



Heavy-element damage seeding in proteins under XFEL illumination

Spencer K. Passmore,^{a,b*} Alaric L. Sanders,^{a,c} Andrew V. Martin^d and Harry M. Quiney^{a*}

^aSchool of Physics, University of Melbourne, Parkville, Victoria 3010, Australia, ^bOptical Sciences Centre, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia, ^cTCM Group, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, United Kingdom, and ^dSchool of Science, STEM College, RMIT University, Melbourne, Victoria 3000, Australia. *Correspondence e-mail: spencerpassmore@swin.edu.au, quiney@unimelb.edu.au

Received 7 November 2024

Accepted 30 June 2025

Edited by M. Yabashi, RIKEN SPring-8 Center, Japan

Keywords: heavy-element damage seeding; XFEL; femtosecond studies; serial crystallography; SFX; diffract-then-destroy; protein structure; single particles; pump–probe; radiation damage.

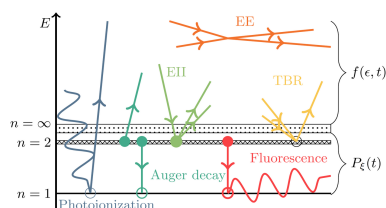
Supporting information: this article has supporting information at journals.iucr.org/s

Serial femtosecond X-ray crystallography (SFX) captures the structure and dynamics of biological macromolecules at high spatial and temporal resolutions. The ultrashort pulse produced by an X-ray free-electron laser (XFEL) ‘outruns’ much of the radiation damage that impairs conventional crystallography. However, the rapid onset of ‘electronic damage’ due to ionization limits this benefit. Here, we distinguish the influence of different atomic species on the ionization of protein crystals by employing a plasma code that tracks the unbound electrons as a continuous energy distribution. The simulations show that trace quantities of heavy atoms ($Z > 10$) contribute a substantial proportion of global radiation damage by rapidly seeding electron ionization cascades. In a typical protein crystal, sulfur atoms and solvated salts induce a substantial fraction of light-atom ionization. In further modeling of various targets, global ionization peaks at photon energies roughly 2 keV above inner-shell absorption edges, where sub-2 keV photoelectrons ejected from these shells initiate ionization cascades that are briefer than the XFEL pulse. These results indicate that relatively small quantities of heavy elements can substantially affect global radiation damage in XFEL experiments.

1. Introduction

Serial femtosecond crystallography (SFX) is a potentially revolutionary structural determination tool for biology that overcomes key challenges faced by conventional X-ray crystallography (Botha & Fromme, 2023). Using X-ray free-electron lasers (XFELs) to serially illuminate small crystals with ultrabright, femtosecond pulses of radiation, the structural signal of a molecule is captured so swiftly that the atomic nuclei are effectively frozen in place (Neutze *et al.*, 2000; Chapman *et al.*, 2014; Suga *et al.*, 2015). This approach not only mitigates structural damage, allowing for the use of much higher fluences, but also facilitates time-resolved crystallography. As such, SFX can be applied to capture molecular movies of rapid biological processes, including photoactivated dynamics (Tenboer *et al.*, 2014; Pande *et al.*, 2016; Suga *et al.*, 2017; Standfuss, 2019; Pandey *et al.*, 2020; Suga *et al.*, 2020; Carrillo *et al.*, 2021) and enzyme catalysis (Olmos *et al.*, 2018; Dasgupta *et al.*, 2019; Hull *et al.*, 2024).

In the ideal limit of SFX, the X-ray diffraction pattern of each crystal is captured by an ultrashort pulse before any secondary damage processes can propagate (Seibert *et al.*, 2011; Boutet *et al.*, 2012a; Spence, 2017; Chapman, 2017; Nass, 2019). In this ‘diffraction before destruction’ picture, damage no longer constrains the X-ray intensities that allow for successful structure determination. This freedom allows for



