

# Towards the theoretical limitations of X-ray nanocrystallography at high intensity: the validity of the effective-form-factor description

Malik Muhammad Abdullah,<sup>a,b,c\*</sup> Sang-Kil Son,<sup>a,b</sup> Zoltan Jurek<sup>a,b</sup> and Robin Santra<sup>a,b,c</sup>

Received 27 April 2018  
Accepted 10 August 2018

Edited by E. E. Lattman, University at Buffalo, USA

**Keywords:** XFELs; X-ray nanocrystallography; effective form factor; ionization; radiation damage.

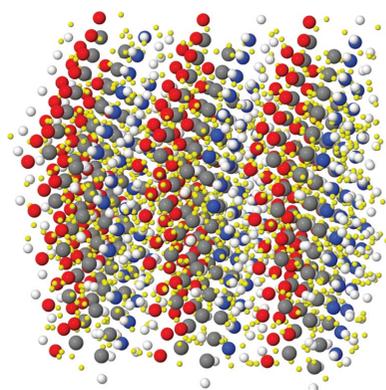
<sup>a</sup>Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany, <sup>b</sup>The Hamburg Centre for Ultrafast Imaging, Luruper Chaussee 149, 22761 Hamburg, Germany, and <sup>c</sup>Department of Physics, University of Hamburg, Jungiusstrasse 9, 20355 Hamburg, Germany. \*Correspondence e-mail: muhammad.abdullah.malik@desy.de

X-ray free-electron lasers (XFELs) broaden horizons in X-ray crystallography. Facilitated by the unprecedented high intensity and ultrashort duration of the XFEL pulses, they enable us to investigate the structure and dynamics of macromolecules with nano-sized crystals. A limitation is the extent of radiation damage in the nanocrystal target. A large degree of ionization initiated by the incident high-intensity XFEL pulse alters the scattering properties of the atoms leading to perturbed measured patterns. In this article, the effective-form-factor approximation applied to capture this phenomenon is discussed. Additionally, the importance of temporal configurational fluctuations at high intensities, shaping these quantities besides the average electron loss, is shown. An analysis regarding the applicability of the approach to targets consisting of several atomic species is made, both theoretically and *via* realistic radiation-damage simulations. It is concluded that, up to intensities relevant for XFEL-based nanocrystallography, the effective-form-factor description is sufficiently accurate. This work justifies treating measured scattering patterns using conventional structure-reconstruction algorithms.

## 1. Introduction

X-ray free-electron lasers (XFELs) (Berrah & Bucksbaum, 2014; Emma *et al.*, 2010) provide ultrashort X-ray pulses with unparalleled luminosity, producing ultrabright diffraction patterns, which enable atomic scale reconstruction of biomolecular structures. Unraveling the structural changes in XFEL-irradiated biomolecules has evoked great interest for decades (Neutze *et al.*, 2000; Chapman *et al.*, 2011; Boutet *et al.*, 2012; Redecke *et al.*, 2013). Recent advances in the technology of X-ray sources have opened new horizons in the field of time-resolved X-ray crystallography. According to a novel experimental scheme called serial femtosecond crystallography (SFX) (Chapman, 2015), biomolecular nanocrystals are individually illuminated by one XFEL pulse each and the scattering patterns are recorded. A set of such patterns is then used for the determination of the electron density. However, an XFEL pulse also induces radiation damage in the targeted sample.

Numerous theoretical and numerical studies have been carried out in order to investigate the effects of radiation damage in single instances and nanocrystals of biological molecules (*e.g.* proteins) and also in non-biological systems



(Murphy *et al.*, 2014; Abdullah *et al.*, 2016, 2017; Neutze, 2014; Curwood *et al.*, 2013; Chapman *et al.*, 2014; Ziaja *et al.*, 2015; Ho *et al.*, 2016). X-ray irradiation causes electronic damage, which affects the atomic form factors (Quiney & Nugent, 2011; Son, Young & Santra, 2011) and may also result in atomic displacements on longer time-scales, leading to the annihilation of the Bragg diffraction spots (Barty *et al.*, 2012).

Ionization reduces the atomic form factors and therefore the scattered signal and thus limits the achievable resolution. In order to distinguish the atomic species of a nanocrystal from reconstruction, the atom type may need to be assigned on the basis of the electron density. Atomic species like carbon, nitrogen and oxygen, which exist in large numbers in biomolecular crystals and have similar atomic numbers, may be difficult to distinguish because of electronic damage. Furthermore, as the patterns accumulate signal from a sample undergoing radiation induced changes, they are no longer in a strict Fourier-transformation relationship with the electron density. Therefore, theoretical investigations are essential in understanding the formation and information content of such non-ideal patterns.

In one approach, it is considered that scattering possesses the statistical characteristics of a partially coherent diffraction pattern (Quiney & Nugent, 2011; Lorenz *et al.*, 2012; Hau-Riege *et al.*, 2007), whereas in the case of molecules containing a single atomic species and assuming a simple linear scaling relation between charge state and atomic form factors, the scattering pattern can be written as a coherent sum based on effective electron densities (Hau-Riege *et al.*, 2007). Recently, in the case of large biomolecular species, a simple approximation of using effective form factors defined by the square root of the time-averaged square of time-dependent scattering factors has also been employed (Lunin *et al.*, 2013; Lunin *et al.*, 2015).

In this article, we will redefine the effective form-factors, emphasizing the implications for the interpretation of the scattering patterns. A time-integrated pattern does not correspond to a static electron density *via* a Fourier transform in a mathematically rigorous manner. It is formed by an incoherent sum of non-identical, individually coherent patterns. Therefore, it is not straightforward that conventional pattern-processing schemes can be expected to work. However, if the temporal-variance-aided effective-form-factor description is proven to be accurate under relevant damage conditions, it also ensures that the time-integrated pattern can be treated as a coherent pattern to good accuracy, and image processing algorithms can be expected to converge and deliver a solution. By using a realistic radiation-damage model including both atomic and environmental effects, we theoretically investigate the limitations of the simple effective-form-factor concept on the example of a glycine (C<sub>2</sub>H<sub>5</sub>NO<sub>2</sub>) organic nanocrystal. By calculating Bragg intensities we analyze the contribution of the temporal variance and the threshold pulse intensity up to which the constructed effective form factors are valid to describe the non-ideal patterns, thus allowing for the use of conventional crystallography processing methods.

