* factor approximation. Thus, equation (1) applies specifically to a narrow wedge of data, as in this experiment. With broader angular ranges, both the anisotropic *B* tensors and the variation in the *scale* parameter with crystal orientation (owing to the varying irradiated crystal volume) would have to be considered.

The parameters *scale* and *B* were estimated by maximizing the likelihood function [equation (18) of Popov & Bourenkov, 2003], as implemented in *BEST*. The function is derived

under the assumption that the intensities obey the acentric Wilson distribution (Wilson, 1949). Lorentz-polarization-corrected un-scaled integrated intensities output by *XDS* (Kabsch, 2010) were used as the input to *BEST*. Thus, for each of the data wedges, the *scale* and *B* values were determined independently of the other wedges, and no scaling between wedges by *XDS* was involved. Applying the scaling procedure to the series of wedges measured on one crystal (or one crystal centering) gave the scaling parameters as a function of dose, $\text{scale}(D)$ and $B(D)$.

Compared with the standard scaling method, this approach does not rely on an assumption of identity in scaled intensities for sequential observations of the same \mathbf{h} . The latter assumption holds at low doses, when standard scaling may provide more accurate parameter estimates, but fails at progressively higher doses. The problems with the interpretability of the results of the standard *B*-factor scaling output in terms of radiation damage metrics have been discussed previously (Southworth-Davies *et al.*, 2007; Barker *et al.*, 2009; De la Mora *et al.*, 2011).

The dependence of the scaling results on the choice of the integration software (*XDS*, *HKL2000* or *MOSFLM*) has been analyzed previously in the course of testing the *BEST* software

(data not shown). No systematic issues were identified.

3.2. Dose dependence of the *B*-factor

The dependence of the *B*-factor on the absorbed dose is presented in Fig. 1(a). For Figs. 1(a), 1(b) and 1(c), for each of the systems studied, one measurement corresponding to the highest total absorbed dose was selected. The reproducibility of the experiments is shown in the supplementary figures¹ where all the measurements are included.

As described in previous studies (Kmetko *et al.*, 2006; Bourenkov & Popov, 2010; Borek *et al.*, 2010; Warkentin & Thorne, 2010) carried out at both cryo or room temperatures, the observed distributions of the *B*-factor *versus* dose can be fitted by a linear function,

$$B(D) = B_0 + \beta D, \quad (2)$$

¹ Supplementary data for this paper are available from the IUCr electronic archives (Reference: XH5035). Services for accessing these data are described at the back of the journal.

and the parameter β can be used as one of the radiation sensitivity characteristics. The average β values observed for each structure and their standard deviations are listed in Table 1 (see also the supplementary material). Whereas at cryogenic temperatures much lower values of β , varying within a narrow range of 0.5–1.2 Å² MGy⁻¹, were observed (Kmetko *et al.*, 2006; Leal *et al.*, 2011), at RT β values are higher by one to two orders of magnitude and show a large variation between structures.

3.3. Dose dependence of the scale factor

The dependence of the *scale* factor on the absorbed dose is shown in Fig. 1(b). Under cryoconditions, only a small and approximately linear reduction in the *scale* factors has been observed, accounting for at most 5–15% of the overall drop in the intensity (Leal *et al.*, 2011). At RT, the decrease in the *scale* factor was pronouncedly non-linear, larger in magnitude and varied strongly between different structures. Empirically, the function

$$\text{scale}(D) = \text{constant} \times \exp(-\gamma^2 D^2) \quad (3)$$

was found to describe the character of the distribution of the *scale* factor *versus* dose reasonably accurately, within the range of the *scale* factors sampled by our experiments. The constant and the parameter γ (Gy⁻¹) have been fitted for all data sets (supplementary figures and table); the statistics for γ per structure are given in Table 1. Equivalent analysis of the cryotemperature data from Leal *et al.* (2011) showed the variation of γ to be between 0.03 and 0.05 MGy⁻¹.

3.4. Dose dependence of the total scattered intensity

Several recent radiation damage studies (*e.g.* Kmetko *et al.*, 2011) have used β as the only metric of radiation damage. With negligibly small γ values at cryotemperatures, such an approach is justified. However, neglecting the strong variation in overall scale factor at RT leads to a systematic underestimation of radiation sensitivity. As discussed, a two-parameter model (β and γ) is necessary and sufficient to construct a predictive model of resolution-dependent intensity variation. However, the use of a two-parametric model as a comparative metric relating the radiation sensitivity of different structures, or one and the same structure under various conditions, is extremely inconvenient.

