Simulations of radiation damage as a function of the temporal pulse profile in femtosecond X-ray protein crystallography

H. Olof Jönsson,[a] Nicuşor Tîmneanu,[a,b] Christofer Östlin,[a] Howard A. Scott[c] and Carl Caleman[a,d]*

[a]Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden, [b]Department of Cell and Molecular Biology, Uppsala University, Biomedical Centre, Box 596, SE-751 24 Uppsala, Sweden, [c]Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA, and [d]Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron, Notkestrasse 85, DE-22607 Hamburg, Germany.
*E-mail: carl.caleman@physics.uu.se

Serial femtosecond X-ray crystallography of protein nanocrystals using ultrashort and intense pulses from an X-ray free-electron laser has proved to be a successful method for structural determination. However, due to significant variations in diffraction pattern quality from pulse to pulse only a fraction of the collected frames can be used. Experimentally, the X-ray temporal pulse profile is not known and can vary with every shot. This simulation study describes how the pulse shape affects the damage dynamics, which ultimately affects the biological interpretation of electron density. The instantaneously detected signal varies during the pulse exposure due to the pulse properties, as well as the structural and electronic changes in the sample. Here ionization and atomic motion are simulated using a radiation transfer plasma code. Pulses with parameters typical for X-ray free-electron lasers are considered: pulse energies ranging from 10^4 to 10^7 J cm^{-2} with photon energies from 2 to 12 keV, up to 100 fs long. Radiation damage in the form of sample heating that will lead to a loss of crystalline periodicity and changes in scattering factor due to electronic reconfigurations of ionized atoms are considered here. The simulations show differences in the dynamics of the radiation damage processes for different temporal pulse profiles and intensities, where ionization or atomic motion could be predominant. The different dynamics influence the recorded diffracted signal in any given resolution and will affect the subsequent structure determination.

Keywords: X-ray free-electron laser; serial femtosecond crystallography; radiation damage; plasma simulations.

1. Introduction

X-ray free-electron lasers enable high-resolution structural determination of biological materials by using short enough pulses to outrun many of the effects of radiation damage (Neutze *et al.*, 2000; Chapman *et al.*, 2011). This relatively new approach makes obtaining usable diffraction patterns from nanocrystals, or even single particles, possible. In fact, the ultrashort free-electron laser (FEL) pulses have been used to successfully obtain diffraction data from such samples in experiments (Lomb *et al.*, 2011; Boutet *et al.*, 2012; Johansson *et al.*, 2012; Koopmann *et al.*, 2012; Redecke *et al.*, 2013; Barends *et al.*, 2013*a*,*b*; Johansson *et al.*, 2013; Kern *et al.*, 2013, 2014; Liu *et al.*, 2013; Hattne *et al.*, 2014; Kupitz *et al.*, 2014; Weierstall *et al.*, 2014; Sawaya *et al.*, 2014) conducted at the

Linac Coherent Light Source (LCLS) (Emma *et al.*, 2010). However, the photon–matter interactions cause heavy ionization and subsequent sample decay, which despite the brief time scales may limit the diffractive possibilities (Chapman *et al.*, 2014). Understanding these processes is thereby paramount to identifying, and overcoming, present limits of serial femtosecond crystallography (SFX).

When a sample is exposed to X-rays the atoms will be ionized. The ionization will cause the molecular bonds to break and leads to subsequent changes in the structure. The ionization leads to both a resolution-dependent reduction in Bragg signal and the addition of uniform diffuse scattering (Barty *et al.*, 2012). In a crystalline sample the ionization is spread out among a large number of individual molecules, which reduces the influence on the diffracted signal. The

atomic cross sections will vary with the electronic state during a high-fluence pulse, further adding to the difficulties of predicting the sample decay (Son *et al.*, 2011).

When the inner-shell electrons are photoionized, short-lived states that render the atom more transparent to photons arise (Young *et al.*, 2010). While a continuous beam may ionize atoms fully, short pulses make it possible to control the ionization process so that the highest charge states are not reached (Hoener *et al.*, 2010). The laser pulse shape may alter these effects, especially when atoms in the transparent state are exposed to high-fluence pulses. In addition to the direct photoionization and following Auger decay, further ionization is created by electron impact when energetic electrons hit atoms and ions. The number of generated secondary electrons may outnumber the primary photoelectrons by a factor of 100 or more (Caleman *et al.*, 2009, 2011*a*).

The first theoretical predictions that were made suggested that a pulse length of 10 fs or shorter and a high number of photons (5×10^{12} photons in 100 nm focus) may be required to successfully determine the structure of a protein with atomic resolution (Neutze *et al.*, 2000) from very small crystals ($5 \times 5 \times 5$ unit cells). The first SFX experiments have used pulses of the order 50–100 fs (Chapman *et al.*, 2011; Boutet *et al.*, 2012; Barty *et al.*, 2012; Redecke *et al.*, 2013) and have suggested that atomic resolution is also reachable with longer pulses. Boutet *et al.* and Redecke *et al.* obtained atomic resolution with lower intensities than mentioned above using crystals considerable larger than $5 \times 5 \times 5$ unit cells. At high intensities (Barty *et al.*, 2012) the pulse length does not need to be shorter, as the diffraction is gated during the pulse by the loss of spatial coherence. As the atoms are disarranged, the ordered structure of the crystal is impaired and its signal-amplifying properties weakened. Eventually, the sample becomes virtually invisible to the late-arriving photons and the remaining tail of the pulse offers no usable contribution to the diffraction pattern (Caleman *et al.*, 2012). The apparent pulse length at a given resolution can thus be considerably shorter, and this can explain the achieved atomic resolution in protein crystal samples (Barty *et al.*, 2012; Caleman *et al.*, 2015).

In a typical SFX experiment, only a fraction of the collected frames are used for three-dimensional structural determination. In some experiments, as few as 10–20% of the detected hits can be indexed and used (Chapman *et al.*, 2011; Boutet *et al.*, 2012). Possible explanations as to why a large number of crystal hits cannot be used for data analysis is the unpredictability from shot to shot, the variation in the quality of the crystals, the quality of the FEL pulse and possibly the damage dynamics.

A FEL beam is generated in an undulator through the principle of self-amplified stimulated emission (SASE). Due to the stochastic nature of the SASE process that starts from random noise in the electron beam, individual pulses will have large statistical fluctuation (Tiedtke *et al.*, 2009). As a consequence, the temporal pulse profile may be different from pulse to pulse (Düsterer *et al.*, 2011), possibly affecting the diffraction output. Efforts have been made to control the lasing process by seeding (Zhao *et al.*, 2012) and self-seeding (Amann *et al.*, 2012; Allaria *et al.*, 2012), that provide a more well defined laser beam at a cost of a lower intensity. Methods to infer the temporal pulse profile from measurement of the electron bunch temporal profiles have been presented (Düsterer *et al.*, 2011). Recently, methods to directly measure the X-ray temporal pulse profile have been presented (Behrens *et al.*, 2014; Helml *et al.*, 2014). These studies show that the temporal pulse profile varies on a shot-to-shot basis. Three idealized pulse shapes have been identified based on Behrens *et al.* (2014) and Schneidmiller & Yurkov (2011). For comparison with earlier work a flat-top shaped pulse is included. We explore how different X-ray pulse profiles influence the damage dynamics and consequently Bragg diffraction.

This paper is organized as follows: §2 describes the plasma code, choice of simulation parameters, and the treatment of ionization and atomic displacement. §3 describes the simulation results. In §4 a general discussion and the limitations of our approach are presented. In §5 the main results and the conclusions are presented.