structure determination from micrometre or sub-micrometre crystals (Chapman *et al.*, 2011; Chapman *et al.*, 2014). Indeed, several XFEL experiments have captured the structures of radiation-sensitive metalloproteins without evidence of significant structural deformation (Chapman *et al.*, 2011; Suga *et al.*, 2015; Suga *et al.*, 2020). The technique can thus be applied to molecules that do not readily form large crystals, such as membrane proteins (Hunter & Fromme, 2011; Gisriel *et al.*, 2019; Martin & Sawyer, 2019; Zhao *et al.*, 2019; Botha & Fromme, 2023). Using smaller crystals is particularly advantageous for time-resolved experiments, which demand larger quantities of sample than static structure determination (Standfuss, 2019), are less tolerant to beam attenuation, and benefit from high crystal surface area to volume ratios when attempting to induce structural changes chemically (Schmidt, 2013).

In practice, ultrafast electronic processes disrupt the simple diffract-and-destroy picture – scattering signals seen in experiments generally remain dose-dependent even at the practical minimum pulse widths achievable (~ 10 fs) (Boutet *et al.*, 2012a; Martin & Quiney, 2016; Abbey *et al.*, 2016; Nass *et al.*, 2020). In particular, the electrons ejected through photoionization and subsequent Auger decays initiate electron impact ionization (EII) cascades that will go on to become the main cause of ionization well before the end of the pulse (Caleman *et al.*, 2011a; Vinko *et al.*, 2012; Chapman *et al.*, 2014). This ‘electronic damage’ has complex, non-linear effects on the scattered wavefield (Quiney & Nugent, 2011; Martin & Quiney, 2016; Abdullah *et al.*, 2018; Kozlov *et al.*, 2020b), and various theoretical treatments have been proposed to model the underlying ultrafast damage processes (Quiney & Nugent, 2011; Caleman *et al.*, 2011a; Hau-Riege, 2013; Abdullah *et al.*, 2018; Kozlov *et al.*, 2020a; Kozlov *et al.*, 2020b). Efficient use of limited XFEL beam time (Gisriel *et al.*, 2019; de la Mora *et al.*, 2020) is expected to depend on identifying experimental parameters that minimize electronic damage (Dickerson *et al.*, 2020).

Heavy atoms ($Z > 10$ in this context) are ubiquitous in protein crystals. Though they typically comprise less than 1% of the illuminated atoms, they play a key role in the biochemistry of metalloproteins. Additionally, their unusual scattering behavior has various applications to X-ray crystallography. Anomalous phasing, a notable example, has been successfully extended to SFX with microcrystals using sulfur native to proteins (Barends *et al.*, 2013; Nakane *et al.*, 2015; Nass *et al.*, 2021), or with much heavier atoms deliberately added for a stronger signal (Barends *et al.*, 2014; Hunter *et al.*, 2016; Gorel *et al.*, 2017). The rapid ionization of heavy atoms under high fluences induces a strong time dependence in their atomic scattering factors. This behavior is central to proposed methods of high-intensity radiation-damage-induced phasing (HI-RIP) (Son *et al.*, 2011; Son *et al.*, 2013; Chapman *et al.*, 2014; Hau-Riege & Bennion, 2015; Galli *et al.*, 2015a; Galli *et al.*, 2015b; Galli *et al.*, 2015c; Nass *et al.*, 2020; Caleman *et al.*, 2020).

A large part of the existing body of work dedicated to simulating atomic dynamics under XFEL illumination follows

either the Monte Carlo molecular dynamics (MD) paradigm (Abbey *et al.*, 2016; Jurek *et al.*, 2016; Ho & Knight, 2017; Ho *et al.*, 2020) or a plasma approach (Caleman *et al.*, 2011b; Leonov *et al.*, 2014; Kozlov *et al.*, 2020a; Ziaja *et al.*, 2023). These frameworks are in some senses complementary; the former provides a detailed picture of damage on the local scale, while the latter focuses on global statistics (Caleman *et al.*, 2020). In the MD case, individual ions and electrons are treated as classical particles and tracked through space. Such modeling is too computationally demanding for system sizes over 100–1000 atoms, so cannot simulate typically sized proteins in full (Hau-Riege, 2013; Abdullah *et al.*, 2016; Abdullah *et al.*, 2018; Nass *et al.*, 2020). In contrast, models that incorporate a zero-dimensional plasma code may capably simulate much larger (or arbitrary) systems (Caleman *et al.*, 2011b; Kozlov *et al.*, 2020a; Caleman *et al.*, 2020; Dawod *et al.*, 2024), mitigating possible difficulties in capturing the effect of trace elements.