The $D_{1/2}$ parameter, the dose at which the total diffraction intensity reduces by a factor of two, was introduced as a radiation sensitivity metric by Garman and co-workers (Owen *et al.*, 2006) and used in a similar form in many subsequent studies (*e.g.* Barker *et al.*, 2009; Sanishvili *et al.*, 2011; Owen *et al.*, 2012).

The relationship between the dose-dependent scaling factors and the total scattered intensity may be derived in a straightforward way. Combining expressions (1), (2) and (3) and integrating over h , we obtain expression (4) for the total scattered intensity as a function of absorbed dose,

$$I_{\Sigma}(D) \propto \exp(-\gamma^2 D^2) \int_0^{1/d_{\min}} h^2 \widehat{J}(h) \exp[-\frac{1}{2}(B_0 + \beta D)h^2] dh. \quad (4)$$

The total recorded intensities, estimated by the direct summation of all integrated intensities (corrected by the Lorenz-polarization factor, within the resolution limit d_{\min}), as a function of dose are shown in Fig. 1(c) (see supplementary figures for all data sets). The solid lines represent the $I_{\Sigma}(D)/I_{\Sigma}(0)$ functions calculated using (4), the fitted values of γ , β and B_0 , and the respective values of d_{\min} for each experiment. The calculated curves match the observations rather well. Corresponding half-dose values $D_{1/2}$, satisfying the equation $2I_{\Sigma}(D_{1/2}) = I_{\Sigma}(0)$, are listed in the supplementary table for each measurement.

3.5. Normalized half-dose

Two examples of the integrand function in equation (4) are given in Fig. 2 for (arbitrarily chosen) values of $B_0 + \beta D = 10$ Å² and 20 Å². In all calculations presented here, the low-resolution integration limit was set at 1/12 Å, from the available $\widehat{J}(h)$ tabulation. This low-resolution truncation had no effect on the analysis.

The integration in (4) can be carried out either to the resolution limit of the data or until the integral converges. The first method corresponds to the regular procedure of determining $D_{1/2}$ by direct summation of measured intensities, the second provides the true estimate of the total scattered intensity independent of the resolution of the experiment. It was found that the high-resolution integration cut-off had a significant systematic influence on the estimated $D_{1/2}$ values. The $D_{1/2}$ obtained by either the direct intensity summation or by the integration (4) within the same resolution limits are systematically higher (hence, the radiation sensitivity is underestimated) when compared with the estimates of $D_{1/2}$ calculated by integrating (4) to convergence. For the THAU and FAE data, where the errors were largest, this amounts to a difference of about 30%.

Furthermore, from consideration of equation (4), Fig. 2 and the supplementary table, the systematic dependence of $D_{1/2}$ on B_0 is apparent. $D_{1/2}$ is systematically larger for more weakly diffracting crystals having large B_0 , simply due to the fact that

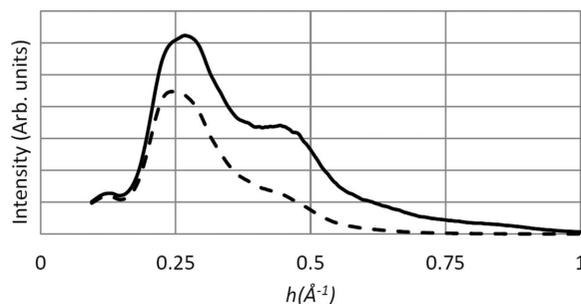


Figure 2
A generic empirical model of total diffraction intensity *versus* resolution, $h^2 \widehat{J}(h) \exp[-(1/2)Bh^2]$, for $B = 10$ Å² (solid line) and $B = 20$ Å² (dashed line).