2. Method

2.1. Simulations

Interaction between the FEL beam and a protein crystal was simulated with the non-local thermodynamic equilibrium (non-LTE) radiation transfer code *CRETIN* (Scott, 2001; Scott & Mayle, 1994). The plasma model used is well suited for simulating warm dense matter where ionization due to high photon fluence is a considerable factor. The code has previously been able to reproduce experimental results from biological systems (density close to 1 g cm^{-3}) exposed to a FEL pulse (Barty *et al.*, 2012). Agreement between simulations and experiment has also been shown with high and low fluence using soft X-rays (Bergh *et al.*, 2008*a*; Andreasson *et al.*, 2011).

The simulation tracks the time evolution of radiation spectra, electronic states, transition rates and electronic level populations, radiation transport, opacities and heating rates (Caleman *et al.*, 2011*b*). A screened hydrogenic model was used for the element composition of the sample implemented, and the code accurately simulates Auger decay processes by taking life times of hollow atoms into account. *CRETIN* models changes in the absorption cross sections due to electron excitation and depletion of electronic states, as well as continuum lowering. Secondary ionization processes like electron–ion collisions are included, and instant thermalization of electrons is assumed where electrons follow a Maxwellian energy distribution. The choice of electron–ion coupling coefficient will affect the dynamics of the atom and ion energies in the system. Here the coefficient is calculated with Spitzer's formula (Spitzer, 1956), using a Coulomb logarithm introduced for dense systems (Gericke *et al.*, 2002). Another effect of a dense system is a lowering of the continuum edges, here calculated by the Stewart–Pyatt formula

(Stewart & Pyatt, 1966), a common approximation that has been tested against both experiments and more detailed models (Nantel *et al.*, 1998).

In the plasma model, the average atomic composition of the sample is simulated with no intrinsic structural information, and any structural changes must be calculated from average properties such as ion populations, transition rates and temperature (see Fig. 1). The model does not treat the breaking of atomic bonds; however, at the intensities studied here the system quickly turns into a plasma (typically within the first 1–2 fs) and the energy scales become larger than the average binding energy in the molecule. The advantage of using a plasma approach is that it can treat large systems such as protein crystals at an affordable computational cost. However, local structural changes, as reported by Lomb *et al.* (2011), and collective motion are not treated. To study these effects, methods such as molecular dynamics description (Neutze *et al.*, 2000; Caleman *et al.*, 2011*a*) or so-called hybrid models (Hau-Riege, 2012) should be employed.

Three different photon energies were simulated; 2 keV, 6 keV and 12 keV, which are relevant both for the currently operating LCLS and the upcoming European XFEL (Schneidmiller & Yurkov, 2011). To span a large set of beam intensities, pulses with intensities of 10^{17}–10^{20} W cm^{-2} were simulated. This corresponds to 10^4–10^7 J cm^{-2} in a 100 fs pulse. Such intensities are available at the LCLS today at the various experimental stations and could be expected using the 0.1 µm-diameter focus at the scientific instrument Single Particles, Clusters and Biomolecules (SPB) at the European XFEL (Mancuso, 2012; Mancuso *et al.*, 2013). For example, the low end of the simulations correspond roughly to the intensity at the CXI endstation with a micrometer focus, where many experiments have been performed to date (Boutet *et al.*, 2012; Redecke *et al.*, 2013; Kern *et al.*, 2014). The high end of the simulations corresponds to intensities in the 100 nm focus (Nass *et al.*, 2015). Note also that, depending on sample delivery technique, the samples will be exposed to various intensities on a shot-to-shot basis in the spatial intensity focus profile.

First, radiation damage dynamics in general are studied on a 100 fs time scale, which is the typical time scale for the thermalization of electrons with the ions (coupling between the free electrons and the ions/atoms in the sample). This time scale will capture the interplay between the ionizations and atomic displacements. Lastly, we will also look more specifically at 50 fs pulses, which are relevant in the context of recent experiments performed at the LCLS (Boutet *et al.*, 2012; Kupitz *et al.*, 2014).

Average ionization and average displacement were studied in a one-dimensional geometry consisting of a 700 nm-thick plane in 11 simulation zones (see Fig. 1). Each zone had the same atomic composition, corresponding to a Photosystem I protein crystal containing 78% solvent water: $H_{141400}O_{57300}C_{16900}N_{3310}S_{89}Fe_{12}Mg_{96}P_3Ca$ with density 1.077 g cm^{-3}. This particular sample is well known and has been studied in previous FEL crystallography experiments (Chapman *et al.*, 2011; Fromme & Spence, 2011; Aquila *et al.*, 2012).

The simulation geometry is illustrated in Fig. 1. Every zone is separately treated as a continuum with neutral net charge and mass conservation, but with radiation and heat transport occurring between neighboring zones. While this geometry is unsuitable for very small samples and at the edges of a protein crystal where escaping electrons and ions must be considered (Caleman *et al.*, 2011*a*), the middle zones describe well the bulk properties of a crystal. The number of zones can be chosen arbitrarily and is a tradeoff between computational efficiency for a small number of zones and a higher accuracy with a large number of simulation zones.

The approach used here will investigate average displacements and ionization, and provides a scaling of Bragg intensities, but it does not treat the coherent dislocations (resulting in moving of Bragg peaks) or crystal explosion (changes in the widths of Bragg peaks). Hydrodynamic expansion was not included in this study. An experimental scenario is assumed where the crystals are delivered with a liquid jet into the interaction region; in this situation crystal expansion is limited by the surrounding water. This has been shown to be viable for the short time scale of the pulse (Caleman *et al.*, 2011*b*). Any coherence effects that could occur due to the rapid ionization of heavy atoms are also not taken into account. Such effects have been suggested by simulations to be useful for phasing (Son *et al.*, 2011).

2.1.1. Pulse shapes. To study the impact of varying temporal pulse profiles a pulse shape function $s(t)$ is introduced. This function describes the variation in flux over time at any given fluence, pulse length and photon energy. The pulse flux at time t is given by $I(t) = I_0 s(t)$, where I_0 is the total pulse fluence and $s(t)$ is the normalized pulse shape factor defined as

$$(1/T) \int_{t=0}^{T} s(t)\, \mathrm{d}t = 1, \qquad (1)$$

where T is the pulse duration. For a specific pulse fluence I_0, a specific photon energy and pulse length, simulations differ only by the shape function $s(t)$ defining when photons arrive during the pulse. Different pulse shapes (see Fig. 1) were simulated, referred to as follows. *Flat*: constant flux throughout the pulse, representing the standard shape that has been modeled previously (Barty *et al.*, 2012); *single*: centered high-flux region (53% of the intensity in the middle 3/20 of the pulse) and broad tails with lower flux, that can represent a single spike above background, which could be expected for non-Gaussian pulses; *large–small*: high initial flux (44.5% of the total intensity in the first 3/20 of the pulse) and a lower final peak (20% of the total intensity in the last 3/20 of the pulse), to model a two-peak structure as observed by Behrens *et al.* (2014); and *multiple*: a comb structure consisting of six high-flux regions evenly interspaced by regions with zero flux, chosen to model various microbunching effects during the SASE process.