Several plasma codes that treat the free electrons as non-thermal (*i.e.* non-Maxwellian) have been applied in studies of single-element, solid-density targets (Hau-Riege, 2013; Leonov *et al.*, 2014; Royle *et al.*, 2017; Le *et al.*, 2019; Ziaja *et al.*, 2023; Ren *et al.*, 2023; Shi *et al.*, 2024). However, plasma codes used in prior studies to simulate the damage sustained by biomolecules all approximate the free-electron energy distribution with a Maxwellian, under an assumption of instantaneous thermalization. This treatment may miss important effects – existing work on single-element, solid-density targets quotes thermalization times in picoseconds (Kho, 1985; Hau-Riege, 2013; Royle *et al.*, 2017; Abdullah *et al.*, 2017; Kitamura, 2019; Ren *et al.*, 2023).

In this work, we employ a non-thermal plasma physics code to examine electronic damage in protein crystals containing heavy elements without assuming instantaneous thermalization. We combine a custom-built frozen-shell Hartree–Fock code (Kozlov & Quiney, 2019) with a non-standard B-spline approach to solving the Boltzmann equation for the time-dependent non-equilibrium energy distribution of the free electrons. The details of this framework are given in Section 2. We apply the model to compare how light (C, N, O) and heavy atoms influence the ionization dynamics of biological matter in Section 3. We go on to identify how trace quantities of heavy elements in biomolecular targets give rise to species-dependent high-damage regions of the pulse parameter space in Section 4. We discuss new avenues for mitigating radiation damage implied by these findings in Section 5.

2. Simulation method

This study introduces a new non-thermal extension of the atomistic collisional-radiative plasma solver *AC4DC* described by Kozlov & Quiney (2019). While the previous version assumed free electrons to instantaneously thermalize after collisional-ionization, the new treatment allows for a non-Maxwellian free-electron energy distribution of arbitrary functional form. The scope of this study is restricted to the ionization dynamics, neglecting nuclear motion. This focus is motivated by two points: (i) ionization drives the ionic motion

(Coulomb explosion) during and immediately following the XFEL pulse; (ii) studies of XFEL radiation damage consistently show that changes in ionic state (electronic damage) make up the bulk of the radiation damage for pulses on the order of 10 fs (Abdullah *et al.*, 2018; Neutze *et al.*, 2000; Nass *et al.*, 2020).

The model couples the free-electron energy distribution, $f(\epsilon, t)$, to the population of possible ionic states, $P_\xi(t)$, through the processes of photoionization, Auger and fluorescent decay, electron impact ionization (EII), three-body recombination (TBR), and pairwise Coulomb electron–electron interactions (EE). These processes are summarized in Fig. 1. Atomic parameters are calculated in the radially averaged Hartree–Fock approximation (Kozlov & Quiney, 2019; Kozlov *et al.*, 2020a). EII and TBR are approximated using the well established binary-encounter dipole model of Kim & Rudd (1994). The equations of motion, obtained from radially averaging the Boltzmann equation, then read

$$\frac{\partial}{\partial t} f(\epsilon, t) = \mathcal{Q}[P_\xi, f](\epsilon), \quad (1)$$

$$\frac{d}{dt} P_\xi(t) = \sum_{\eta \neq \xi} \Gamma_{\eta \rightarrow \xi} P_\eta(t) - \Gamma_{\xi \rightarrow \eta} P_\xi(t), \quad (2)$$

where ξ and η denote electron configurations, and $\mathcal{Q}[P_\xi, f](\epsilon)$ and Γ represent the couplings to processes affecting f and P , respectively (illustrated in Fig. 1). Details for the calculations of the atomic cross-sections, Γ , are given by Kozlov & Quiney (2019). In terms of separate processes, $\mathcal{Q}[P_\xi, f](\epsilon)$ has the form