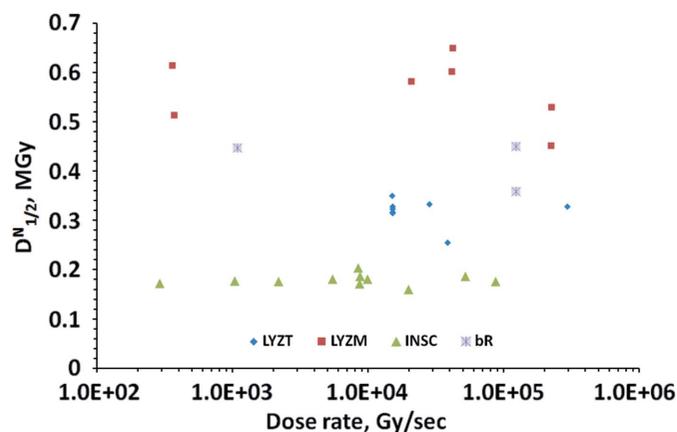


Figure 4
Normalized half-dose $D_{1/2}^N$ versus the dose rate for LYZT, LYZM, bR and INSC.

remarkably constant over the whole dose-rate range covered. The random noise is somewhat higher for LYZM and LYZT, where measurements were performed on different crystals.

4. Discussion

4.1. Application of the workflow

The reliable determination of the radiation sensitivity of macromolecular crystals is not a trivial experimental task. When assessing radiation damage *via* retrospective analysis of regular data sets collected in the process of structure solution, we are systematically confronted with a number of unaccountable critical details of the experiment (Krojer & von Delft, 2011). As demonstrated by Nowak *et al.* (2009), reproducibility issues are also evident in characterization experiments. We conclude that standard data collection protocols are not suitable for the measurement of radiation sensitivity. Therefore, an automated characterization procedure was developed and extensively tested under cryogenic conditions (Leal *et al.*, 2011). The procedure is currently implemented in an integrated workflow, transparently connecting experiment and data analysis at the ESRF beamlines (Brockhauser *et al.*, 2012). It has proven to be a convenient and reliable tool delivering reproducible data.

Here the workflow has been applied to a large set of model structures at RT. The use of the capillary-free mounting technique with the humidity control device HC1 (Sanchez-Weatherby *et al.*, 2009) simplified sample handling in these experiments. Furthermore, it excluded uncontrolled dehydration or crystal slippage.

4.2. An updated decay model

The main purpose of the radiation sensitivity measurements was to estimate the parameters of a model describing resolution-dependent intensity variation with dose, in a form suitable for optimization of data collection in *BEST*. The model used in *BEST* was previously constructed on the basis of empirical data collected at cryotemperature, where a linear increase in overall *B*-factors is the totally dominant factor.

4.2.1. The Debye–Waller factor model. All the RT experiments presented here (77 decay curves in total) showed a linear increase in *B*-factor with dose, essentially up to the point where the loss of the signal-to-noise of high-resolution reflections made accurate determination of *B*-factors impossible. From multiple measurements, it is estimated that the relative error in determining the coefficient β is about 15% in this procedure. The average value of $\beta = 20 \text{ \AA}^2 \text{ MGy}^{-1}$ for THAU is between those of Warkentin & Thorne (2010) who reported $\beta = 35 \text{ \AA}^2 \text{ MGy}^{-1}$, and Rajendran *et al.* (2011) who reported $\beta = 16 \text{ \AA}^2 \text{ MGy}^{-1}$. For INSC, $\beta = 32 \text{ \AA}^2 \text{ MGy}^{-1}$ agrees broadly with the $44 \text{ \AA}^2 \text{ MGy}^{-1}$ reported (Rajendran *et al.*, 2011).

4.2.2. The scale factor model. The experiments described here indicate that due to the presence of a strong decay component which is independent of the scattering angle, the linear *B*-factor model alone significantly underestimates the intensity variation in all resolution shells at RT. The noise in the *scale(D)* dependence is significantly higher than that in *B(D)*. This noise is ascribed to the variation in the three-dimensional spot profile parameters that are independently adjusted by *XDS* (Kabsch, 2010) for each processed wedge. Note that such adjustments would occur with any processing package; they have no negative effect on any of the standard crystallographic data set statistics. The accuracy of *scale(D)* is also a dominating factor defining the match of the $I_{\Sigma}(D)$ curves with the approximation in equation (4). Noticeable deviations of about 10% are observed specifically at low doses for several examples (*e.g.* INSC in the supplementary figures, panel *c*). We speculate that a sharp increase in the crystal mosaicity after the first exposures leads to a significant change in the integrated intensity values which are unaccounted for by the current analysis.