2.2. Analysis

The intensity of the instantaneous Bragg signal originating from a crystal exposed to a FEL pulse decreases on average
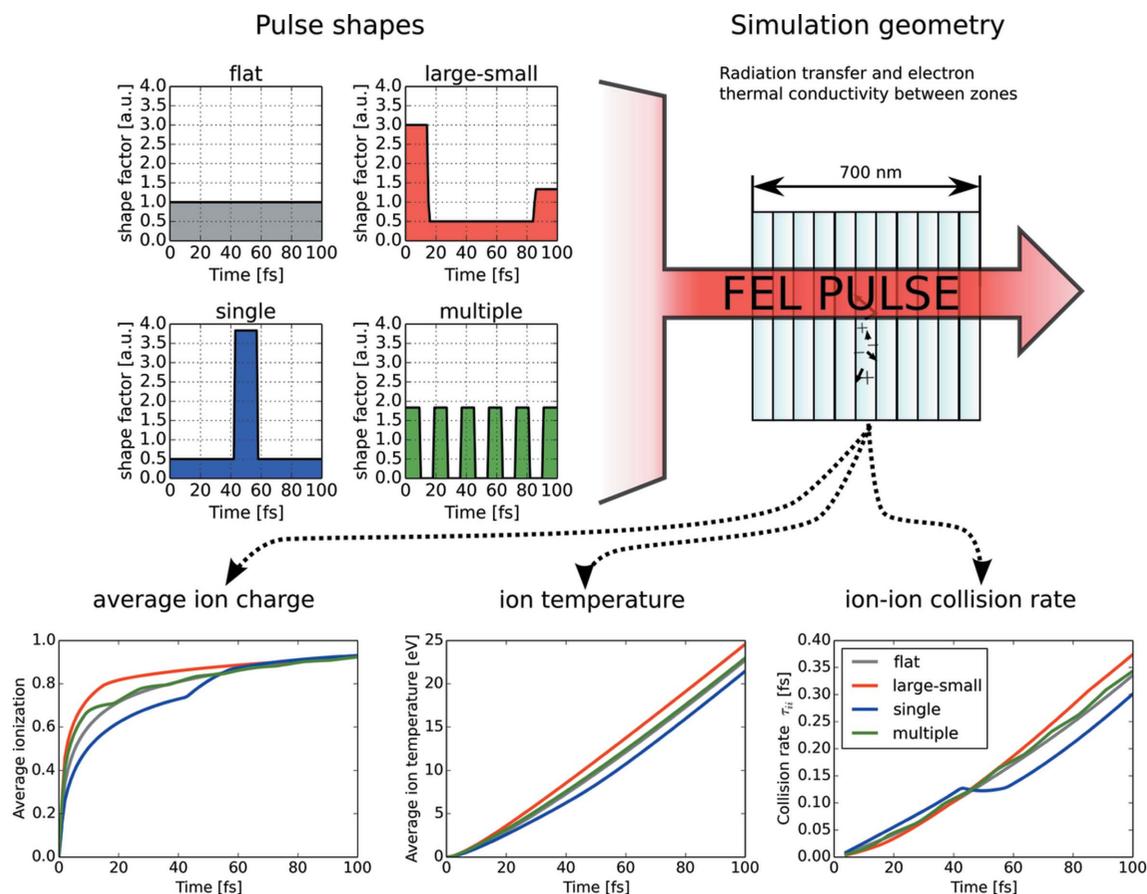
Figure 1
Simulation setup with pulse profiles, simulation geometry and typical output parameters. Top left: graphical representations of the four pulse shapes that were simulated. All pulses have a length of 100 fs and the same total fluence. The color scheme is kept throughout the paper: flat pulse (gray), large–small (red), single pulse (blue) and multiple (green). Top right: simulation geometry for protein nanocrystals in one dimension, divided into zones with the same chemical composition. The X-ray laser pulse is simulated in all the zones, which interact with each other through radiation transfer and thermal conductivity. Bottom row: plasma parameters from the simulations that are used to model radiation damage: average ion charge (charge 1 for fully ionized atoms), ion temperature and ion–ion collision rate. The plot shows the time dependence for the selected pulse profiles with a photon energy of 6 keV and pulse intensity of 2×10^6 J cm^{-2}.

during the exposure due to both ionization and displacement of the atoms. These two phenomena are introduced separately below, and they are considered together for their effect on the Bragg signal.

2.2.1. Ionization. The scattering factors of an atom will change depending on charge state. Degradation due to ionization $k(q, t)$, calculated as a function of time t and scattering vector **q**, is calculated from the atomic form factors and the population of states (Hau-Riege, 2007). This decay factor is the fraction of signal in relation to the neutral system,

$$k(q, t) = \langle f \rangle^2 / \langle f_0 \rangle^2, \qquad (2)$$

where the atomic form factor f is estimated by the parametrization introduced by Cromer & Mann (1968):

$$f[\sin(\theta)/\lambda] = \sum_{i=0}^{4} a_i \exp\{b_i[\sin(\theta)/\lambda]^2\} + c, \qquad (3)$$

where a_i, b_i and c are defined by Caleman *et al.* (2015), θ is the Bragg angle and λ is the wavelength. The scattering factors for ions decrease as they lose electrons. Hence, the factor $k(q, t)$ will vary with the ionization states present in the sample within the range [0, 1] and will decrease with increasing charge state

(as seen in Fig. 1). These dynamics of the decay factor $k(q, t)$ are shown in Figs. 2(b) and 2(d).

2.2.2. Displacement. To describe the degradation in Bragg diffraction due to displacement of atoms in the protein crystal lattice a previous approach (Barty *et al.*, 2012) was used. Displacement is described by atomic velocity (temperature) and collisions with atoms and ions within the Debye length. Ion temperature T_i and ion–ion collision rate $\tau_i(t)$ will increase during the pulse (see Fig. 1), as calculated from the plasma code. Root-mean-square displacement as a function of time $\sigma_i(t)$ is given by

$$\sigma_i(t) = \left[2N \int_0^t D_i(t')\, dt' \right]^{1/2}, \qquad (4)$$

where N is the number of dimensions (for our particular case, along the direction of the scattering vector, the dimension is taken to be $N = 1$). The diffusion coefficient $D_i(t)$ for an ion i is dynamic and at time t can be estimated as (Book, 2007)

$$D_i(t) = \frac{k_B T_i(t)\tau_i(t)}{m_i}, \qquad (5)$$

where k_B is Boltzmann's constant and m_i is the ion mass. Displacement of atoms in a crystal lattice will degrade the diffraction pattern. The scattering angles corresponding to high spatial frequencies are those first affected by atomic displacement. The degradation due to displacement as a function of time t and scattering vector q can be described with the following decay function (Caleman *et al.*, 2015),

$$h(q, t) = \exp\left[-4\pi^2 q^2 \sigma^2(t)\right], \quad (6)$$

where h is in the range [0, 1] and is decreasing with increasing q or σ. The magnitude of the scattering vector $\mathbf{q}$ is calculated as $q = 2\sin(\theta)/\lambda$, where θ is the Bragg angle and λ is the wavelength. Figs. 2(a) and 2(c) show the dynamics of the decay factor $h(q, t)$.

2.2.3. Bragg signal. The Bragg signal will be the integrated scattering over the range of the full pulse. The full Bragg signal from an undamaged sample at scattering vector $\mathbf{q}$ is calculated as

$$I_{\text{ideal}}(\mathbf{q}) = r_e^2 \Delta\Omega \left|F_0(\mathbf{q})\right|^2 I_0 \, (1/T) \int_{t=0}^{T} s(t) \, dt. \quad (7)$$

Here r_e is the classical electron radius, $\Delta\Omega$ is the solid angle of a pixel detector, $|F_0(\mathbf{q})|^2$ is the form factor particular for a sample (Caleman *et al.*, 2011a), I_0 is the beam intensity and $s(t)$ is the shape function. Under this ideal assumption of no damage, the Bragg signal at all angles is always directly proportional to the fluence. In a more realistic model, radiation damage will affect the Bragg signal depending on the scattering angle as
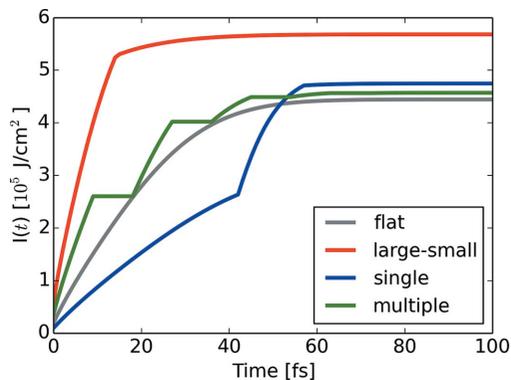