$$\begin{aligned} \mathcal{Q}[P_\xi, f](\epsilon) = & \mathcal{Q}^{\text{Photo}}[P_\xi](\epsilon) + \mathcal{Q}^{\text{Auger}}[P_\xi](\epsilon) + \mathcal{Q}^{\text{EII}}[P_\xi, f](\epsilon) \\ & + \mathcal{Q}^{\text{TBR}}[P_\xi, f, f](\epsilon) + \mathcal{Q}^{\text{EE}}[f, f](\epsilon). \end{aligned} \quad (3)$$

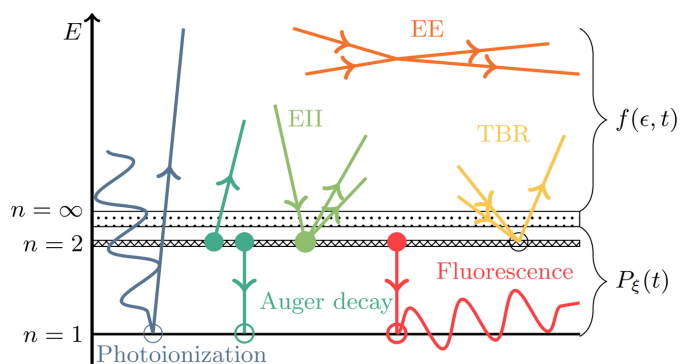


Figure 1 Bound–free, bound–bound and free–free transitions in biomolecular plasma. Acronyms denote electron–electron scattering (EE), electron–impact ionization (EII) and three-body recombination (TBR). Bound energy levels are reminiscent of carbon for illustrative purposes. Processes are labeled to indicate the populations they affect: one or both of \mathbf{P} (the ionic states) and f (the free electrons). The dotted section at $E = 0$ represents the weakly bound molecular structure that is ignored here. Filled circles represent initial-state bound electrons, and hollow circles their final states.

The notation adopted in equation (3) is such that each $\mathcal{Q}[\dots]$ is multilinear in all arguments. Explicit expressions for these source terms and collision kernels are given in Appendix A.

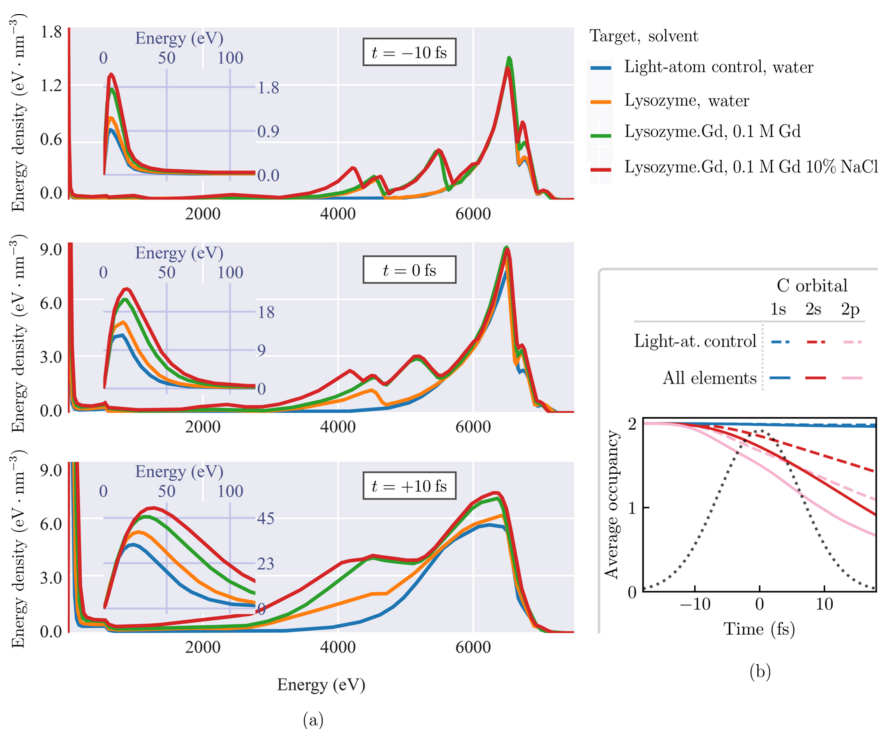
The new code tracks an arbitrary non-Maxwellian distribution by expanding $f(\epsilon)$ with piecewise-polynomial B-splines $B_k(\epsilon)$ (de Boor, 1978). These basis functions are non-zero only for a small region of energy, allowing for an efficient sparse internal representation of \mathcal{Q} without sacrificing the differentiability of f . As the simulation progresses, the spline grid is dynamically adapted to have increased density across the energy range spanning the bulk of the thermal electrons and in the vicinity of the main primary ionization peaks (see Appendix A for further details). This adaptive grid allowed the code to perform full dynamical non-thermal plasma simulations of lysozyme in under an hour on a contemporary desktop.