Nevertheless, at this noise level one can confidently see the concave character of the *scale(D)* dependency, with a linear decay rate $\partial \text{scale}(D)/\partial D$ increasing with dose. The analytical function (3) chosen to model this dependence represents a simplest uni-parametric function satisfying the second-order kinetic equation

$$\frac{1}{\text{scale}(D)} \frac{\partial \text{scale}(D)}{\partial D} = -\frac{1}{2} \gamma^2 D. \quad (5)$$

Equation (3) may only be considered an empirical parameterization useful to quantify the intensity decay within the range sampled by our experiments, typically $\text{scale}(D) > 0.6$. The nature of the variation in *scale(D)* and its concave shape is rather obscure. Within the framework of the standard crystallographic methodology, *scale* is proportional to the number of exposed unit cells and (3) could describe the formation of extended non-crystalline ‘void’ areas within the sample. The increase in the linear decay rate could then be connected to de-stabilization of the crystalline areas in the vicinity of the voids. In Hendrickson (1976) such a destabilization was proposed to be responsible for the time-dependent intensity decay component, but naturally not for the dose-dependent one. Thus, the model is at odds with the dose-rate independence of our experiments. So far it is clear only that the

observed changes in the scale factors are connected to defects of the crystal lattice. At this point neither the exact type(s) of the defects involved and their relationship to the observed integrated intensities, nor the defect kinetics with the dose, can be assigned. Answering this question would require complex experiments unrelated to crystallographic data collection.

4.2.3. Applications of the model. Despite its unclear physical meaning, the γ parameter was determined reproducibly in multiple measurements with an r.m.s.d. error of approximately 20%. The decay model used in the program *BEST* in versions 3.4.4 and higher was adopted to make use of the two-parameter (β , γ) model. The accuracy of determining β and γ , as may be achieved in a ‘sacrificial crystal’ experiment, is clearly sufficient for the optimization of data collection parameters: the signal-to-noise ratio in the data as a function of the absorbed dose varies only slowly in the vicinity of its attainable maximum (Bourenkov & Popov, 2010). At the same time, neither $D_{1/2}^N$ nor the $D_{1/2}$ parameter alone would be sufficient to define an optimal exposure time or dose in a particular ‘sacrificial crystal’ experiment (due to the strong resolution dependence in both the decay rate and in the signal-to-noise ratio). One should note that the β , γ model and the optimization methods mentioned above are only applicable at a relatively high resolution. The lowest resolution data set presented here was at 4.0 Å (LACV). Presumably, a practical limit of applicability would be at around 4–5 Å. Real-life examples of optimized RT data collections using the two-parameter model will be published elsewhere. The sensitivity calibration at RT may require two iterations, since the sensitivity may deviate strongly from initial expectations.

4.3. $D_{1/2}^N$ as a comparative metric

In order to represent the results of the sensitivity measurements of different types of crystals in a form suitable for comparison, the model based on two highly correlated parameters was numerically reduced to a single parameter $D_{1/2}^N$, which is equivalent to the half-dose parameter previously used in many studies. Appropriate normalization removes severe systematic dependence on the experimental conditions and the initial diffraction power of the crystal, both of which are inherent to $D_{1/2}$ determined by direct intensity summation. Note that to determine the half-dose our treatment uses the same number of fit parameters (two) as the direct summation method, but does not suffer from a systematically occurring inconsistency when $I_{\Sigma}(D)/I_{\Sigma}(0) \neq 1$ at $D = 0$. We propose the use of the normalized $D_{1/2}^N$ metric in radiation damage studies in general, or at least in those where decay dependence on the resolution is not an explicit subject of the experiment. The software for this analysis will be made available along with *BEST*.

For the LYST, INSC and THAU crystal systems, room-temperature radiation sensitivity data have been reported previously. Within the difference in the definitions and analysis, $D_{1/2}^N$ of LYST compare well with $D_{1/2}$ reported by Barker *et al.* (2009) (0.3 versus 0.2 MGy), INSC (0.18 versus 0.2 MGy) and THAU (0.4 versus 0.35 MGy) by Rajendran *et*

al. (2011), and THAU (0.4 versus 0.15 MGy) by Kmetko *et al.* (2011). Here, we use averaged reported values for comparison. From re-analysing the cryotemperature data reported by Leal *et al.* (2011), the observed normalized half-dose values were between 12 and 23 MGy for seven different crystal systems.