Figure 3
The accumulated Bragg signal $I(t) = (1/T) \, I_0 \int_0^t k(q, t') \, h(q, t') \, s(t') \, dt'$ as a function of time, from carbon at $q = 0.62$ Å^{-1}. A 700 nm PSI crystal exposed to a 100 fs-long pulse at 6 keV and intensity 2×10^6 J cm^{-2} was simulated. The accumulated intensity $I(t)$ is shown in J cm^{-2} and can be interpreted as an 'effective' photon intensity that includes pulse shape and damage effects, and should be multiplied by the term $r_e^2 \Delta\Omega |F_0(\mathbf{q})|^2$.

$$I_{\text{Bragg}}(\mathbf{q}) = r_e^2 \Delta\Omega \left|F_0(\mathbf{q})\right|^2 I_0 \, (1/T) \int_{t=0}^{T} k(q, t) \, h(q, t) \, s(t) \, dt. \quad (8)$$

This equation describes the signal I_{Bragg} from the full pulse as a function of the scattering vector $\mathbf{q}$. The full signal will depend on the fluence from the term $r_e^2 \Delta\Omega |F_0(\mathbf{q})|^2$, the incoming intensity I_0, and is limited by the degradation caused by the ionization function $k(q, t)$ and displacement function $h(q, t)$. The factors $k(q, t)$ and $h(q, t)$ both depend on the state of the plasma (ionization, temperature, collisions), and changes in the plasma dynamics will affect both factors.

All times t are weighted by the shape function $s(t)$, making it possible to quantify the effects of the pulse shape on the Bragg diffraction. An example of how the accumulated Bragg signal depends on the time t during the pulse is shown in Fig. 3.

3. Results

The change in scattering power is described by the displacement decay factor h and the ionization decay factor k: examples for a given intensity and energy are shown in Fig. 2. These factors depend on time and scattering angle. For the four different pulse shapes studied here the difference in k and h is seemingly small. The displacement in the sample is not strongly dependent of what pulse shape is employed (compare Fig. 2), when the number of photons is kept the same for all simulated pulse shapes. Displacement is driven by ion temperature and ion–ion collisions, and these parameters are slowly varying
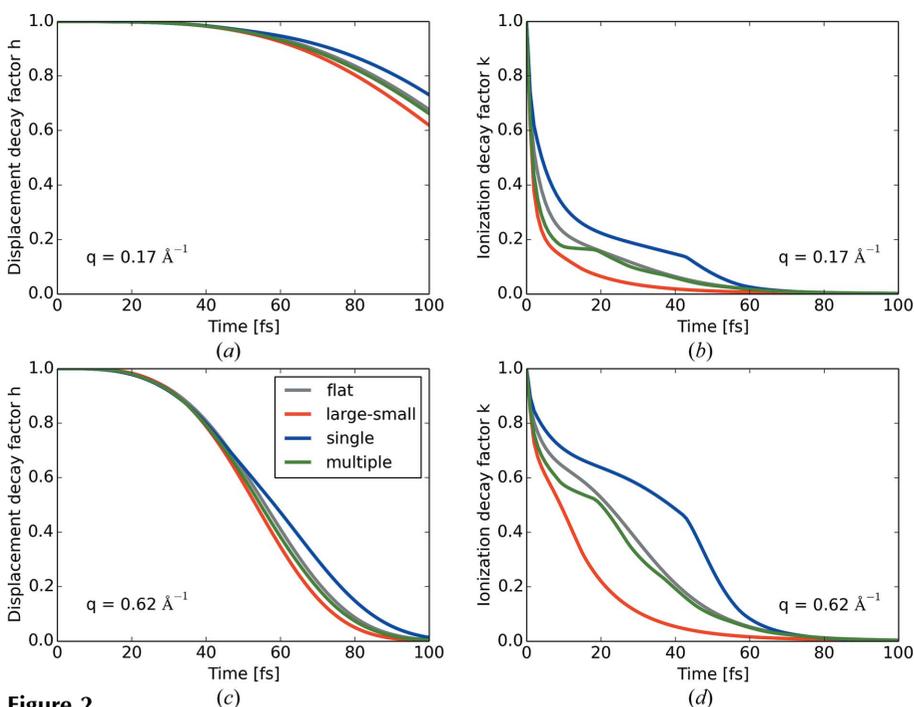

Figure 2
Decay factors for a scattered Bragg signal from carbon for different scattering vectors as a function of time. A 700 nm PSI crystal exposed to a 100 fs-long pulse with four pulse shapes at 6 keV and intensity 2×10^6 J cm^{-2} was simulated. The factors have ranges from 1 for an unaffected crystal to 0 for a full termination of Bragg signal. (a) Atomic displacement $h(q, t)$ for $q = 0.17$ Å^{-1}, (b) ionization $k(q, t)$ for $q = 0.17$ Å^{-1}, (c) atomic displacement $h(q, t)$ for $q = 0.62$ Å^{-1}, (d) ionization $k(q, t)$ for $q = 0.62$ Å^{-1}.

with time in a similar way for all the simulated pulse shapes (see Fig. 1).

The ionization decay factor k on the other hand shows a stronger dependence on the pulse shape (Fig. 2). The ionization is more sensitive to photon peaks in the pulse, as seen in Fig. 1, although at the end of the pulse the overall ionization of the bulk is still the same. These variations are clearer in the factor k at low q, where the atomic form factors are very sensitive to the total number of electrons that scatter X-rays. The most obvious difference in the factor k is between the *large–small* and *single* temporal pulse profiles (Fig. 2). The difference is explained by the strong photon flux arriving either at the beginning of the pulse or in the middle of it.

Multiplying h by k with the pulse shape factor s [equation (8)], the actual effect of the X-ray pulse on the diffracted pattern can be described. Fig. 3 shows all these effects over the length of the pulse; the jumps in the overall signal show the onset of various peaks in the pulse and subsequent ionization. The saturation in the overall signal is due to the gating of Bragg scattering. Scattering at a specific angle will eventually be terminated as these sample-modifying processes lead to a lack of spatial coherence starting at small length scales (Fig. 2). Remaining parts of the pulse will fail to contribute to the Bragg signal, regardless of the pulse shape or the flux of photons.

Fig. 3 already indicates that the best-case scenario is a pulse that begins with a high flux of photons (in this case *large–small*).