## 2. Theoretical methods

The scattering pattern of a crystal affected by severe radiation damage at high X-ray intensity is calculated by an incoherent summation over all possible electronic and nuclear configurations weighted by the corresponding probabilities of occurrence at a given time and then accumulated over the whole X-ray pulse. When a crystal is exposed to a high-intensity X-ray beam with fluence  $\mathcal{F}$  and photon energy  $\omega$  (we employ atomic units), the scattering intensity at the momentum transfer  $\mathbf{Q}$  is given by

$$\frac{dI(\mathbf{Q}, \mathcal{F}, \omega)}{d\Omega} = \mathcal{F}C(\Omega) \int_{-\infty}^{\infty} dt g(t) \sum_{\hat{I}, \hat{R}} P_{\hat{I}, \hat{R}}(\mathcal{F}, \omega, t) \times \left| \sum_X \sum_{j=1}^{N_X} f_{X, I_j^X}(\mathbf{Q}, \omega) \exp(i\mathbf{Q} \cdot \mathbf{R}_j^X) \right|^2, \quad (1)$$

where  $X$  indicates the atomic species and  $j$  represents the atomic index of that species.  $C(\Omega)$  represents a factor depending on the polarization of the X-ray pulse, while  $g(t)$  is the normalized temporal envelope of the pulse. We assume a uniform fluence distribution within the irradiated part of the crystal (Abdullah *et al.*, 2016).  $\hat{I} = \{I_j^X\}$  is the global electronic configuration of the crystal, which is given by specifying the electronic configuration  $I_j^X$  of all individual atoms, and  $\hat{R} = \{\mathbf{R}_j^X\}$  is the global nuclear configuration of all atomic positions  $\mathbf{R}_j^X$  in the nanocrystal. The atomic form factor differs for different atomic species  $X$  and different electronic configurations, so it is given by  $f_{X, I_j^X}$ .  $P_{\hat{I}, \hat{R}}$  is the time-dependent probability of  $\hat{I}$  and  $\hat{R}$ , which also depends on  $\mathcal{F}$  and  $\omega$ . Note that it is critical to obtain the time evolution of  $P_{\hat{I}, \hat{R}}$  in order to evaluate the scattering intensity of equation (1).

In contrast, the scattering intensity for an undamaged sample is calculated simply by using a coherent sum (the dependence on  $\mathbf{Q}$ ,  $\mathcal{F}$  and  $\omega$  is omitted for the sake of convenience),

$$\frac{dI}{d\Omega} = \mathcal{F}C(\Omega) \left| \sum_X f_X^0 \left[ \sum_{j=1}^{N_X} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^X) \right] \right|^2, \quad (2)$$

where  $f_X^0$  is the atomic form factor of the atomic species  $X$  in the neutral ground state. Here we consider nonresonant X-ray scattering only.

Our goal is to approximate the scattering intensity for XFEL-irradiated crystals by using a simple coherent form as in equation (2). The simplest solution can be obtained by replacing  $f_X^0$  with the time-averaged atomic form factor,

$$\bar{f}_X = \int_{-\infty}^{\infty} dt g(t) \tilde{f}_X(t), \quad (3)$$

where  $\tilde{f}_X(t) = \sum_{I_X} P_{I_X}(t) f_{I_X}$  is the time-dependent atomic form factor during the X-ray pulse and  $f_{I_X}$  is the atomic form factor of the  $I_X$ th electronic configuration of the given atomic species  $X$ .  $P_{I_X}(t)$  is the configurational population at a given time  $t$ , which was considered within the independent-atom

model (Son, Chapman & Santra, 2011, 2013). The time-averaged atomic form factor  $\bar{f}_X$  at  $\mathbf{Q} = 0$  is typically interpreted as the effective charge during the X-ray pulse for the given atomic species. The effective charge (time-averaged electron loss) is enhanced as the intensity increases (see Fig. 2) because of ionization dynamics, thus reducing the time-averaged form factor.

On the other hand, it has been suggested that the time-averaged atomic form factor is not enough to describe the scattering intensity in the case of high-intensity X-ray fields (Son, Chapman & Santra, 2011, 2013; Galli, Son, Barends *et al.*, 2015). Since the time-dependent atomic form factor varies dramatically during an intense X-ray pulse, the temporal variance needs to be taken into account (Son, Chapman & Santra, 2013). For a single atomic species, it is trivial to derive the following effective form factor from the generalized Karle-Hendrickson equation (Son, Chapman & Santra, 2011, 2013):

$$f_X^{\text{eff}} = \left( |\bar{f}_X|^2 + V_X^{\text{time}} \right)^{1/2}, \quad (4)$$

where  $V_X^{\text{time}} = [\int_{-\infty}^{\infty} dt g(t) |\tilde{f}_X(t)|^2] - |\int_{-\infty}^{\infty} dt g(t) \tilde{f}_X(t)|^2$ . If the XFEL-irradiated crystal consists of more than one atomic species, it can be shown that the scattering intensity may be approximated by a coherent sum as in equation (2), with the effective atomic form factors defined in equation (4) (see Appendix A for details). With this definition, the distinction between the effective form factor and that derived from the effective charge can be clearly seen. Since  $V_X^{\text{time}} > 0$ , the time-averaged form factor  $\bar{f}_X$  always underestimates the effective form factor  $f_X^{\text{eff}}$ . We will present a detailed numerical analysis for those form factors in the following section, based on realistic radiation-damage simulations of nanocrystals irradiated by intense X-ray pulses. Note that the form of equation (4) is equivalent to that proposed in Lunin *et al.* (2015):  $f_X^{\text{eff}} = [\int_{-\infty}^{\infty} dt g(t) |\tilde{f}_X(t)|^2]^{1/2}$ .

