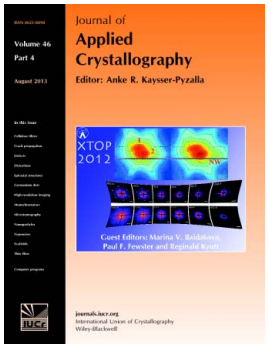
Journal of
Applied Crystallography
ISSN 0021-8898

RADDOSE-3D: time- and space-resolved modelling of dose in macromolecular crystallography

Oliver B. Zeldin, Markus Gerstel and Elspeth F. Garman

J. Appl. Cryst. (2013). **46**, 1225–1230



Journal of Applied Crystallography covers a wide range of crystallographic topics from the viewpoints of both techniques and theory. The journal presents papers on the application of crystallographic techniques and on the related apparatus and computer software. For many years, the *Journal of Applied Crystallography* has been the main vehicle for the publication of small-angle scattering papers and powder diffraction techniques. The journal is the primary place where crystallographic computer program information is published.

Journal of
Applied Crystallography
ISSN 0021-8898

Received 22 January 2013
Accepted 26 April 2013

RADDOSE-3D: time- and space-resolved modelling of dose in macromolecular crystallography

Oliver B. Zeldin, Markus Gerstel and Elspeth F. Garman*

Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK. Correspondence e-mail: elspeth.garman@bioch.ox.ac.uk

RADDOSE-3D allows the macroscopic modelling of an X-ray diffraction experiment for the purpose of better predicting radiation-damage progression. The distribution of dose within the crystal volume is calculated for a number of iterations in small angular steps across one or more data collection wedges, providing a time-resolved picture of the dose state of the crystal. The code is highly modular so that future contributions from the community can be easily integrated into it, in particular to incorporate online methods for determining the shape of macromolecular crystals and better protocols for imaging real experimental X-ray beam profiles.

1. Introduction

Radiation damage is one of the most common causes for the failure of X-ray macromolecular crystallography (MX) experiments (Holton, 2009; Garman, 2010). This damage is caused by absorption of some of the incident X-ray photons by the crystal. The energy deposited in the crystal leads to chemical transformations which eventually lead to a decrease in the intensity of the diffraction pattern and ultimately compromise the quality of the final calculated electron density. In some cases, damage can be so severe that it is not possible to solve the structure and obtain reasonable electron density using the diffraction images collected while the damage progresses. A number of metrics have been proposed to quantify the progression of radiation damage to protein crystals, and since the earliest days of MX radiation-damage research, dose (D, energy deposited per unit mass) has been accepted as the appropriate calculated metric against which to evaluate these radiation-damage indicators (Blake & Phillips, 1962).

The program *RADDOSE* has evolved through three iterations: v1 (Murray *et al.*, 2004), v2 (Paithankar *et al.*, 2009) and v3 (Paithankar & Garman, 2010). It provides a one-dimensional model of absorbed dose in a non-rotating macromolecular crystal, and allows predictive estimates of crystal lifetime to be made through comparison of the results with an experimentally determined dose limit (Owen *et al.*, 2006). For cases where the crystal is smaller than the beam, and where the beam has a uniform profile, this provides a valuable piece of information and input for use by strategy optimization tools such as *BEST* (Bourenkov & Popov, 2010).

However, modern advances in synchrotron technology, in particular data collection using micro-beams between 1 and 30 μm in size (*e.g.* Riekel *et al.*, 2005; Axford *et al.*, 2012), have led to the dose distribution in crystals often being highly inhomogeneous (Zeldin *et al.*, 2013). This paper presents a new predictive software tool, *RADDOSE-3D*, which allows full three-dimensional modelling of dose distributions for MX.

1.1. Interactions of X-rays with macromolecular crystals: *RADDOSE* v1 to v3

At the energies relevant for standard MX experiments (5–20 keV), there are three physical processes that contribute to the attenuation of the X-ray beam within the crystal: elastic scattering, which gives rise to the diffraction pattern observed at the detector, the photoelectric effect, which dominates the absorption cross section, and Compton scattering, which has a low probability at these energies but becomes important at energies above 20 keV, rising to 67% of all the interaction events at 40 keV (Paithankar & Garman, 2010). Each atom type from which the crystal is constituted has an individual energy-dependent cross section for these phenomena. Heavier atoms generally have larger absorption coefficients and contribute disproportionately to the total energy absorbed from the X-ray beam by a crystal.

In order to calculate an overall X-ray absorption cross section for the crystal, the unit-cell composition must be known. *RADDOSE* calculates this from the unit-cell size, the number of monomers, the monomer primary sequence length (protein or DNA), the number of heavy atoms in the monomer (sulfur from cysteines and methionines, or selenium from seleno-methionine substitutions, and/or known bound metal ions), the millimolar concentration of heavy atoms in the solvent and, optionally, the solvent content. If no solvent content is provided, the program calculates this from steric and volume considerations and fills the remaining volume of the unit cell with water.