Fig. 4 gives an overall view of the simulation results for the different pulse shapes and wavelengths, as a function of the magnitude of the scattering vector q and incident intensity I_0. The same range of intensities 10^4–10^7 J cm^{-2} is used for different wavelengths. The color scale in the figure describes the expected relative photon intensity of a Bragg peak (I_{Bragg}) without the structure factors $|F_0(\mathbf{q})|^2$ [equation (8)] for a specific scattering vector. In an experiment the actual intensity in a Bragg peak on the detector is also dependent of the characteristics of the sample, such as crystal quality, sample size and molecule type. The y-axis in each panel describes the incoming pulse intensity, I_0. Each row in the figure corresponds to a specific photon energy and each column to a specific pulse shape, as defined in Fig. 1. Going to higher intensities generally leads to more Bragg-scattered photons. The q values (x-axis) are chosen to correspond to Bragg angles from $0°$ to $90°$, at different energies. The same scattering angles enables higher q values at higher photon energy. Moving from small scattering angles (low q) to larger scattering angles at a specific pulse intensity will in some cases increase $I_{\text{Bragg}}/|F_0(\mathbf{q})|^2$ (see, for example, *large–small* at 6 keV) and sometimes decrease (*single* at 12 keV). This variation of
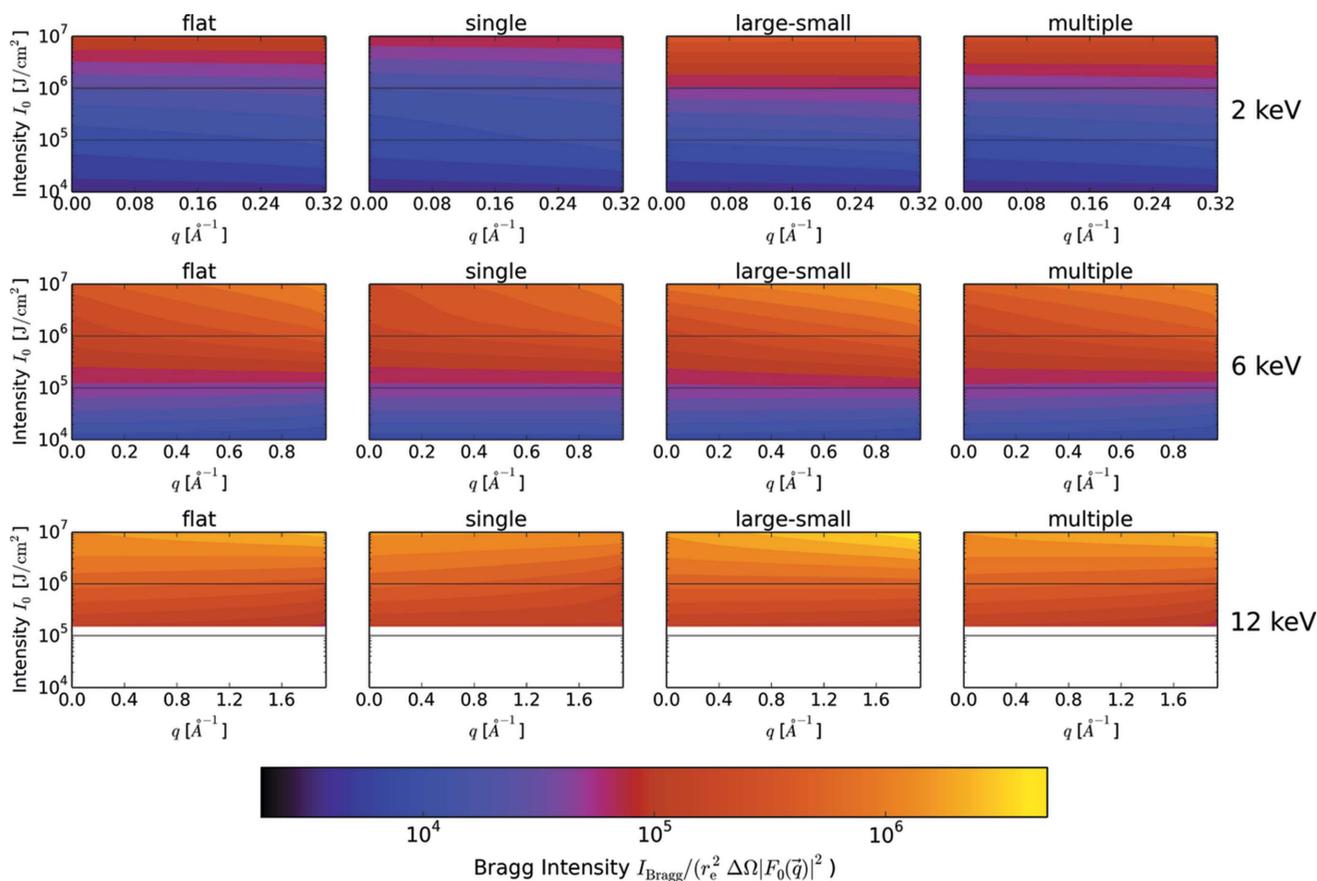

Figure 4
The accumulated Bragg signal as a function of resolution (q) and pulse intensity (I_0) for four different pulse shapes and photon energies (2, 6 and 12 keV). The Bragg intensity is calculated over the entire pulse length (100 fs) based on equation (8), without taking into account the term $r_e^2 \Delta\Omega |F_0(\mathbf{q})|^2$. For 12 keV, and low intensities, the expected ionization is very low. Here the plasma model is not valid. Top row: 2 keV. Middle row: 6 keV. Bottom row: 12 keV.

the signal depends on whether ionization or atomic displacement is the dominating damage mechanism (Caleman *et al.*, 2015). The absorption cross section at 2 keV is much (about a factor 140) larger than at 12 keV. For the same given intensity the onset of damage is faster, the subsequent damage is higher and the absolute Bragg intensities will be smaller at 2 keV compared with at 6 and 12 keV.

The *flat* pulse gives rise to a scattered signal that increases with the total intensity. Ionization causes a decay in the Bragg signal that is uneven for different scattering angles, dictated by the changes in the atomic scattering factor. Eventually displacement will lead to a total termination of the signal starting at the high scattering angles. As a result, the end of the pulse will not contribute to the signal (Fig. 3).

The *single* pulse will initially give a signal directly proportional to the intensity, which is lower than that of the other pulses. Ionization will be less extensive because of the lower flux, but this gain is countered by the lower number of scattering photons. Even if the temperature is slightly lower during the pulse than for the other pulse profiles, a plasma is still formed; meaning that eventually there is a displacement-caused termination of signal. For large scattering angles,

especially at high intensities, this cutoff will be reached *before* the arrival of the high-flux pulse peak, meaning most of the photons will be wasted. At low intensities the main peak in the pulse will contribute to the signal due to slow damage dynamics. However, in these cases the absolute intensity could be too low to achieve good quality data from small crystals.

The *large–small* pulse maximizes the initial flux and will give a proportionally stronger signal from the unaffected system. Ionization will happen faster because of the higher flux, but this is mostly countered by the higher number of photons in the beginning of the pulse. The resulting ion temperature will be higher but the signal termination due to displacement will occur *after* the majority of the photons have already scattered.

The *multiple* pulse produces a signal very similar to the flat pulse, but it builds up in steps.

To obtain a better qualitative picture of the differences between the pulses, the expected signal is compared with a flat-top pulse and is calculated as a relative signal,