### 3. Numerical analysis

#### 3.1. Simulation methods

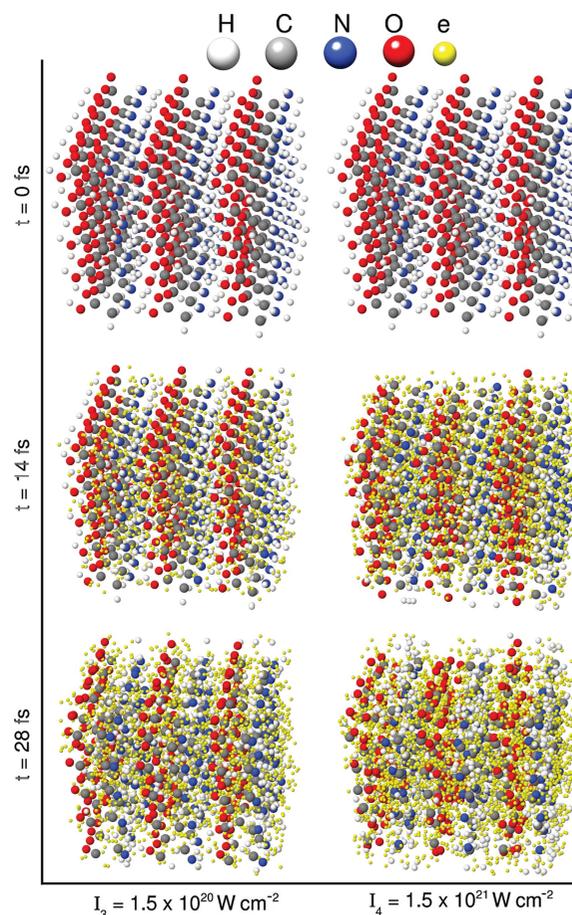
In order to perform a simulation of a nanocrystal exposed to an intense X-ray pulse, we subdivide the nanocrystal into supercells and simulate the ionization and nuclear dynamics for the supercells using *XMDYN* (Jurek *et al.*, 2016; Murphy *et al.*, 2014; Tachibana *et al.*, 2015), applying periodic boundary conditions. *XMDYN* is a radiation-damage-simulation tool that takes into account the inner-shell processes, such as photoionization, Auger and fluorescent relaxation, as well as phenomena caused by the environment, such as collisional ionization, recombination and dynamics driven by Coulomb interaction between charged particles. This supercell approach of *XMDYN* has been applied before to bulk systems (Abdullah *et al.*, 2016, 2017). To construct a scattering pattern from the nanocrystal, we employ the code *XSINC* (Abdullah *et al.*, 2016).

In our investigation, for each Bragg reflection, *XSINC* analyzes the scattering intensity in equation (1) with  $P_{I,\hat{R}}(t)$  obtained from realistic simulations of *XMDYN*, including

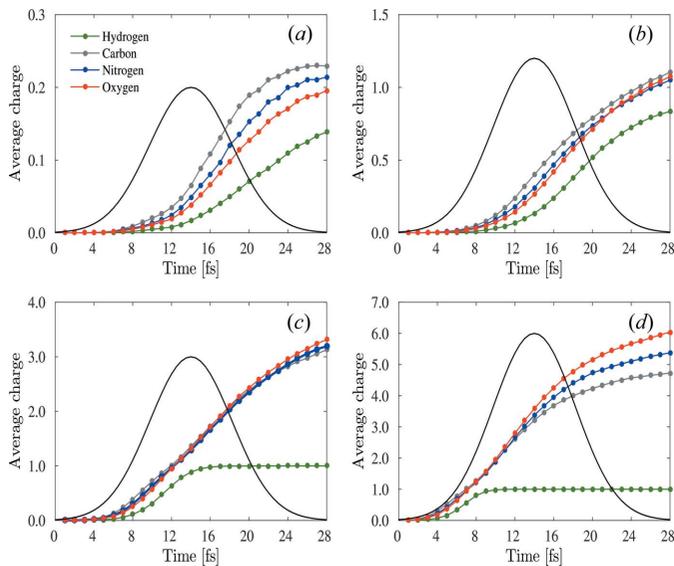
both impact ionization (Bekx *et al.*, 2018) and recombination, which are critical in a dense matter environment (Abdullah *et al.*, 2016). With  $P_{I,\hat{R}}(t)$  derived from  $P_{I,\hat{R}}(t)$ , the time-averaged atomic form factor  $\bar{f}_X$  in equation (3) and the effective atomic form factor  $f_X^{\text{eff}}$  in equation (4) are calculated using *XSINC*.

#### 3.2. Results

In our analysis, we considered a nanocrystal of the amino acid glycine. The virtually assembled crystallographic unit cell is orthorhombic with cell parameters of  $U_x = 5.7248$  Å,  $U_y = 2.986$  Å and  $U_z = 1.912$  Å, containing one molecule. In our simulation, we constructed a supercell with dimensions of  $S_x = 17.174$  Å,  $S_y = 14.93$  Å and  $S_z = 13.384$  Å, containing 105 glycine molecules. For scattering pattern calculations, we considered a nanocrystal consisting of  $29 \times 33 \times 37$  supercells. We used a photon energy of 10 keV and four different X-ray peak intensities:  $I_1 = 1.5 \times 10^{18}$  W cm<sup>-2</sup>,  $I_2 = 1.5 \times 10^{19}$  W cm<sup>-2</sup>,  $I_3 = 1.5 \times 10^{20}$  W cm<sup>-2</sup> and  $I_4 = 1.5 \times 10^{21}$  W cm<sup>-2</sup>. The temporal pulse envelope is Gaussian with 10 fs full width at half-maximum (FWHM) and we assumed spatially uniform irradiation. For each peak intensity, 150 *XMDYN* trajectories were calculated. Fig. 1



**Figure 1** Real-space snapshots of ionization dynamics of a supercell comprising 105 molecules of glycine. The photon energy is 10 keV; the peak intensities are  $I_3 = 1.5 \times 10^{20}$  and  $I_4 = 1.5 \times 10^{21}$  W cm<sup>-2</sup>. The temporal pulse envelope is Gaussian with 10 fs FWHM. The pulse is centered at  $t = 14$  fs.