The total number of modeled atomic configurations for an element scales like the factorial of the total number of subshells of the ground state, making heavier elements very costly to simulate. For computational expedience, $n \geq 2$ shells of $Z \geq 30$ elements were modeled with a single angular momentum (*i.e.* a single energy level). Test simulations showed that ionization rates of light and heavy elements were barely affected by this approximation in the regime considered in this work (see Fig. S3 of the supporting information).

We quantitatively compared this new version of *AC4DC* against a number of published results in the literature, which are presented in Appendix C. The code achieved excellent agreement with non-Maxwellian Monte Carlo simulations by *ddcMD* of amorphous carbon (Hau-Riege, 2013) for the evolution of the ionic states and free electrons, and predicted similar ionic state dynamics to simulations by *XMDYN* of a glycine crystal (Jurek *et al.*, 2016; Abdullah *et al.*, 2018). Somewhat surprisingly, the code also sees good agreement with the particle-in-cell DFT code *PICLS* (Royle *et al.*, 2017) when modeling aluminium plasma for both the ion and free-electron populations. *AC4DC* did not reproduce results for simulations of silicon (Leonov *et al.*, 2014) by a plasma code that uses a similar physical framework to *AC4DC* but which fits $f(\epsilon)$ to a standard grid rather than using adaptive splines (note that we found the adaptive grid crucial to achieving convergence, see Appendix A).

3. Heavy-seeded ionization cascades – lysozyme.Gd

We first study the impact of heavy atoms in a representative model system – gadolinium-derivative hen egg-white lysozyme [lysozyme.Gd, derived from PDB entry 4et8 (Boutet *et al.*, 2012a; Boutet *et al.*, 2012b)] – subjected to a 15 fs FWHM Gaussian pulse of fluence 1.75×10^{12} 7.112 keV photons μm^{-2} . The target contains 35.1% (v/v) solvent [10% m/v NaCl, 0.1 M Na acetate, 0.1 M Gd (Nass *et al.*, 2020)], which we hereafter refer to as the 0.1 M Gd 10% NaCl solvent for brevity. The resulting unit cell, including disordered solvent atoms, has a chemical composition of $\text{H}_{13259}\text{C}_{5153}\text{N}_{1596}\text{O}_{4009}\text{S}_{80}\text{Gd}_{21}\text{Na}_{93}\text{Cl}_{87}$.


Figure 2

Effect of heavy atoms on electronic damage in lysozyme. (a) Snapshots of the free-electron energy distribution for the light-atom control (blue), lysozyme in water (orange), lysozyme.Gd in 0.1 M Gd solvent (green) and the target of lysozyme.Gd in 0.1 M Gd 10% NaCl (Nass *et al.*, 2020) (red). Inset plots show the distributions at the scale of the thermalized electrons. Snapshot times are denoted relative to the pulse's peak intensity ($t = 0$ fs). (b) Corresponding evolution of the average occupancy of the electron orbitals of carbon in the light-atom control (broken lines) and lysozyme.Gd (solid lines); the black dotted line traces the temporal pulse profile. Each simulation used a 15 fs FWHM Gaussian pulse with a fluence of 1.75×10^{12} 7.112 keV photons μm^{-2} .