4.4. Correlation with solvent content

Joint analysis of the average normalized $D_{1/2}^N$ showed a remarkable increase in sensitivity with the crystal solvent content. To our knowledge, we present here only the second reliably documented systematic correlation of radiation sensitivity with a crystal property, the first being the dependence on absorbance described by Owen *et al.* (2006). The most obvious explanation of our data is that this dependence is due to the higher number of water-radiolysis radicals created per protein atom: $D_{1/2}^N \propto (1/\text{solvent content}) - 1$. The relationship holds (approximately) for different crystal forms of insulin and trypsin. Considerations of concentration alone would suggest smaller differences in sensitivities between the monoclinic and trigonal lysozyme. Apparently, the lower lattice energies of high-solvent crystals also play a role. Such an explanation would also be consistent with the large variation in sensitivities between unrelated structures with similar solvent content.

4.5. Dose rate dependence

Finally, our observation of the independence of the decay rates on the dose rate within the range between 0.5 and 300 kGy s⁻¹ agrees well with the results by Warkentin *et al.* (2012) and Owen *et al.* (2012) made at similar temperature, dose rates and time scales of the experiments. Recently, dark progression of radiation damage on the second scale at 240 K (Warkentin *et al.*, 2012) and on the millisecond time scale at RT, as reported by Owen *et al.* (2012), was observed. In such ‘outrunning’ conditions, reliable and reproducible empirical characterization of radiation sensitivity will require considerably more sophisticated experimental procedures and data analysis protocols compared with the dose-dependent regime considered in this work.

The authors would like to thank G. Kachalova, V. Borshchevskiy, K. Polyakov, J. Reguera and R. Wierenga for kindly providing the crystals 6HNLO, bR, TvNiR, LACV and TIM, respectively. We thank Ulrike Kapp for provision of the standard test crystal systems. We thank Sean McSweeney for his useful suggestions and help with the preparation of the manuscript. We gratefully acknowledge partial access to the beam time at the ESRF through the Radiation Damage BAG.

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.
- Alahuhta, M. & Wierenga, R. K. (2010). *Proteins Struct. Funct. Bioinform.* **78**, 1878–1888.