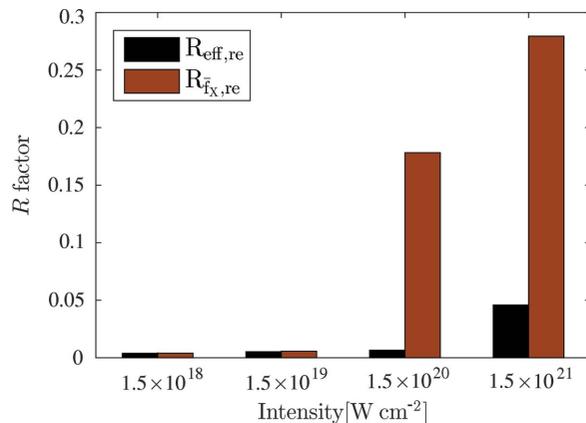


**Figure 2** Average charge as a function of time at the intensity of (a)  $I_1 = 1.5 \times 10^{18}$ , (b)  $I_2 = 1.5 \times 10^{19}$ , (c)  $I_3 = 1.5 \times 10^{20}$  and (d)  $I_4 = 1.5 \times 10^{21} \text{ W cm}^{-2}$ . The black curve represents the temporal Gaussian envelope of 10 fs FWHM, centered at  $t = 14$  fs.

shows real-space snapshots of the atoms in a single supercell undergoing ionization as a function of time, for the intensities  $I_3$  and  $I_4$ . It can be seen from the increasing number of ejected electrons that the structure is substantially ionized by the end of the pulse for the  $I_4$  case. However, during the short time duration of the pulse, the atomic displacements did not exceed 0.15 Å for C, N, and O, being well below the resolution. Hence the Bragg reflections are not affected by the atomic movement. The free-electron contribution to the coherent signal is also negligible because of the fairly uniform average spatial distribution.

Fig. 2 shows the time evolution of the charge for different atomic species at different intensities. For the lowest intensity ( $I_1$ ), almost all the species remain neutral (charges  $< +0.3$ ) after irradiation, whereas for the highest intensity ( $I_4$ ), carbon, nitrogen and oxygen are ionized up to charge states of +4.7, +5.2 and +6.0, respectively. To saturate single-photon absorption for light atoms (carbon, nitrogen and oxygen) at 10 keV, the intensity at 10 fs FWHM must be larger than  $10^{21} \text{ W cm}^{-2}$ . Therefore, X-ray multiphoton ionization does not play a significant role in the intensity regime under consideration, except for the highest intensity. The drastic changes in the charge states shown in the high-intensity cases in Fig. 2 are mainly caused by electron-impact ionization (Abdullah *et al.*, 2016), resulting in severe radiation damage.

The accuracy of the effective-form-factor approximation is verified by the crystallographic  $R$  factor, which is widely used as a measure of the agreement between calculated patterns based on a crystal-

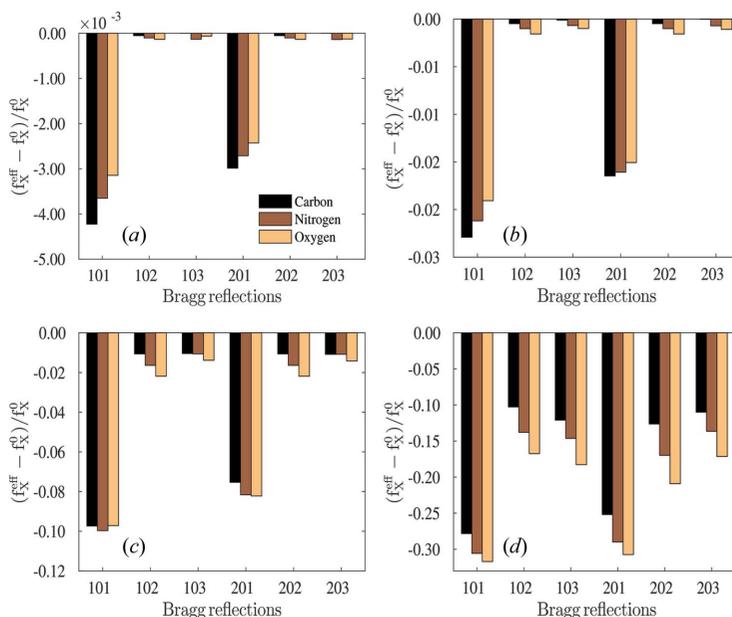


**Figure 3** Crystallographic  $R$  factor in two different cases as a function of intensity. The black bars represent  $R_{\text{eff, re}}$ , the brown bars represent  $R_{f_X, \text{re}}$ .