Fig. 2(a) shows the free-electron energy distribution of the lysozyme.Gd system at times -10 fs, 0 fs and $+10$ fs from the pulse peak. To infer the effect of the heavy atoms on the dynamics, simulations of three ‘toy’ variants of the system are also shown: lysozyme.Gd in 0.1 M Gd solution ($\text{H}_{13846}\text{C}_{5143}\text{N}_{1596}\text{O}_{4300}\text{S}_{80}\text{Gd}_{21}$), lysozyme in water ($\text{H}_{13942}\text{C}_{5056}\text{N}_{1576}\text{O}_{4386}\text{S}_{80}$) and a ‘light-atom control’ ($\text{H}_{13942}\text{C}_{5056}\text{N}_{1656}\text{O}_{4386}$). These have the same solvent density (1.1 g mol^{-1}) and protein concentration as the complete system. It can be seen in Fig. 2(a) that heavy elements make significant contributions to the high-energy region of the distribution relative to their small presence. In addition, the inset plots show a substantially larger thermal (low-energy) electron population for the complete target than the light-atom control. The vast majority of the additional thermal electrons come from additional secondary ionization events in the *light* atoms. This is reflected by Fig. 2(b), which shows that the carbon 2s and 2p orbitals deplete much more rapidly in the complete target.

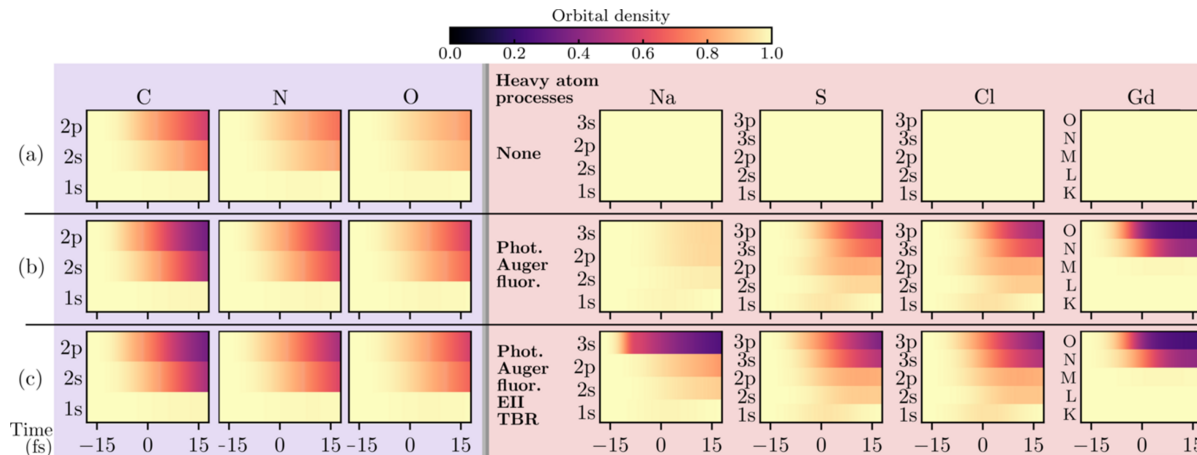
3.1. Comparing the ionization behavior of light and heavy elements

Fig. 3 shows that the influence of the heavy ions on global ionization almost entirely originates with their primary ionization processes. Modeling that neglects the secondary ionization of heavy atoms [Fig. 3(b)] still produces a nearly

identical evolution in the carbon ionic states to those of the complete simulation shown in Fig. 3(c). In other words, the effect of heavy elements on the dynamics is almost entirely due to their primary electron emissions.