The first version of *RADDOSE* only took into account photoelectric absorption, which is by far the dominant effect at the relevant incident X-ray energies. Photoelectron ejection leaves the interacting atom in an excited state, from which it can decay *via* one of two paths: Auger electron emission or fluorescent photon emission. The probability of fluorescent escape (as compared to Auger emission) is small for the atoms that constitute typical proteins and thus was not considered in the first version of *RADDOSE*. However, for larger Z atoms, fluorescent escape can occur at a significant probability (*e.g.* 30% of K-shell de-excitations for an iron atom are by fluorescent decay and 70% by Auger emission). The second version of *RADDOSE* (Paithankar *et al.*, 2009) calculated both the fluorescent emission probability and the probability that these fluorescent photons might escape the crystal rather than be reabsorbed, making the simplifying assumption that all the fluorescent photons are produced in the centre of the crystal. The energy loss in the crystal due to the Compton effect (inelastic X-ray scattering) was included in version 3 of *RADDOSE* with the appropriate angular distribution and self-absorption treatment (Paithankar & Garman, 2010). This version was made available on request to groups working with high-energy incident beams ($E_{\text{X-ray}} > 20$ keV) for MX.

RADDOSE v2 has become a widely used tool for estimating the absorbed dose in MX experiments, owing to its convenient use of biologically relevant parameters as inputs and its detailed consideration of the relevant physics. An alternative approach (Warkentin *et al.*, 2012), yielding results within 5% of those output by *RADDOSE*, is to compile the crystal composition manually and use mass energy-absorption coefficients from the NIST tables (http://www.nist.gov/pml/data/xray_gammaray.cfm).

To calculate dose, *RADDOSE* takes the total absorption coefficient (μ_{abs}) and calculates the energy absorbed by the crystal using a simple exponential absorption function of the crystal thickness, z:

$$\text{Absorbed energy} = E_{\text{Incident}}[1 - \exp(-\mu_{\text{abs}} z)]. \quad (1)$$

This treatment is correct for a crystal that is completely immersed in a beam with a flat top-hat profile. In the case of a Gaussian beam, *RADDOSE* uses the incident photon flux and beam FWHM x and y values to calculate the peak beam intensity, and then uses this flux to provide a maximum dose: a worst case scenario for damage. This dose is correct for the central part of the irradiated crystal which is subjected to the beam peak intensity for the entire rotation range. However, this approximation has obvious limitations in an era when many beams are smaller than the crystals being irradiated and when the beams can have highly featured profiles (Krojer & Von Delft, 2011). One solution to this problem is to use a sacrificial crystal to experimentally characterize radiation-damage sensitivity before determining an optimal data collection strategy. This approach is one of the options implemented in the program *BEST* (Bourenkov & Popov, 2010). Experimental characterization has the obvious disadvantage of requiring a spare crystal of the desired type to be available and can add a time burden since a data collection and burn series must be performed before the desired data collection can proceed.

1.2. Limitations of current predictive models of radiation damage

A simple one-dimensional single-body treatment of the crystal as a non-rotating object presents four major limitations for the optimal prediction of crystal lifetimes. (i) The photoelectron may escape from the crystal surface and the accompanying dose reduction is not taken into account. (ii) The range of doses present within the diffracting volume of a crystal, as a consequence either of rotation or of non-uniform beam profiles, is not taken into account. (iii) Realistic crystal shapes and/or beam profiles cannot be incorporated into the calculations. (iv) If the crystal is larger than the beam, new volumes of the crystal rotate into the beam during the experiment.

Previous Monte Carlo simulations and experimental work have shown that photoelectrons produced by the absorption of 18.5 keV X-rays in a protein crystal cause damage up to about 4 µm from the site of ionization (Nave & Hill, 2005; Sanishvili *et al.*, 2011). As a consequence, for very small crystals, a significant proportion of the initially absorbed energy may escape the crystal surface and thus reduce the effective dose. Nave & Hill (2005) calculated that for a 4 × 4 × 4 µm crystal, approximately seven-eighths of the energy would escape the crystal volume. Although this is an issue for platforms using micro- and nano-crystals such as the X-ray free-electron laser (X-FEL) (Chapman *et al.*, 2011), for crystals of the order of 100 µm in size or more, these effects are small. The probability of photoelectron escape is not included in the work reported here; however, the full three-dimensional treatment of a rotating crystal described below provides the framework necessary for such considerations to be realistically modelled in the future.