$$J_{\mathrm{Bragg}}(\mathbf{q}, I_0, \lambda) = \frac{I_{\mathrm{Bragg}}(\mathbf{q}, I_0, \lambda)}{I_{\mathrm{Bragg}}(\mathbf{q}, I_0, \lambda)_{\mathrm{flat}}} - 1. \qquad (9)$$
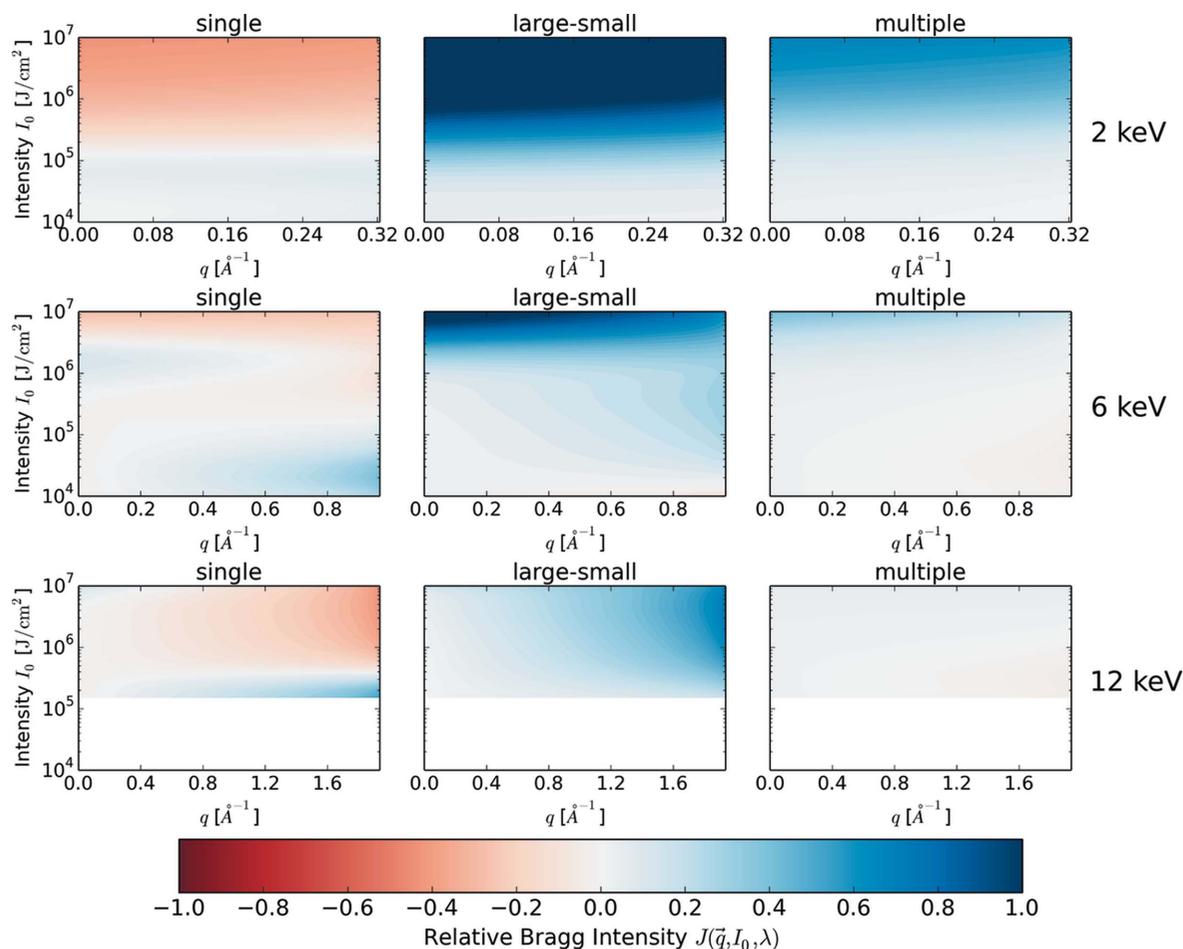

Figure 5
Relative Bragg intensities for three pulse shapes compared with a flat pulse (shown in columns), given at varying pulse intensities (*y*-axis), scattering vectors (*x*-axis) and photon energies (shown in rows), as defined in equation (9). A negative *J* (red scale) means that the pulse shape gives worse Bragg scattering compared with a flat pulse, and a positive *J* (blue scale) means that it gives better Bragg scattering compared with the flat case. Positive *J* values are shown up to 1. For 12 keV, and low intensities, the expected ionization is very low. Here the plasma model is not valid. Top row: 2 keV. Middle row: 6 keV. Bottom row: 12 keV.

The factor J_{Bragg} is displayed in Fig. 5 and was defined such that $J > 0$ means a better signal for the pulse shape compared with flat, and $J < 0$ implies a better Bragg signal for the flat pulse. Here it is clearly seen that radiation damage processes will vary with q differently, depending on pulse shape and total incoming intensity. Photon energy will influence what q regions are affected most by changing pulse shape. One example is the high q region at 12 keV, where the differences are very prominent compared with low q. For 6 keV, the difference between pulse shapes is evenly distributed for all q values.

From the pulses simulated here, the *single* pulse gives the lowest scattered signal (see Fig. 5). This is due to the initial low flux that interacts with the sample to start ionization and displacement, without giving a useful signal. When the high-flux main peak arrives, the sample is already disordered. The *large–small* pulse is clearly a better shape, due to the initial high flux of photons that probes the system before degradation. The general trend is that the *single* gives a lower signal than *flat*, whereas *large–small* gives a higher signal. *Multiple* is similar to *flat*.

Previous studies have used flat pulses (Caleman *et al.*, 2011*b*, 2012, 2015; Barty *et al.*, 2012; Chapman *et al.*, 2014). The simulations presented in the first column of Fig. 4 are consistent with those studies. In an experiment at the LCLS (Barty *et al.*, 2012) we noted that simulations using a flat-top pulse agree better with the experimental results if the simulated pulse was shorter (~40 fs) than the nominal value (~70 fs) inferred from the electron bunch length (70 fs). A similar observation was made independently (Young *et al.*, 2010). There are limitations in using a flat pulse to compare different scenarios, and it would be best to compare with experimental results. There are no such experiments to date; however, schemes to shape and control XFEL pulses are emerging.

Simulations have also been performed here with pulse lengths shorter than 100 fs, to study the damage dynamics on a 50 fs time scale. Similar temporal pulse profiles were investigated as described above (Fig. 1), only scaled down in time by a factor of two. The simulations with 50 fs pulses are presented in Fig. 6 and show in general the same trends as the 100 fs pulses. Overall it is expected that scaling down the pulse length will emphasize more the damage due to ionization
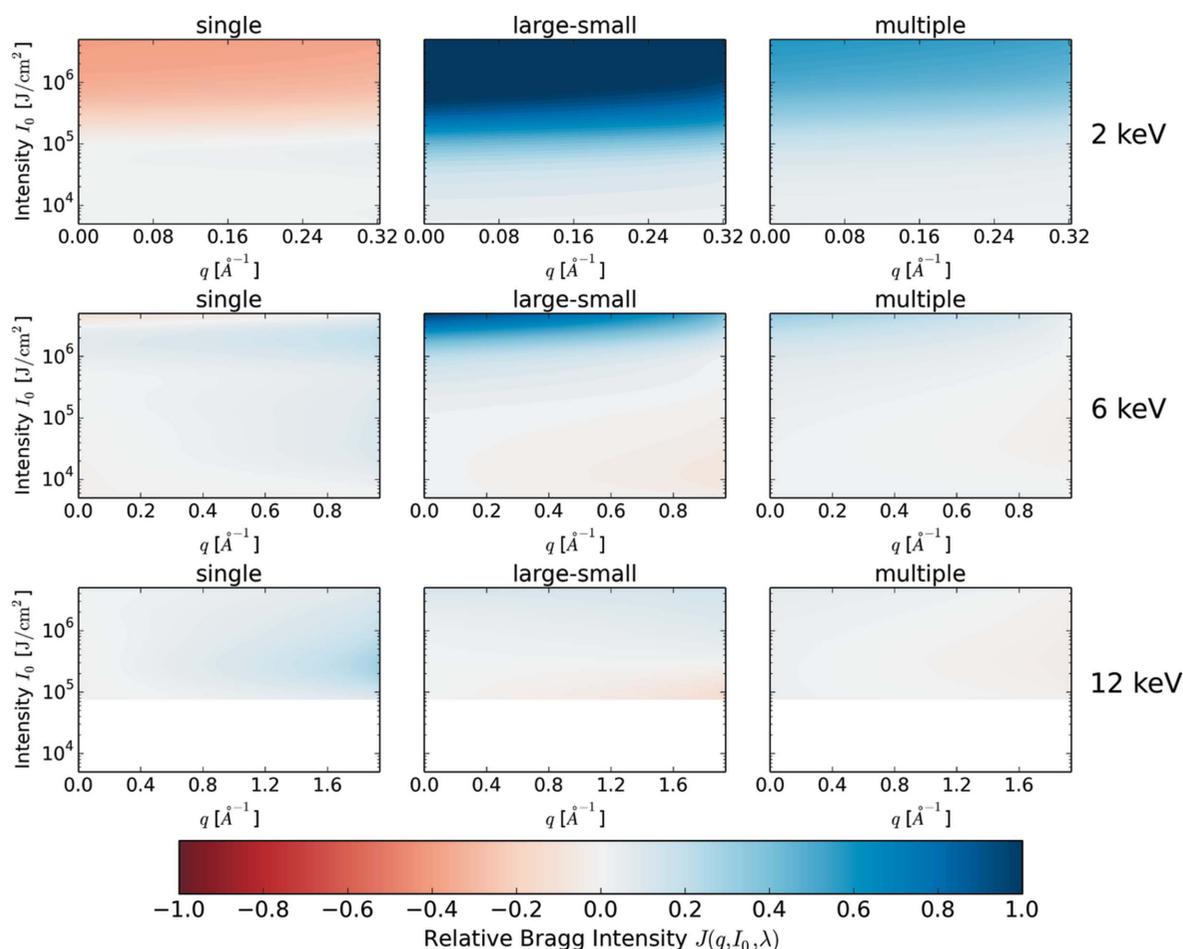