lographic model and the experimental patterns. In Fig. 3 we compare the goodness of fit of two different approximations using the  $R$  factor at several intensities.  $R_{\text{eff, re}}$  is defined by

$$R_{\text{eff, re}} = \frac{\sum_{\mathbf{Q}} | [I_{\text{eff}}(\mathbf{Q})]^{1/2} - [I_{\text{re}}(\mathbf{Q})]^{1/2} |}{\sum_{\mathbf{Q}} [I_{\text{re}}(\mathbf{Q})]^{1/2}}, \quad (5)$$

where the real intensities  $I_{\text{re}}(\mathbf{Q})$  are calculated from the incoherent sum, with full dynamics calculations, in equation (1) and  $I_{\text{eff}}(\mathbf{Q})$  is calculated from the coherent sum in equation (2) by replacing  $f_X^0$  with the effective form factors  $f_X^{\text{eff}}$  [equation (4)]. Similarly,  $I_{f_X}(\mathbf{Q})$  is obtained by replacing  $f_X^0$  with  $\tilde{f}_X$ . Then,  $R_{f_X, \text{re}}$  is calculated from  $I_{f_X}(\mathbf{Q})$  and  $I_{\text{re}}(\mathbf{Q})$ . The crystallographic  $R$  factor is calculated using reflections up to



**Figure 4** Relative differences of the effective form factor ( $f_X^{\text{eff}}$ ) compared with the ideal form factor ( $f_X^0$ ) for different atomic species. The peak intensity for each panel is the same as used in Fig. 2.

Miller index (666) which corresponds to the resolution of 1.58 Å. The  $R$ -factor value required for successful structural determination at atomic resolution is suggested to be  $R \leq 0.15$  as a rule of thumb (Neutze *et al.*, 2000), supported also by statistical analysis of deposited solutions in the Protein Data Bank (PDB) (Urzhumtsev *et al.*, 2009). The minimum possible value of  $R$  factor is zero, indicating perfect agreement between the considered cases. It can be seen that for the highest intensity ( $I_4$ ),  $R_{\text{eff, re}}$  is still only about 0.05, which indicates good agreement between  $I_{\text{eff}}$  and  $I_{\text{re}}$ . Hence, the coherent sum with the effective atomic form factors used here can describe the radiation damage in a nanocrystal even for the highest intensity ( $I_4$ ). On the other hand,  $R_{\bar{f}_X, \text{re}}$  increases much more rapidly as a function of the intensity, indicating that the time-averaged atomic form factor  $\bar{f}_X$  is a poor choice when attempting to approximate the non-ideal pattern in terms of a coherent pattern;  $f_X^{\text{eff}}$  [equation (4)] provides a much better fit, particularly at the highest intensities considered here.

To further explore the changes caused by radiation-damage dynamics using the effective form factors, we analyzed the relative difference between the effective and ideal (undamaged) form factors,  $(f_X^{\text{eff}} - f_X^0)/f_X^0$ , as shown in Fig. 4. The effective atomic form factors are always reduced because of the radiation damage, so all plots in Fig. 4 are negative. The relative differences are almost negligible at low intensities (see Figs. 4*a* and 4*b*), but no longer at high intensities; the maximum difference is about 10% in Fig. 4(*c*) and 30% in Fig. 4(*d*). Moreover, these relative differences are not constant for different Bragg reflections and different atomic species. For example, at the lowest intensity in Fig. 4(*a*), the effective form factors of carbon at the (101) and (201) reflections are more reduced than those of oxygen, even though the percentage is very small. At the highest intensity in Fig. 4(*d*), the  $f_X^{\text{eff}}$  of oxygen are more reduced than those of carbon, and the relative differences fluctuate between 10% and 30% for different Bragg reflections. Hence, the effective form factors cannot, in general, be obtained by multiplying the standard form factors  $f_X^0$  by a single uniform scaling factor.

#### 4. Conclusions

In summary, we have discussed the generalization of the effective-form-factor approximation applied to describe scattering patterns from XFEL-irradiated samples consisting of multiple atomic species. We have shown that these quantities are mainly shaped by the average electron loss caused by stochastic ionization events and dynamical configurational fluctuations. We have demonstrated *via* realistic numerical simulations that the role of the latter contribution becomes more prominent with increasing X-ray intensity. Still, up to intensities relevant for XFELs, the effective-form-factor description is acceptable, also implying that conventional structure-reconstruction algorithms dealing with purely coherent scattering signals can be expected to work in this intensity regime as well.

#### APPENDIX A

##### Effective atomic form factor

First, let us define the effective form factor in equation (4). In a similar fashion to Galli, Son, Barends *et al.* (2015), the effective atomic form factor is defined by the square root of the scattering intensity given by only one atomic species  $X$  after averaging over time and configurations:

$$f_X^{\text{eff}} = \left( \frac{\mathcal{F}C(\Omega) \int_{-\infty}^{\infty} dt g(t) \sum_{\hat{i}, \hat{R}} P_{\hat{i}, \hat{R}}(\mathcal{F}, \omega, t) \left| \sum_{j=1}^{N_X} f_{X, I_j^X} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^X) \right|^2}{\mathcal{F}C(\Omega) \int_{-\infty}^{\infty} dt g(t) \sum_{\hat{i}, \hat{R}} P_{\hat{i}, \hat{R}}(\mathcal{F}, \omega, t) \left| \sum_{j=1}^{N_X} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^X) \right|^2} \right)^{1/2}. \quad (6)$$

We assume that the nanocrystal is exposed to a homogeneous fluence distribution (Abdullah *et al.*, 2016). Assuming that no nuclear motions are involved during the short pulse duration and radiation-damage dynamics of individual atoms happen individually, the global population is given by the product of the individual atomic populations with the corresponding electronic configuration,

$$P_{\hat{i}, \hat{R}}(\mathcal{F}, \omega, t) = \prod_X \prod_{j=1}^{N_X} P_{I_j^X}(\mathcal{F}, \omega, t). \quad (7)$$