- Axford, D., Owen, R. L., Aishima, J., Foadi, J., Morgan, A. W., Robinson, J. I., Nettleship, J. E., Owens, R. J., Moraes, I., Fry, E. E., Grimes, J. M., Harlos, K., Kotecha, A., Ren, J., Sutton, G., Walter, T. S., Stuart, D. I. & Evans, G. (2012). *Acta Cryst.* **D68**, 592–600.
- Barker, A. I., Southworth-Davies, R. J., Paithankar, K. S., Carmichael, I. & Garman, E. F. (2009). *J. Synchrotron Rad.* **16**, 205–216.
- Bartunik, H. D., Summers, L. J. & Bartsch, H. H. (1989). *J. Mol. Biol.* **210**, 813–828.
- Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X., Ebner, S. G., Isenegger, A., Schneider, R., Schneider, J., Glettig, W., Pradervand, C., Panepucci, E. H., Tomizaki, T., Wang, M. & Schulze-Briese, C. (2011). *Cryst. Growth Des.* **11**, 916–923.
- Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Philipps, D. C. & Sarma, V. R. (1965). *Nature (London)*, **206**, 757–761.
- Bode, W. & Huber, R. (1978). *FEBS Lett.* **90**, 265–269.
- Borek, D., Cymborowski, M., Machius, M., Minor, W. & Otwinowski, Z. (2010). *Acta Cryst.* **D66**, 426–436.
- Borshchevskiy, V. I., Round, E. S., Popov, A. N., Buldt, G. & Gordeliy, V. I. (2011). *J. Mol. Biol.* **409**, 813–825.
- Bourenkov, G. P. & Popov, A. N. (2010). *Acta Cryst.* **D66**, 409–419.
- Boyko, K. M., Polyakov, K. M., Tikhonova, T. V., Slutsky, A., Antipov, A. N., Zvyagilskaya, R. A., Bourenkov, G. P., Popov, A. N., Lamzin, V. S. & Popov, V. O. (2006). *Acta Cryst.* **F62**, 215–217.
- Brockhauser, S., Svensson, O., Bowler, M. W., Nanao, M., Gordon, E., Leal, R. M. F., Popov, A., Gerring, M., McCarthy, A. A. & Gotz, A. (2012). *Acta Cryst.* **D68**, 975–984.
- Charron, C., Kadri, A., Robert, M.-C., Giegé, R. & Lorber, B. (2002). *Acta Cryst.* **D58**, 2060–2065.
- De la Mora, E., Carmichael, I. & Garman, E. F. (2011). *J. Synchrotron Rad.* **18**, 346–357.
- Dhouib, K., Khan Malek, C., Pflöging, W., Gauthier-Manuel, B., Duffait, R., Thuillier, G., Ferrigno, R., Jacquamet, L., Ohana, J., Ferrer, J. L., Théobald-Dietrich, A., Giegé, R., Lorber, B. & Sauter, C. (2009). *Lab Chip*, **9**, 1412–1421.
- Dubnovitsky, A. P., Ravelli, R. B., Popov, A. N. & Papageorgiou, A. C. (2005). *Protein Sci.* **14**, 1498–1507.
- Dunlop, K. V., Irvin, R. T. & Hazes, B. (2005). *Acta Cryst.* **D61**, 80–87.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Gabadinho, J., Beteva, A., Gujjarro, M., Rey-Bakaikoa, V., Spruce, D., Bowler, M. W., Brockhauser, S., Flot, D., Gordon, E. J., Hall, D. R., Lavault, B., McCarthy, A. A., McCarthy, J., Mitchell, E., Monaco, S., Mueller-Dieckmann, C., Nurizzo, D., Ravelli, R. B. G., Thibault, X., Walsh, M. A., Leonard, G. A. & McSweeney, S. M. (2010). *J. Synchrotron Rad.* **17**, 700–707.
- Garman, E. F. (2010). *Acta Cryst.* **D66**, 339–351.
- Halle, B. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 4793–4798.
- Hansen, C. L., Classen, S., Berger, J. M. & Quake, S. R. (2006). *J. Am. Chem. Soc.* **128**, 3142–3143.
- Hargreaves, D. (2012). *J. Appl. Cryst.* **45**, 138–140.
- Hendrickson, W. A. (1976). *J. Mol. Biol.* **106**, 889–893.
- Hogle, J., Rao, S. T., Mallikarjunan, M., Beddell, C., McMullan, R. K. & Sundaralingam, M. (1981). *Acta Cryst.* **B37**, 591–597.
- Holton, J. M. (2009). *J. Synchrotron Rad.* **16**, 133–142.
- Incardona, M.-F., Bourenkov, G. P., Levik, K., Pieritz, R. A., Popov, A. N. & Svensson, O. (2009). *J. Synchrotron Rad.* **16**, 872–879.
- Juers, D. H. & Matthews, B. W. (2001). *J. Mol. Biol.* **311**, 851–862.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kachalova, G. S., Bourenkov, G. P., Mengesdorf, T., Schenk, S., Maun, H. R., Burghammer, M., Riekel, C., Decker, K. & Bartunik, H. D. (2010). *J. Mol. Biol.* **396**, 785–799.