The consequence of not being able to model factors (ii)–(iv) above for systematic radiation-damage studies has been that crystals must always be chosen that are smaller than the beam and that experimental design must be carefully optimized to create an even dose profile (Kmetko *et al.*, 2006; Owen *et al.*, 2006). Thus, existing guidelines for conducting experiments that take into account the effect of dose on data quality are largely derived for even-dose cases. Examples of such guidelines for 100 K data collection include the experimental dose limit of 30 MGy, at which the diffraction intensity will have dropped to 70% or less of its initial value (Owen *et al.*, 2006), and the coefficient of sensitivity, which gives a value for the expected rate of increase in relative B value, B_{rel}, as radiation damage

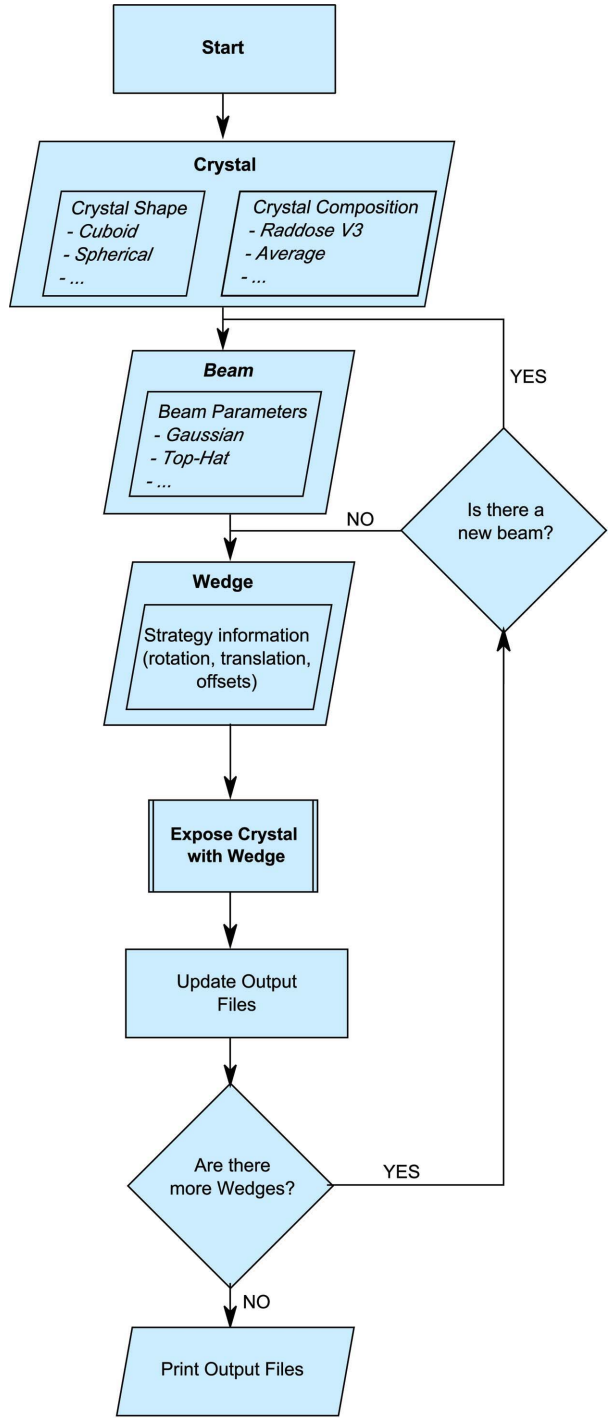


Figure 1
Flow chart illustrating the structure of the code.

progresses: $S_{AD} = \Delta B_{rel}/8\pi^2\Delta D \simeq 0.014$ Å^2 MGy^{-1} (Kmetko *et al.*, 2006).

For the general application of dose-lifetime guidelines to the crystallographic determination of novel structures, this one-dimensional model is appropriate within the regime for which it was designed. However, it offers limited predictive power for cases where dose contrast is high. For instance, for the example of the use of the program presented in §5, the dose at the centre of the crystal is ∼11 times that of the average for the whole crystal. In cases such as this, a one-dimensional model is not appropriate, and there is a clear need for a more sophisticated three-dimensional treatment as described here.

Additionally, a minor improvement to the accuracy of the simulation has been made by distinguishing between the absorption (μ_{abs}) and the attenuation coefficient (μ_{att}) when calculating the dose at each depth point in the crystal. Previous versions of *RADDOSE* have treated μ_{abs} and μ_{att} as equal since the difference is very small for the conditions typically used in MX where 2% absorption of the beam is typical (Garman, 2010). Because *RADDOSE-3D* calculates a full three-dimensional dose field, the new program treats these two coefficients separately and thus corrects this approximation.