We also assume that the dynamical profiles of individual atomic populations are similar to each other for a given atomic species,  $P_{I_j^X}(\mathcal{F}, \omega, t) \simeq P_{I_X}(\mathcal{F}, \omega, t)$ . Then the effective atomic form factor goes over into

$$\begin{aligned} f_X^{\text{eff}} &= \left( \int_{-\infty}^{\infty} dt g(t) |\tilde{f}_X(t)|^2 + \frac{N_X \left[ \sum_{I_X} \bar{P}_{I_X} |f_{I_X}|^2 - \int_{-\infty}^{\infty} dt g(t) |\tilde{f}_X(t)|^2 \right]}{\left| \sum_{j=1}^{N_X} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^X) \right|^2} \right)^{1/2} \\ &\simeq \left[ \int_{-\infty}^{\infty} dt g(t) |\tilde{f}_X(t)|^2 \right]^{1/2} \\ &= \left( |\bar{f}_X|^2 + V_X^{\text{time}} \right)^{1/2}, \end{aligned} \quad (8)$$

where  $\bar{P}_{I_X} = \int_{-\infty}^{\infty} dt g(t) P_{I_X}(t)$  (the dependence of  $\mathcal{F}$ ,  $\mathbf{Q}$  and  $\omega$  is omitted). A similar analysis was performed in Son, Chapman & Santra (2011). The term within the brackets in equation (8) diminishes when  $N_X$  becomes large, because at Bragg peaks  $\left| \sum_{j=1}^{N_X} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^X) \right|^2 \propto (N_X)^2$ . It is worthwhile noting that this definition of the effective atomic form factor is directly connected to the MAD coefficient  $\tilde{a}_X$  in Son, Chapman & Santra (2011):  $f_X^{\text{eff}} = f_X^0 (\tilde{a}_X)^{1/2}$ .

Next, we demonstrate how the scattering intensity may be approximated by a coherent sum using the effective form factors in equation (2). We start from equation (2) in Appendix A in Galli, Son, White *et al.* (2015). For simplicity, we consider only two atomic species,  $A$  and  $B$  (an extension to many atomic species is straightforward):

$$\begin{aligned} \frac{dI}{d\Omega} = & \mathcal{FC}(\Omega) \int_{-\infty}^{\infty} dt g(t) \sum_i P_i(\mathcal{F}, \omega, t) \\ & \times \left| \sum_{j=1}^{N_A} f_{A,j^A}(\mathbf{Q}, \omega) \exp(i\mathbf{Q} \cdot \mathbf{R}_j^A) \right. \\ & \left. + \sum_{k=1}^{N_B} f_{B,k^B}(\mathbf{Q}, \omega) \exp(i\mathbf{Q} \cdot \mathbf{R}_k^B) \right|^2. \end{aligned} \quad (9)$$

The above scattering-intensity expression can be written as the extended Karle–Hendrickson equation, following the expressions in Galli, Son, White *et al.* (2015),

$$\begin{aligned} \frac{dI}{d\Omega} = & \mathcal{FC}(\Omega) \left[ |F_A^0|^2 \tilde{a}_A + N_A |f_A^0|^2 (a_A - \tilde{a}_A) + |F_B^0|^2 \tilde{a}_B \right. \\ & + N_B |f_B^0|^2 (a_B - \tilde{a}_B) + |F_A^0| |F_B^0| B_{AB} \cos \Delta\varphi_{AB}^0 \\ & \left. + |F_A^0| |F_B^0| C_{AB} \sin \Delta\varphi_{AB}^0 \right]. \end{aligned} \quad (10)$$

When anomalous scattering contributions are small enough (for example, light atoms at hard X-rays as demonstrated in the main text), the term with  $C_{AB}$  may be neglected, while the term with  $B_{AB}$  still needs to be evaluated because of the coupling of  $\tilde{f}_A(t)$  and  $\tilde{f}_B(t)$ . Here the molecular form factor is defined by

$$F_X^0 = f_X^0 \sum_{j=1}^{N_X} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^X) = |F_X^0| \exp(i\varphi_X^0), \quad (11)$$

and the phase difference is  $\Delta\varphi_{AB}^0 = \varphi_A^0 - \varphi_B^0$ . Note that the dependence on  $\mathbf{Q}$  and  $\omega$  is omitted for simplicity. The atom-specific MAD coefficients are given by

$$a_X = \frac{1}{(f_X^0)^2} \sum_{I_X} \bar{P}_{I_X} |f_{I_X}|^2, \quad (12)$$

$$\tilde{a}_X = \frac{1}{(f_X^0)^2} \int_{-\infty}^{\infty} dt g(t) |\tilde{f}_X(t)|^2, \quad (13)$$

and the biatom-specific MAD coefficient  $B_{AB}$  is given by

$$B_{AB} = \frac{2}{f_A^0 f_B^0} \int_{-\infty}^{\infty} dt g(t) \tilde{f}_A(t) \tilde{f}_B(t). \quad (14)$$

After plugging the effective form factor  $f_X^{\text{eff}}$  into equation (10), the scattering intensity is recast as

$$\begin{aligned} \frac{dI}{d\Omega} = & \mathcal{FC}(\Omega) \left[ \left| \frac{F_A^0}{f_A^0} f_A^{\text{eff}} + \frac{F_B^0}{f_B^0} f_B^{\text{eff}} \right|^2 + N_A \bar{V}_A^{\text{config}} + N_B \bar{V}_B^{\text{config}} \right. \\ & \left. + |F_A^0| |F_B^0| \left( B_{AB} - \frac{2f_A^{\text{eff}} f_B^{\text{eff}}}{f_A^0 f_B^0} \right) \cos \Delta\varphi_{AB}^0 \right], \end{aligned} \quad (15)$$

where  $\bar{V}_X^{\text{config}} = (f_X^0)^2 (a_X - \tilde{a}_X)$  (Son, Chapman & Santra, 2013). Under Bragg conditions, the terms with  $\bar{V}_X^{\text{config}}$  are smaller than others as  $N_X$  becomes larger. In addition, let us assume that time profiles of the dynamical behavior of different atomic species are proportional to  $h(t)$ , such that