Figure 6
Relative Bragg intensities for three 50 fs pulse shapes compared with a flat pulse of 50 fs (shown in columns), given at varying pulse intensities (*y*-axis), scattering vectors (*x*-axis) and photon energies (shown in rows), as defined in equation (9). A negative J (red scale) means that the pulse shape gives worse Bragg scattering compared with a flat pulse, and a positive J (blue scale) means that it gives better Bragg scattering compared with the flat case. Positive J values are shown up to 1. These same trends can be observed for 50 fs pulses as in the 100 fs case. For 12 keV, and low intensities, the expected ionization is very low. Here the plasma model is not valid. Top row: 2 keV. Middle row: 6 keV. Bottom row: 12 keV.

compared with displacement; in other words the factor $k(q, t)$ will have a high relative importance compared with $h(q, t)$ for the same pulse intensity. When the pulses are shorter, the self-termination effect due to displacement will come at higher intensities when more of the photons have diffracted. As a consequence, the single pulse shape is not as inefficient as reported for the 100 fs case, basically because the single spike in this pulse arrives before the atoms have moved substantially.

4. Discussion

Any sample investigated with XFEL pulses will quickly turn into plasma. The transition between a solid state in equilibrium and a plasma state is relevant for the process of structural determination that relies on X-ray diffraction, and the question is how long the system will retain its structure and provide significant data. Recent work that bridges a gap between the solid approach and plasma (Vinko et al., 2014) suggests that a condensed matter approach for calculating ionization potentials gives a good description of the state of warm dense matter induced by femtosecond X-ray lasers, where the ions have barely started to move, have a high degree of ionization and are surrounded by hot free electrons. Thus, a non-LTE plasma approach seems like a good tool to consider for exploring the effects of an XFEL on a biological crystal, as it gives a proper description of the absorption cross sections during the exposure. On femtosecond time scales, this model captures the non-equilibrium between free electrons and ions, with hot electrons that slowly couple to the ion dynamics (time scales of 100 fs). These provide a more complete picture as the sample turns from solid to plasma. Beyond the ionization and sample dynamics, a more complete study of the effect of the pulse shape would require making calculations on the achievable resolution that can be obtained for a particular biomolecular material. This would require modeling the diffraction from a particular protein crystal and the process of structural determination from simulated data, and, although this is desirable, it is currently beyond the scope of this work.

To improve on the structural determination, the ideal would be to find a way to compensate for damage by rescaling the diffraction data. A Debye–Waller factor does not provide a good scaling for the diffraction data from an exploding crystal, and a better scaling can be obtained considering ultrafast heating in a plasma description (Barty et al., 2012). We have recently discussed the possibility of obtaining a formula for an analytical scaling (Caleman et al., 2015), which depends on a few parameters that describe both the ionization and displacement of the sample.

In the limit of single particle imaging, the approach presented here should still be valid, in the sense that radiation damage induces a degradation of the continuous scattering pattern. The motion-induced blurring of the scattering pattern has been studied earlier (Bergh et al., 2008b) and the signal at high resolution degrades with a factor analogous to $h(q, t)$. In the same context of single particles, a work that studied the ionization dynamics (Hau-Riege et al., 2007) suggested a way

to correct the diffraction pattern for ionization damage using a similar factor as $k(q, t)$, showing that such correction would lead to a reduced residual factor.

5. Summary and conclusions

The study of ultrafast dynamics of biomolecular systems is a growing research field pushed by the development of X-ray free-electron lasers, and there is a need for rigorous studies founded in quantum electrodynamics. Here a phenomenological approach has been used based on a plasma model to describe the (average) effects of X-ray interaction with biological matter on the expected diffracted signal. We have previously shown good agreement between experiments and a theoretical description of radiation damage in protein nanocrystals in free-electron laser crystallography (Barty et al., 2012). Building on that model, predictions could be made of the expected Bragg signals and radiation damage affecting the signal from varying flux during single X-ray pulses. Earlier studies (Bergh et al., 2008a; Caleman et al., 2011b) have indicated that pulse profiles have little impact on the ionization level at the end of the pulse. While this is still true (see Fig. 1), in a diffraction experiment the relevant variable is the number of photons that scatter, which could be quite sensitive to the ionization dynamics. In this paper we more closely consider the effects on Bragg diffraction. Even if the end state of the system may be the same or very similar between different pulse shapes (Fig. 1), the differences in intermediate states, and the weighting in importance in these states due to flux variation, will indeed impact the degradation of Bragg signal.

In addition to the effects from varying intensity and photon energy (Caleman et al., 2015), here features that depend only on the temporal pulse profile have been identified. It is found that resolution-dependent self-termination of the Bragg signal is driven by the atomic displacement and this effect is amplified further by differences in pulse shape. Ionization becomes an important factor at higher photon fluxes. Ionization is dependent on integrated flux, making the instanteneous electronic state very shape-dependent even if the end state is the same (especially at high intensities where the system is almost fully ionized). In the cases where the pulse starts with a low-flux tail (as in the single pulse), early photons might start the displacement process, making large parts of the pulse irrelevant.

Here it is found that flat- or Gaussian-shaped pulses are *not* necessarily optimal, unless it is guaranteed that the pulse is short and well defined so that no photons arrive early to start a plasma. The best pulse shape in terms of efficiency would instead be front-loaded, with an early strong maximum in flux. This assumes, however, that the signal-to-noise ratio will not drench the signal at the end of the pulse due to diffuse scattering. Conversely, pulses with low early fluxes give the lowest efficiency in terms of number of photons in the pulse that diffracts. When a plasma is created, the variation in the temperature gradient between different early fluxes will not cause extreme variations in atomic disorder growth. All pulses will start the radiation damage processes, but these early

photons in the low flux case might not contribute enough to the Bragg signal. If the maximum in flux happens after the self-termination caused by atomic displacement, most of the pulse will no longer contribute to diffraction but to diffuse scattering.

It has been shown that higher intensity will always increase the Bragg signal, but not necessarily with the same factor at the same scattering angles. This is controlled by electron configuration and atomic displacement, which both contribute to the state of the system when the photons arrive.

We find that a high pulse intensity will not necessarily lead to good diffraction patterns. The variation in quality between diffraction patterns obtained experimentally may be partially explained by varying pulse shape, even if the pulses have the same total intensity. In particular, the variation in dependence on θ may give incorrect electron densities. These variations should be taken into account when carrying out efficient data collection for high-resolution crystallography experiments at X-ray free-electron sources.