2. Logical structure of the code

The design principle behind *RADDOSE-3D* was to create an open-source package that not only provides a full three-dimensional simulation of MX experiments for the calculation of dose profiles but also creates a general framework that could be extended to treat other macroscopic phenomena in MX, such as absorption corrections for low-energy incident photons. In its most general sense, this package takes a voxel field representation of a crystal and a computational model of a beam, and then simulates the interaction of the two *via* a strategy here described by a wedge which specifies the angular range and translations associated with a single exposure within an MX experiment. The crystal has various properties associated with each voxel: position in three-dimensional space, dose and fluence (photons per unit area). For a given small volume element in the crystal, the dose is a function of the energy absorbed in that element. This is calculated for each step using the intensity at the leading edge of the volume[1] and the proportion of the beam absorbed by that volume element: parameterized by the absorption coefficient. The beam is described by the photon energy, an intensity profile and the total flux.

The logical flow of the code, as shown in Fig. 1, starts by using the parameters defined in the input file or through the online interface available at http://www.raddo.se to set up the three components (crystal, beam and wedge) described above. Input files describe a single crystal and one or more beams and wedges of contiguous data collection; an example is shown in Fig. 2. The full grammar for the input files is specified in detail in the supplementary material.[2] A single run of the program thus always acts on a single crystal but can simulate several wedges (*e.g.* in different places on the crystal), each wedge being associated with the most recently defined beam. For instance, different beams might describe changes in energy for multiple-wavelength anomalous diffraction data collection, or changes in flux due to a change in transmission or an electron storage ring refill. The code works its way through the simulation in the order described in the input file. Pseudo-code for the calculation performed during the exposure routine is shown in Fig. 3. The output files are

```
# Complicated input file for RADDOSE-3D
Crystal
Type Cuboid
Dimensions 50 300 50
AngleP 45
AngleL 0
PixelsPerMicron 0.48

CoefCalc RDV3
unitcell 78 78 78
NumResidues 51
nummonomers 24
ProteinHeavyAtoms S 6 Zn 2
SolventHeavyConc P 0.425

Beam
Type Gaussian
Flux 1e12
FWHM 20 20
Energy 12.4
Collimation Rectangular 60 60

Wedge 0 360
ExposureTime 60
AngularResolution 1
StartOffset -75 -75

Wedge 0 360
ExposureTime 60
StartOffset -45 -45
TranslatePerDegree 0.125 0.125

Beam
Type Gaussian
Flux 4e11
FWHM 15 15
Energy 12.4
Collimation Rectangular 15 15

Wedge 0 360
ExposureTime 60
StartOffset 50 50
RotAxBeamOffset 0

Wedge 0 360
ExposureTime 60
StartOffset 75 75
RotAxBeamOffset 20
```

Figure 2
Structure of a *RADDOSE-3D* input file. The crystal is first described by a 'Crystal' block (salmon coloured in the electronic version of the journal), which remains active throughout. Next, a beam (the green block), followed by one or more wedges (blue), is described. As illustrated in Fig. 1, each wedge defines how to simulate an exposure of the previously specified beam with the crystal. Here the first two wedges will act with the first beam and the last two with the second beam. The full input syntax is described in the supplementary material.

```
for each angle within range and resolution set in Wedge
  for each voxel in Crystal
    let coords be voxel coordinates of voxel
    Rotate coords by angle
    Translate coords by angle*(transX, transY)
    Calculate fluence at coords using Beam
    Find depth of coords along beam to upstream face of crystal
    Calculate dose and fluence at coords
    Update dose and fluence for voxel
  end
end
Calculate aggregate metrics
Output Wedge aggregate metrics
```

Figure 3
Pseudo-code showing how the dose for each wedge is calculated by *RADDOSE-3D*.

[1] Itself calculated by taking the intensity at the face of the crystal and attenuating this for the depth of the volume element within the crystal.
[2] Supplementary material is available from the IUCr electronic archives (Reference: RR5041). Services for accessing this material are described at the back of the journal.

Table 1
Default output files produced by *RADDOSE-3D*.

File	Contents	Class
`output-Summary.txt` and STDOUT	Human readable description of experiment, output metrics and histograms of final dose state for each wedge.	`OutputSummaryText` implements `Output` and `ExposeObserver`
`output-Summary.csv`	Easily parsable summary of output metrics for each wedge.	`OutputSummaryCSV` implements `Output`
`output-DoseState.csv`	*x*, *y*, *z*, dose, fluence data for the final state of the crystal. For plotting dose maps.	`OutputFinalDoseStateCSV` implements `Output`

Table 2
Metrics output in the summary file of *RADDOSE-3D*.

Further discussion of the applicability of each of these metrics to different circumstances is presented by Zeldin *et al.* (2013).