$\tilde{f}_A(t) = f_A^0 h(t)$  and  $\tilde{f}_B(t) = f_B^0 h(t)$ . The factor  $[B_{AB} - (2f_A^{\text{eff}} f_B^{\text{eff}})/(f_A^0 f_B^0)]$  then becomes

$$\begin{aligned} B_{AB} - \frac{2f_A^{\text{eff}} f_B^{\text{eff}}}{f_A^0 f_B^0} = & \frac{2}{f_A^0 f_B^0} \int_{-\infty}^{\infty} dt g(t) \tilde{f}_A(t) \tilde{f}_B(t) \\ & - \frac{2}{f_A^0 f_B^0} \left[ \int_{-\infty}^{\infty} dt g(t) |\tilde{f}_A(t)|^2 \right]^{1/2} \left[ \int_{-\infty}^{\infty} dt g(t) |\tilde{f}_B(t)|^2 \right]^{1/2} \\ = & 2 \int_{-\infty}^{\infty} dt g(t) [h(t)]^2 - 2 \left\{ \left[ \int_{-\infty}^{\infty} dt g(t) |h(t)|^2 \right]^2 \right\}^{1/2} \\ = & 0. \end{aligned} \quad (16)$$

Consequently, if we assume small anomalous scattering signals and similar dynamical behavior for different atomic species, and neglect small  $N_X \bar{V}_X^{\text{config}}$  terms, then the scattering intensity may be expressed as the conventional coherent sum,

$$\begin{aligned} \frac{dI}{d\Omega} = & \mathcal{FC}(\Omega) \left| \frac{F_A^0}{f_A^0} f_A^{\text{eff}} + \frac{F_B^0}{f_B^0} f_B^{\text{eff}} \right|^2 \\ = & \mathcal{FC}(\Omega) \left| f_A^{\text{eff}} \left[ \sum_{j=1}^{N_A} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^A) \right] \right. \\ & \left. + f_B^{\text{eff}} \left[ \sum_{j=1}^{N_B} \exp(i\mathbf{Q} \cdot \mathbf{R}_j^B) \right] \right|^2. \end{aligned} \quad (17)$$

## Acknowledgements

We thank Thomas White and John Jasper Bekx for fruitful discussions.

## Funding information

This work has been supported by the excellence cluster ‘The Hamburg Centre for Ultrafast Imaging (CUI): Structure, Dynamics and Control of Matter at the Atomic Scale’ of the Deutsche Forschungsgemeinschaft.

## References

- Abdullah, M. M., Anurag, Jurek, Z., Son, S.-K. & Santra, R. (2017). *Phys. Rev. E*, **96**, 023205.
- Abdullah, M. M., Jurek, Z., Son, S.-K. & Santra, R. (2016). *Struct. Dyn.* **3**, 054101.
- Barty, A., Caleman, C., Aquila, A., Nicusor, T., Lomb, L., White, T. A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T. R. M., Barthelmess, M., Bogan, M. J., Bostedt, C., Bozek, J. D., Coffee, R., Coppola, N., Davidsson, J., DePonte, D. P., Bruce Doak, R., Ekeberg, T., Elser, V., Epp, S. W., Erk, B., Fleckenstein, H., Foucar, L., Fromme, P., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C. Y., Hartmann, R., Hartmann, A. A., Hauser, G., Hirsemann, H., Holl, P. A., Hunter, M. S., Johansson, L., Kassemeyer, S., Kimmel, N., Kirian, R. A., Liang, M., Maia, F. R. N. C., Malmerberg, E., Marchesini, S., Martin, A. V., Nass, K., Neutze, R., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Scott, H., Schlichting, I., Schulz, J., Seibert, M. M., Shoeman, R. L., Sierra, R. G., Soltau, H., Spence, J. C. H., Stellato, F., Stern, S., Strüder, L., Ullrich, J., Wang, X.,