- Kisselman, G., Qiu, W., Romanov, V., Thompson, C. M., Lam, R., Battaile, K. P., Pai, E. F. & Chirgadze, N. Y. (2011). *Acta Cryst.* **D67**, 533–539.
- Kmetko, J., Hussein, N. S., Naides, M., Kalinin, Y. & Thorne, R. E. (2006). *Acta Cryst.* **D62**, 1030–1038.
- Kmetko, J., Warkentin, M., English, U. & Thorne, R. E. (2011). *Acta Cryst.* **D67**, 881–893.
- Krojer, T. & von Delft, F. (2011). *J. Synchrotron Rad.* **18**, 387–397.
- Leal, R. M. F., Bourenkov, G. P., Svensson, O., Spruce, D., Gujjarro, M. & Popov, A. N. (2011). *J. Synchrotron Rad.* **18**, 381–386.
- Li, L. & Ismagilov, R. F. (2010). *Annu. Rev. Biophys.* **39**, 139–158.
- Maire, A. le, Gelin, M., Pochet, S., Hoh, F., Pirocchi, M., Guichou, J.-F., Ferrer, J.-L. & Labesse, G. (2011). *Acta Cryst.* **D67**, 747–755.
- Marquart, M., Walter, J., Deisenhofer, J., Bode, W. & Huber, R. (1983). *Acta Cryst.* **B39**, 480–490.
- Morris, R. J., Blanc, E. & Bricogne, G. (2004). *Acta Cryst.* **D60**, 227–240.
- Mueller-Dieckmann, C., Panjikar, S., Schmidt, A., Mueller, S., Kuper, J., Geerlof, A., Wilmanns, M., Singh, R. K., Tucker, P. A. & Weiss, M. S. (2007). *Acta Cryst.* **D63**, 366–380.
- Murray, J. W., Garman, E. F. & Ravelli, R. B. G. (2004). *J. Appl. Cryst.* **37**, 513–522.
- Nanao, M. H., Sheldrick, G. M. & Ravelli, R. B. G. (2005). *Acta Cryst.* **D61**, 1227–1237.
- Nave, C. & Garman, E. F. (2005). *J. Synchrotron Rad.* **12**, 257–260.
- Nowak, E., Brzuszkiewicz, A., Dauter, M., Dauter, Z. & Rosenbaum, G. (2009). *Acta Cryst.* **D65**, 1004–1006.
- Nurizzo, D., Mairs, T., Gujjarro, M., Rey, V., Meyer, J., Fajardo, P., Chavanne, J., Biasci, J.-C., McSweeney, S. & Mitchell, E. (2006). *J. Synchrotron Rad.* **13**, 227–238.
- Owen, R. L., Axford, D., Nettleship, J. E., Owens, R. J., Robinson, J. I., Morgan, A. W., Doré, A. S., Lebon, G., Tate, C. G., Fry, E. E., Ren, J., Stuart, D. I. & Evans, G. (2012). *Acta Cryst.* **D68**, 810–818.
- Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4912–4917.
- Paithankar, K. S., Owen, R. L. & Garman, E. F. (2009). *J. Synchrotron Rad.* **16**, 152–162.
- Polyakov, K. M., Boyko, K. M., Tikhonova, T. V., Slutsky, A., Antipov, A. N., Zvyagilskaya, R. A., Popov, A. N., Bourenkov, G. P., Lamzin, V. S. & Popov, V. O. (2009). *J. Mol. Biol.* **389**, 846–862.
- Popov, A. N. & Bourenkov, G. P. (2003). *Acta Cryst.* **D59**, 1145–1153.
- Prates, J. A., Tarbouriech, N., Charnock, S. J., Fontes, C. M., Ferreira, L. M. & Davies, G. J. (2001). *Structure*, **9**, 1183–1190.
- Rajendran, C., Dworkowski, F. S. N., Wang, M. & Schulze-Briese, C. (2011). *J. Synchrotron Rad.* **18**, 318–328.
- Reguera, J., Weber, F. & Cusack, S. (2010). *PLoS Pathog.* **6**, e1001101.
- Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G. & Cipriani, F. (2009). *Acta Cryst.* **D65**, 1237–1246.
- Sanishvili, R., Yoder, D. W., Pothineni, S. B., Rosenbaum, G., Xu, S., Vogt, S., Stepanov, S., Makarov, O. A., Corcoran, S., Benn, R., Nagarajan, V., Smith, J. L. & Fischetti, R. F. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 6127–6132.
- Smith, G. D., Pangborn, W. A. & Blessing, R. H. (2005). *Acta Cryst.* **D61**, 1476–1482.
- Southworth-Davies, R. J., Medina, M. A., Carmichael, I. & Garman, E. F. (2007). *Structure*, **15**, 1531–1541.
- Tikhonova, T. V., Slutsky, A., Antipov, A. N., Boyko, K. M., Polyakov, K. M., Sorokin, D. Y., Zvyagilskaya, R. A. & Popov, V. O. (2006). *Biochim. Biophys. Acta*, **1764**, 715–723.
- Warkentin, M., Badeau, R., Hopkins, J. B., Mulichak, A. M., Keefe, L. J. & Thorne, R. E. (2012). *Acta Cryst.* **D68**, 124–133.
- Warkentin, M., Badeau, R., Hopkins, J. & Thorne, R. E. (2011). *Acta Cryst.* **D67**, 792–803.
- Warkentin, M. & Thorne, R. E. (2010). *Acta Cryst.* **D66**, 1092–1100.
- Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.