Average dose–whole crystal, AD–WC	Total absorbed energy (J) divided by the mass of the whole crystal.
Average dose–exposed region, AD–ER	Total absorbed energy (J) divided by the mass of a region defined as any voxel having nonzero dose.
Threshold average dose, TAD	The average dose inside a volume bounded by the isosurface calculated to enclose a fraction of the total absorbed energy: the default is 95%.
Maximum dose, max dose	The dose in the voxel with the highest dose value anywhere in the crystal.
Dose contrast	Maximum dose divided by threshold average dose: a measure of the contrast in the dose distribution.
Used volume	The fraction of the crystal voxels with a nonzero dose.
Absorbed energy	The sum of the absorbed energy for all the voxels in the crystal.
Final dose histogram	A histogram showing the fraction of voxels in each dose range.

updated after calculating each wedge, so that a step-by-step report of the dose state of the crystal can be provided at the end of the simulation. This step-by-step output allows information to be obtained on the time course of the dose accumulation.

3. Technical details of the code

RADDOSE-3D is written in Java and follows an object-oriented programming architecture. In this programming paradigm, concepts are described by classes which contain program code and a description of attributes. Objects are instances of such classes, wherein program code and associated data are combined into a consistent unit with an internal state. Objects can be used to represent a part of the actual experiment, such as a specific crystal in an experiment. The overall program structure is documented in Figs. S1–S6 in the supplementary material. Major elements of the diffraction experiment such as the crystal, the X-ray beam and wedges of X-ray exposure are represented by objects. These objects are created from configuration files, or standard input (STDIN), as illustrated in Fig. S2. This generic input structure can easily be extended to accommodate other sources, such as a future graphical user interface or automated streaming input from other programs in a beamline environment (*e.g.* beam profile and crystal topography).

`Crystal` objects contain a representation of the physical crystal volume as a set of evenly spaced voxel coordinates, which can be rotated and translated using functions supplied by the object. For each voxel the associated dose and fluence values are stored. The size of the voxels can be user defined. The recommended voxel size for convergence of the dose metrics is approximately one-tenth of the beam FWHM (minimum of *x* and *y*), as demonstrated by the simulation results shown in Fig. S7 of the supplementary material.

For any crystal type, the attenuation and absorption coefficients can be calculated from the unit-cell parameters using a previous version of *RADDOSE* (v1, 2 or 3), or alternatively average values of the absorption coefficient of 0.237 mm^{-1} and attenuation coefficient of 0.281 mm^{-1}, which are appropriate for an incident beam energy of 12.4 keV (1 Å), can be used. These values were calculated (Holton & Frankel, 2010) using an average protein composition from a survey of the Protein Data Bank (Berman *et al.*, 2002). When using these averages, the crystal density is taken as 1.2 g ml^{-1} with an assumed 50% solvent content with a solvent mass of ∼1 g ml^{-1} and a protein mass of 1.41 g ml^{-1} (Fischer *et al.*, 2004).

The `Beam` interface facilitates access to information describing the beam profile, energy and flux (Fig. S4). `Wedge` objects contain all the information pertaining to the data collection strategy: rotational range, any offset either in the starting orientation and position of the crystal or between the rotation axis and the beam axis, and any translation that is to take place during the wedge exposure.

The output of *RADDOSE-3D* is created by modules following the observer design pattern (Gamma *et al.*, 1994). This program structure allows easy creation of new output modules that can produce any required statistic in any format. Depending on the user's needs, different output modules can be selected. The default output module selection is shown in Table 1. A diagram of all the `Output` classes is given in Fig. S6.

4. Program output

The raw output of the three-dimensional dose simulation is a scalar field describing the dose state of the crystal at every voxel and at every time step. Because each wedge is broken into time steps corresponding to a defined angular resolution (default 2°), a fully time-resolved picture yields a dose field for each angular step computed during the simulation. Since it is neither practical nor very illuminating to manually look at several hundred (*e.g.* 180 dose fields for a single 360° wedge at default resolution) four-dimensional arrays, each with hundreds of thousands of entries (typical number of voxels ∼125 000), it is vital to compute aggregate metrics.

Because there is, as of yet, no experimental consensus regarding what appropriate dose-based metric should be used to describe uneven dose distributions (Zeldin *et al.*, 2013), the code currently produces several metrics, as summarized in Table 2. They range from the highly intuitive average dose–whole crystal (AD–WC), a simple measure of the average dose in the whole crystal volume, or max dose, the highest dose value at any point in the crystal volume, to the more sophisticated metrics, such as threshold average dose (TAD) or average dose–exposed volume (AD–EV), which provide potential routes to better describe the dose field of a crystal. Contrast metrics, such as dose contrast (max dose divided by TAD), are also output as a way to assess the approximate pattern of dose within the diffracting volume.