- Weidenspointner, G., Weierstall, U., Wunderer, C. B. & Chapman, H. N. T. (2012). *Nature Photon.* **6**, 35–40.
- Bekx, J. J., Son, S.-K., Santra, R. & Ziaja, B. (2018). *Phys. Rev. A*, **98**, 022701.
- Berrah, N. & Bucksbaum, P. H. (2014). *Sci. Am.* **310**, 64–71.
- Boutet, S., Lomb, L., Williams, G. J., Barends, T. R. M., Aquila, A., Bruce Doak, R., Weierstall, U., DePonte, D. P., Steinbrener, J., Shoeman, R. L., Messerschmidt, M., Barty, A., White, T. A., Kassemeyer, S., Kirian, R. A., Seibert, M. M., Montanez, P. A., Kenney, C., Herbst, R., Hart, P., Pines, J., Haller, G., Gruner, S. M., Philipp, H. T., Tate, M. W., Hromalik, M., Koerner, L. J., van Bakel, N., Morse, J., Ghonsalves, W., Arnlund, D., Bogan, M. J., Caleman, C., Fromme, R., Hampton, C. Y., Hunter, M. S., Johansson, L. C., Katona, G., Kupitz, C., Liang, M., Martin, A. V., Nass, K., Redecke, L., Stellato, F., Timneanu, N., Wang, D., Zatsepin, N. A., Schafer, D., Defever, J., Neutze, R., Fromme, P., Spence, J. C. H., Chapman, H. N. & Schlichting, I. (2012). *Science*, **20**, 362–364.
- Chapman, H. N. (2015). *Synchrotron Radiat. News*, **28**, 20–24.
- Chapman, H. N., Caleman, C. & Timneanu, N. (2014). *Philos. Trans. Roy. Soc. B Biol. Sci.* **369**, 20130313.
- Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmeß, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.-U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Røcker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R. M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature*, **470**, 73–77.
- Curwood, E. K., Quiney, H. M. & Nugent, K. A. (2013). *Phys. Rev. A*, **87**, 053407.
- Emma, P., Akre, R., Arthur, J., Bionta, R., Bostedt, C., Bozek, J., Brachmann, A., Bucksbaum, P., Coffee, R., Decker, F.-J., Ding, Y., Dowell, D., Edstrom, S., Fisher, A., Frisch, J., Gilevich, S., Hastings, J., Hays, G., Hering, Ph., Huang, Z., Iverson, R., Loos, H., Messerschmidt, M., Miahnahri, A., Moeller, S., Nuhn, H.-D., Pile, G., Ratner, D., Rzepiela, J., Schultz, D., Smith, T., Stefan, P., Tompkins, H., Turner, J., Welch, J., White, W., Wu, J., Yocky, G. & Galayda, J. (2010). *Nature Photon.* **4**, 641–647.
- Galli, L., Son, S.-K., Barends, T. R. M., White, T. A., Barty, A., Botha, S., Boutet, S., Caleman, C., Doak, R. B., Nanao, M. H., Nass, K., Shoeman, R. L., Timneanu, N., Santra, R., Schlichting, I. & Chapman, H. N. (2015). *IUCrJ*, **2**, 627–634.
- Galli, L., Son, S.-K., White, T. A., Santra, R., Chapman, H. N. & Nanao, M. H. (2015). *J. Synchrotron Rad.* **22**, 249–255.
- Hau-Riege, S. P., London, R. A., Chapman, H. N., Szoke, A. & Timneanu, N. (2007). *Phys. Rev. Lett.* **98**, 198302.
- Ho, P. J., Knight, C., Tegze, M., Faigel, G., Bostedt, C. & Young, L. (2016). *Phys. Rev. A*, **94**, 063823.
- Jurek, Z., Son, S.-K., Ziaja, B. & Santra, R. (2016). *J. Appl. Cryst.* **49**, 1048–1056.
- Lorenz, U., Kabachnik, N. M., Weckert, E. & Vartanyants, I. A. (2012). *Phys. Rev. E*, **86**, 059906.
- Lunin, V. Y., Grum-Grzhimailo, A. N., Gryzlova, E. V., Sinityn, D. O., Balabaev, N. K., Lunina, N. L., Petrova, T. E., Tereshkina, K. B., Abdunasyrov, E. G., Stepanov, A. S. & Krupyanskii, Y. F. (2013). *Mater. Biol. Bioinform.* **8**, 93–118.
- Lunin, V. Y., Grum-Grzhimailo, A. N., Gryzlova, E. V., Sinityn, D. O., Petrova, T. E., Lunina, N. L., Balabaev, N. K., Tereshkina, K. B., Stepanov, A. S. & Krupyanskii, Y. F. (2015). *Acta Cryst. D* **71**, 293–303.
- Murphy, B. F., Osipov, T., Jurek, Z., Fang, L., Son, S.-K., Mücke, M., Eland, J. H. D., Zhaunerchyk, V., Feifel, R., Avaldi, L., Bolognesi, P., Bostedt, C., Bozek, J. D., Grilj, J., Guehr, M., Frasiniski, L. J., Glowina, J., Ha, D. T., Hoffmann, K., Kuk, E., McFarland, B. K., Miron, C., Sistrunk, E., Squibb, R. J., Ueda, K., Santra, R. & Berrah, N. (2014). *Nat. Commun.* **5**, 4281.
- Neutze, R. (2014). *Philos. Trans. Roy. Soc. B Biol. Sci.* **369**, 20130318.
- Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature*, **406**, 752–757.
- Quiney, H. M. & Nugent, K. A. (2011). *Nat. Phys.* **7**, 142–146.
- Redecke, L., Nass, K., DePonte, D. P., White, T. A., Rehders, D., Barty, A., Stellato, F., Liang, M., Barends, T. R. M., Boutet, S., Williams, G. J., Messerschmidt, M., Seibert, M. M., Aquila, A., Arnlund, D., Bajt, S., Barth, T., Bogan, M. J., Caleman, C., Chao, T.-C., Doak, R. B., Fleckenstein, H., Frank, M., Fromme, R., Galli, L., Grotjohann, I., Hunter, M. S., Johansson, L. C., Kassemeyer, S., Katona, G., Kirian, R. A., Koopmann, R., Kupitz, C., Lomb, L., Martin, A. V., Mogk, S., Neutze, R., Shoeman, R. L., Steinbrener, J., Timneanu, N., Wang, D., Weierstall, U., Zatsepin, N. A., Spence, J. C. H., Fromme, P., Schlichting, I., Duszynko, M., Betzel, C. & Chapman, H. N. (2013). *Science*, **339**(6116), 227–230.
- Son, S.-K., Chapman, H. N. & Santra, R. (2011). *Phys. Rev. Lett.* **107**, 218102.
- Son, S.-K., Chapman, H. N. & Santra, R. (2013). *J. Phys. B At. Mol. Opt. Phys.* **46**, 164015.
- Son, S.-K., Young, L. & Santra, R. (2011). *Phys. Rev. A*, **83**, 033402.
- Tachibana, T., Jurek, Z., Fukuzawa, H., Motomura, K., Nagaya, K., Wada, S., Johnsson, P., Siano, M., Mondal, S., Ito, Y., Kimura, M., Sakai, T., Matsunami, K., Hayashita, H., Kajikawa, J., Liu, X.-J., Robert, E., Miron, C., Feifel, R., Marangos, J. P., Tono, K., Inubushi, Y., Yabashi, M., Son, S.-K., Ziaja, B., Yao, M., Santra, R. & Ueda, K. (2015). *Sci. Rep.* **5**, 10977.
- Urzhumtsev, A., Afonine, P. V. & Adams, P. D. (2009). *Acta Cryst. D* **65**, 1283–1291.
- Ziaja, B., Jurek, Z., Medvedev, N., Saxena, V., Son, S.-K. & Santra, R. (2015). *Photonics*, **2**, 256–269.