Efforts have been made to limit the variation in possible pulse shapes by employing seeded SASE techniques at the expense of lowered total intensity. The individual tradeoff between control and total fluence should be investigated very closely in light of the results presented here. A pulse-to-pulse shape monitoring system implemented at the beamline would be useful for various types of experiments, as it would be beneficial for understanding the ultrafast damage mechanims and it may even reduce the concern of sample damage during the process of structural determination.

References

Allaria, E. *et al.* (2012). *Nat. Photon.* **6**, 699–704.

Amann, J. *et al.* (2012). *Nat. Photon.* **6**, 693–698.

Andreasson, J. *et al.* (2011). *Phys. Rev. E*, **83**, 016403.

Aquila, A. *et al.* (2012). *Opt. Express*, **20**, 2706–2715.

Barends, T. R. M., Foucar, L., Botha, S., Doak, R. B., Shoeman, R. L., Nass, K., Koglin, J. E., Williams, G. J., Boutet, S. & Messerschmidt, M. (2013a). *Nature (London)*, **505**, 244–247.

Barends, T. R. M. *et al.* (2013b). *Acta Cryst.* D**69**, 838–842.

Barty, A. *et al.* (2012). *Nat. Photon.* **6**, 35–40.

Behrens, C., Decker, F.-J., Ding, Y., Dolgashev, V. A., Frisch, J., Huang, Z., Krejcik, P., Loos, H., Lutman, A., Maxwell, T. J., Turner, J., Wang, J., Wang, M.-H., Welch, J. & Wu, J. (2014). *Nat. Commun.* **5**, 3762.

Bergh, M., Huldt, G., Tîmneanu, N., Maia, F. R. N. C. & Hajdu, J. (2008b). *Q. Rev. Biophys.* **41**, 181.

Bergh, M., Tîmneanu, N., Hau-Riege, S. P. & Scott, H. A. (2008a). *Phys. Rev. E*, **77**, 026404.

Book, D. L. (2007). *Plasma Formulary*. Washington: Naval Research Laboratory.

Boutet, S. *et al.* (2012). *Science*, **337**, 362–364.

Caleman, C., Bergh, M., Scott, H. A., Spence, J. C. H., Chapman, H. N. & Tîmneanu, N. (2011b). *J. Mod. Opt.* **58**, 1486–1497.

Caleman, C., Huldt, G., Ortiz, C., Maia, F. R. N. C., Parak, F. G., Hajdu, J., van der Spoel, D., Chapman, H. N. & Tîmneanu, N. (2011a). *ACS Nano*, **5**, 139–146.

Caleman, C., Ortiz, C., Marklund, E., Bultmark, F., Gabrysch, M., Parak, F. G., Hajdu, J., Klintenberg, M. & Tîmneanu, N. (2009). *Europhys. Lett.* **85**, 18005; erratum: **88**, 29901.

Caleman, C., Timneanu, N., Martin, A. V., Aquila, A., Barty, A., Scott, H. A., White, T. A. & Chapman, H. N. (2015). *Opt. Express*, **23**, 1213–1231.

Caleman, C., Timneanu, N., Martin, A. V., White, T. A., Scott, H. A., Barty, A., Aquila, A. & Chapman, H. N. (2012). *Proc. SPIE*, **8504**, 85040H.

Chapman, H. N., Caleman, C. & Timneanu, N. (2014). *Philos. Trans. R. Soc. B*, **369**, 20130313.

Chapman, H. N. *et al.* (2011). *Nature (London)*, **470**, 73–78.

Cromer, D. T. & Mann, J. B. (1968). *Acta Cryst.* A**24**, 321–324.

Düsterer, S. *et al.* (2011). *New J. Phys.* **13**, 093024.

Emma, P. *et al.* (2010). *Nat. Photon.* **4**, 641–647.

Fromme, P. & Spence, J. C. (2011). *Curr. Opin. Struct. Biol.* **21**, 509–516.

Gericke, D., Murillo, M. S. & Schlanges, M. (2002). *Phys. Rev. E*, **65**, 036418.

Hattne, J. *et al.* (2014). *Nat. Methods*, **11**, 545–548.

Hau-Riege, S. P. (2007). *Phys. Rev. A*, **76**, 042511.

Hau-Riege, S. P. (2012). *Phys. Rev. Lett.* **108**, 217402.

Hau-Riege, S. P, London, R., Chapman, H., Szoke, A. & Timneanu, N. (2007). *Phys. Rev. Lett.* **98**, 198302.

Helml, W. *et al.* (2014). *Nat. Photon.* **8**, 950–957.

Hoener, M. *et al.* (2010). *Phys. Rev. Lett.* **104**, 253002.

Johansson, L. C. *et al.* (2012). *Nat. Methods*, **9**, 263–265.

Johansson, L. C. *et al.* (2013). *Nat. Commun.* **4**, 2911.

Kern, J. *et al.* (2013). *Science*, **340**, 491–495.

Kern, J. *et al.* (2014). *Nat. Commun.* **5**, 4371.

Koopmann, R. *et al.* (2012). *Nat. Methods*, **9**, 259–262.

Kupitz, C. *et al.* (2014). *Nature (London)*, **513**, 261–265.

Liu, W. *et al.* (2013). *Science*, **342**, 1521–1524.

Lomb, L. *et al.* (2011). *Phys. Rev. B*, **84**, 214111.

Mancuso, A. P. (2012). *Conceptual Design Report Scientific Instrument Single Particles, Clusters, and Biomolecules (SPB)*, XFEL.EU TR-2011-007. Europen XFEL, Hamburg, Germany.

Mancuso, A. P., Aquila, A., Borchers, G., Giewekemeyer, K. & Reimers, N. (2013). *Technical Design Report: Scientific Instrument SPB*, TR-2013-004. Europen XFEL, Hamburg, Germany.

Nantel, M., Ma, G., Gu, S., Côté, C. Y., Itatani, J. & Umstadter, D. (1998). *Phys. Rev. Lett.* **80**, 4442–4445.

Nass, K. *et al.* (2015). *J. Synchrotron Rad.* **22**, 225–238.

Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature (London)*, **406**, 752–757.

Redecke, L. *et al.* (2013). *Science*, **339**, 227–230.

Sawaya, M. R. *et al.* (2014). *Proc. Natl. Acad. Sci. USA*, **111**, 12769–12774.

Schneidmiller, E. A. & Yurkov, M. V. (2011). *Photon Beam Properties at the European XFEL*, XFEL.EU TN-2011-006. Europen XFEL, Hamburg, Germany.

Scott, H. A. (2001). *J. Quant. Spectrosc. Radiat. Transf.* **71**, 689.

Scott, H. A. & Mayle, R. W. (1994). *Appl. Phys. B*, **58**, 35–43.

Son, S.-K., Chapman, H. N. & Santra, R. (2011). *Phys. Rev. Lett.* **107**, 218102.

Spitzer, L. (1956). *Physics of Fully Ionized Gases.* New York: Interscience.

Stewart, J. C. & Pyatt, K. D. (1966). *Astrophys. J.* **144**, 1203.

Tiedtke, K. *et al.* (2009). *New J. Phys.* **11**, 023029.

Vinko, S. M., Ciricosta, O. & Wark, J. S. (2014). *Nat. Commun.* **5**, 3533.

Weierstall, U. *et al.* (2014). *Nat. Commun.* **5**, 3309.

Young, L. *et al.* (2010). *Nature (London)*, **466**, 56–61.

Zhao, Z. T. *et al.* (2012). *Nat. Photon.* **6**, 360–363.