Consider a fully, evenly exposed crystal and a crystal with only one section highly damaged: both of these cases can have absorbed the same total number of photons and so have the same AD–WC. However it is clear that they will produce very different diffraction patterns. In the case of TAD, a better estimate is achieved by only including the volume that encloses 95% of the absorbed energy, so that unexposed or weakly exposed regions that are associated with the very weak tails of a typical Gaussian beam are thus discounted, leading to a more informative dose estimate. A full discussion of the various metrics listed in Table 2 is presented by Zeldin *et al.* (2013).

Because wedges can overlap each other, all the metrics except for absorbed energy are calculated for the whole crystal volume at the end of each wedge; they represent the dose state at that time point during the experiment. However, absorbed energy for a given wedge thus refers to the energy absorbed only during the specified wedges, not during earlier wedges (*i.e.* it is not cumulative).

In addition to a summary file containing these metrics which are calculated after each wedge, the code can currently output two additional files in comma separated value (CSV) format, as detailed in Table 1. Firstly it provides an `output-summary.csv` file, which contains the same information as the summary file but with one column per metric and one row per wedge. This is an easy to parse file for convenient use in batch processing and scripting. The second additional file is an `output-DoseField.csv` file which is in the format X, Y, Z coordinate, dose, and finally fluence, with one row for each voxel in the crystal. This provides a full map of the final dose field, after all the wedges have been exposed. It can be used to create three-dimensional figures and to visualize the state of the crystal using software such as *ParaView* by KitWare (http://www.paraview.org/) or the R programming language.

5. Example of the use of *RADDOSE-3D*

To illustrate the workflow of using the software, a simple 90° wedge data collection is simulated. The input file for this is described in Fig. 4(*a*). As described above, the crystal block states that the crystal is a cuboid of size 100 × 100 × 100 µm, and the absence of any explicit crystal composition instructions will cause the program to default to average absorption and attenuation coefficients. Pixel resolution, defined in pixels per micrometre, is a particularly important parameter to be specified. As detailed in §3, the resolution should not be smaller than one-tenth of the beam FWHM, which for a 20 × 40 µm FWHM beam such as the one considered here is 2 µm per voxel, hence 0.5 voxels µm^{-1}. Setting this value too low will lead to errors in dose estimation, particularly for max dose, and too high a value will increase the runtime needlessly. The beam block describes a Gaussian profile with a vertical FWHM of 40 µm and a horizontal FWHM of 20 µm, with 10^{12} 12.1 keV photons per second and an outer beam collimation of 80 × 40 µm (vertical by horizontal). This collimation refers to the size of the slits used to collimate the beam, but for uncollimated Gaussians, setting this to twice the beam FWHM will speed up the runtime of the code, while only introducing an error that is much smaller than the likely fidelity of the true beam profile to a perfect Gaussian. Finally, the wedge block defines a 90° data collection over 30 s that will be processed in angular steps of 2°. It is only worth changing the 2° default angular resolution in the case of very small wedges (*e.g.* 5° wedges used to probe relative B factors). Fig. 4(*b*) shows the summary output file for this simulation and Fig. 5 a dose map created from the `DoseField.csv` file using *Paraview*. As can be seen both visually from the dose map and through the variation in the dose metrics in the summary file, this is a highly inefficient use of the crystal volume, with a maximum dose of 12.9 MGy and an AD–WC of 1.2 MGy. The TAD of 3.7 MGy represents the average dose within the region that has absorbed 95% of the energy, and so offers an average that effectively weights the weakly diffracting tails of the Gaussian to zero.

6. Discussion

Thanks to the highly extensible approach chosen when designing the *RADDOSE-3D* software, there is tremendous scope to take into account experimentally determined beamline and crystal parameters. New `Beam` classes can be written to take an experimental profile measurement as input; `Crystal` classes could model accurate crystal shapes using experimental data from imaging techniques such as tomography (Brockhauser *et al.*, 2008) or online confocal microscopy (Khan *et al.*, 2012). The main dose calculation function has been

```
Crystal
Type Cuboid
Dimensions 100 100 100
PixelsPerMicron 0.5
Beam
Type Gaussian
Flux 1e12
FWHM 40 20
Energy 12.1
Collimation Rectangular 80 40
Wedge 0 90
ExposureTime 30
AngularResolution 2
```

(*a*)

```
Cuboid Crystal of size [100, 100, 100] um [x, y, z] at a resolution
of 2.00 microns per voxel edge.
Simple Dose Decay Model used. Diffraction intensity does not decay
with increasing dose.
Gaussian beam, 80.0x40.0 um with 40.00 by 20.00 FWHM (x by y) and
1.0e+12 photons per second at 12.10 keV.
Wedge 1:
Collecting data for a total of 30.0s from phi = 0.0 to 90.0 deg.
Dummy crystal coefficients used.
Absorption coefficient: 2.37e-04 /um.
Attenuation coefficient: 2.81e-04 /um.
Elastic coefficient: 1.80e-05 /um.
Density: 1.20 g/ml.
Average Dose (Whole Crystal)      : 1.17 MGy
Average Dose (Exposed Region)     : 2.89 MGy
Max Dose                          : 12.865 MGy
Average Dose (95.0 % of total absorbed energy threshold (0.60 MGy)):
3.65 MGy
Dose Contrast (Max/Threshold Av.) : 3.53
Used Volume                       : 40.4%
Absorbed Energy (this Wedge)      : 1.49e-03 J.
Dose Inefficiency (Max Dose/mJ Absorbed) : 8.7 1/g
Final Dose Histogram:
Bin  1,  0.0 to  0.1 MGy:  6.8 %
Bin  2,  0.1 to  3.4 MGy: 58.9 %
Bin  3,  3.4 to  6.7 MGy: 23.0 %
Bin  4,  6.7 to 10.1 MGy:  9.0 %
Bin  5, 10.1 to 13.4 MGy:  2.2 %
Bin  6, 13.4 to 16.7 MGy:  0.0 %
Bin  7, 16.7 to 20.0 MGy:  0.0 %
Bin  8, 20.0 to 23.4 MGy:  0.0 %
Bin  9, 23.4 to 26.7 MGy:  0.0 %
Bin 10, 26.7 to 30.0 MGy:  0.0 %
Bin 11, 30.0 MGy upwards:  0.0 %
```

(*b*)

Figure 4
(*a*) Simple input file describing a 90° wedge of data, for which the final dose field is visualized in Fig. 5. (*b*) Summary output file for the input shown in (*a*). Note the large spread between the maximum dose (12.9 MGy) and the average dose–whole crystal (1.2 MGy), and the inefficient use of the crystal volume: only 40% has a nonzero dose, and, as can be seen from the dose histogram, most of this is within the 0.1–3.4 MGy bin, so is weakly exposed.

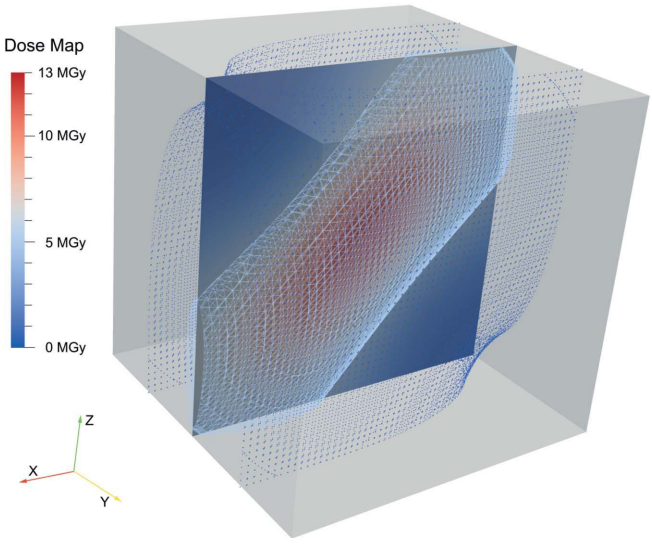


Figure 5
Dose map for the input configuration described in Fig. 4. Isosurfaces are shown at 0.05 and 5 MGy, and a slice map is shown in the plane perpendicular to the rotation axis and along the beam axis. The crystal has been rotated about the *y* axis, with the beam starting down the *z* axis and ending along the *x* axis, as labelled in the figure. Visualization produced using *Paraview*.

implemented such that it could be adapted to monitor a real-time feed of the beamline status (shutter status and goniometer position), allowing users to monitor dose during data collection and thus to make more informed decisions relating to the collection of multiple wedges from one crystal. An immediate example of this is that scientists at synchrotron light sources could implement their own `Beam` classes to simulate beam profiles from the experimental source at that particular beamline. Because these additions are encapsulated within broader interfaces, the way they interact with the rest of the code is pre-determined, and so they can be added without re-testing the broader functionality of the code.

More generally, the entire framework could be repurposed for the calculation of absorption corrections in cases where this becomes an important factor (*e.g.* for long-wavelength beamlines; Armin Wagner, personal communication).

We envisage this code being used on multiple levels: firstly as a routine online application for optimizing experimental protocols for radiation-sensitive crystals such as membrane proteins or multi-domain complexes with tenuous crystal contacts. Secondly, it will allow a more accurate reporting of the damage state of solved crystal structures, and thirdly, it provides a powerful characterization tool for radiation-damage studies, solving one of the major challenges in collecting systematic and statistically significant data for radiation-damage investigations (Allan *et al.*, 2013). It thus allows a broader range of protocols to be used and much more accurate estimation of dose to be implemented in the analysis.

To allow the general use of *RADDOSE-3D*, a free web service is provided at http://www.raddo.se. Users can submit configuration files as individual jobs which are then run from a queue. Once finished, the users are notified by e-mail and can then download the associated result files. Submitted files and results are only visible for the submitting user and are only retained for a limited number of days.

As the web service is a shared system, submitted jobs might be queued for some time depending on concurrent demand and estimated job runtime. Users who require the immediate processing of their job requests, have a large number of job requests or want to extend the operation of *RADDOSE-3D* with their customized output modules can obtain a copy of *RADDOSE-3D* by contacting the corresponding author.

7. Conclusions

We have presented a tool that takes the crystal, beam and experimental parameters of a macromolecular crystallography diffraction experiment and generates a step-by-step picture of the dose profile within the crystal as the data collection proceeds. This basic output can then be used to describe the damage profile within a crystal through the use of appropriate summary metrics, which serve either as a routine tool for predictively evaluating the damage state of a given macromolecular crystal or as a more accurate method of estimating dose for radiation-damage studies.

Given the heterogeneous nature of dose fields, it is vital that future work focuses on experimentally characterizing the metrics that can be extracted from a dose field and establishing those which best describe the damage state of the crystal, as observed in the diffraction pattern. Current research efforts focus on answering this question by performing large-scale analyses in test proteins using strategies that result in a wide range of dose profiles.

We gratefully acknowledge the Engineering and Physical Sciences Research Council (EPSRC) in the UK for funding through studentships at the Life Science Interface (OZ) and the Systems Biology (MG) programmes of the University of Oxford's Doctoral Training Centre. We also thank James Holton for valuable discussions during the development of this software and Jean-Paul Ebejer for the development of the online web interface.

References

Allan, E. G., Kander, M. C., Carmichael, I. & Garman, E. F. (2013). *J. Synchrotron Rad.* **20**, 23–36.
Axford, D. *et al.* (2012). *Acta Cryst.* D**68**, 592–600.
Berman, H. M. *et al.* (2002). *Acta Cryst.* D**58**, 899–907.
Blake, C. & Phillips, D. C. (1962). *Proceedings of the Symposium on the Biological Effects of Ionizing Radiation at the Molecular Level*, pp. 183–191. Vienna: International Atomic Energy Agency.
Bourenkov, G. P. & Popov, A. N. (2010). *Acta Cryst.* D**66**, 409–419.
Brockhauser, S., Di Michiel, M., McGeehan, J. E., McCarthy, A. A. & Ravelli, R. B. G. (2008). *J. Appl. Cryst.* **41**, 1057–1066.
Chapman, H. N. *et al.* (2011). *Nature*, **470**, 73–77.
Fischer, H., Polikarpov, I. & Craievich, A. F. (2004). *Protein Sci.* **13**, 2825–2828.
Gamma, E., Helm, R., Johnson, R. & Vlissides, J. (1994). *Design Patterns: Elements of Reusable Object-Oriented Software.* Reading: Addison Wesley.
Garman, E. F. (2010). *Acta Cryst.* D**66**, 339–351.
Holton, J. M. (2009). *J. Synchrotron Rad.* **16**, 133–142.
Holton, J. M. & Frankel, K. A. (2010). *Acta Cryst.* D**66**, 393–408.
Khan, I., Gillilan, R., Kriksunov, I., Williams, R., Zipfel, W. R. & Englich, U. (2012). *J. Appl. Cryst.* **45**, 936–943.
Kmetko, J., Husseini, N. S., Naides, M., Kalinin, Y. & Thorne, R. E. (2006). *Acta Cryst.* D**62**, 1030–1038.
Krojer, T. & von Delft, F. (2011). *J. Synchrotron Rad.* **18**, 387–397.
Murray, J. W., Garman, E. F. & Ravelli, R. B. G. (2004). *J. Appl. Cryst.* **37**, 513–522.
Nave, C. & Hill, M. A. (2005). *J. Synchrotron Rad.* **12**, 299–303.
Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4912–4917.
Paithankar, K. S. & Garman, E. F. (2010). *Acta Cryst.* D**66**, 381–388.
Paithankar, K. S., Owen, R. L. & Garman, E. F. (2009). *J. Synchrotron Rad.* **16**, 152–162.
Riekel, C., Burghammer, M. & Schertler, G. (2005). *Curr. Opin. Struct. Biol.* **15**, 556–562.
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.
Warkentin, M., Badeau, R., Hopkins, J. B., Mulichak, A. M., Keefe, L. J. & Thorne, R. E. (2012). *Acta Cryst.* D**68**, 124–133.
Zeldin, O. B., Gerstel, M. & Garman, E. F. (2013). *J. Synchrotron Rad.* **20**, 49–57.