The primary electron production rate, from photoelectron and Auger decay, rises substantially with higher Z . This is partly due to the huge scaling of the photoabsorption cross-section [Z^5 for the K -shell (Bethe & Salpeter, 1957)], and partly due to shorter Auger lifetimes of heavier elements — ≈ 1 fs for elements heavier than sulfur (Campbell & Papp, 2001; Nass, 2019; Son *et al.*, 2020). Fig. 3(c) shows that the M -shells of the Gd ions are nearly full for the entire simulation. The rapid ejection of photoelectrons from the M -shell is sustained by the electrons in higher orbitals, which fill core holes on a subfemtosecond timescale through Auger decay and fluorescence. Previous experimental studies have noted that such ‘Auger cycling’ significantly elevates primary ionization under XFEL pulses (Ho *et al.*, 2020; Rudenko *et al.*, 2017), though here it also elevates the number of secondary ionization cascades seeded by Gd.

To quantify the contribution of each element to global ionization in terms of the damage processes that they seed through primary ionization, we construct a measure of the response of the full free-electron energy distribution f to a particular source of primary electrons. We introduce auxiliary distributions $g_n(\epsilon, t)$ that partition f , *i.e.* $f(\epsilon, t) = \sum_n g_n(\epsilon, t)$, such that they represent the free-electron density associated


Figure 3

Impact of heavy element electronic processes on global ionization for lysozyme.Gd in 0.1 M Gd 10% NaCl solvent. Plots show average orbital (1s, 2s, ...) or shell (K, L, ...) densities normalized to 1 in the initial undamaged state, for each element in the target. All atomic processes in the physical model are enabled for light atoms (C, N, O); while the modeled processes in heavy atoms (S, Gd, Na, Cl) are denoted (a) no heavy atom processes, (b) heavy atom photoionization, Auger decay and fluorescence, (c) all atomic processes in the physical model. The elevated depletion of the light atom orbitals in (b) and (c) is caused by additional collisional-ionization processes stemming from heavy atom primary ionization. No such difference in the carbon states is apparent between (b) and (c), indicating that heavy atom secondary ionization is unimportant to light atom ionization. Pulse parameters match those given in Fig. 2.

with specific primary ionization processes, e.g. $n = \text{C}$ Auger or $n = \text{O}$ photoionization. Explicitly, we decompose the right-hand side of equation (3) as

$$\begin{aligned} \frac{\partial g_n(\epsilon, t)}{\partial t} = & Q_n^{\text{Photo}}[P] + Q_n^{\text{Auger}}[P] + Q^{\text{EII}}[P, g_n] \\ & + \frac{1}{2} \sum_m \left(Q^{\text{TBR}}[P, g_n, g_m] + Q^{\text{TBR}}[P, g_m, g_n] \right. \\ & \left. + Q^{\text{EE}}[g_n, g_m] + Q^{\text{EE}}[g_m, g_n] \right). \end{aligned} \quad (4)$$

This choice of decomposition assigns all EII-sourced secondary electrons to the cascade of the impactor (g_n), while the density arising from TBR and EE processes between different cascades is evenly divided between the two.

Fig. 4 shows the aggregate density of the cascades seeded by each atomic-species population. Comparing the light-atom charges [Fig. 4(a)] with the density of the light-atom-seeded cascades [Fig. 4(b)] highlights that the global ionization induced by an atomic species is decoupled from its ionization rate. C, N and O are ionized at similar rates, largely by EII; however, the O atoms seed a significantly larger fraction of the EII cascades, owing to the Z^5 scaling of the K-shell photo-absorption cross-section. It can be further seen that the free-electron density seeded by each heavy-element population [Fig. 4(c)] is comparable with that seeded by the far more abundant light-atom populations. Strikingly, the combined density of the Gd-seeded cascades is near or above that of the O-seeded cascades at all times, despite the system containing O and Gd in a ratio of 190:1. Combined, the cascades seeded by heavy elements account for the majority of the freed electrons in the system for the entirety of the pulse.

From this perspective, it can be understood why the heavy elements have a disproportionate influence on the light atom

ionization. The rapid primary ionization of heavy atoms means they seed a significant number of EII cascades in the target despite their trace presence. Since the vast majority of the ionization in the target is due to EII (Vinko *et al.*, 2012; Chapman *et al.*, 2014), a substantial fraction of ionization in the light-atom structure occurs through cascades seeded by the heavy elements.

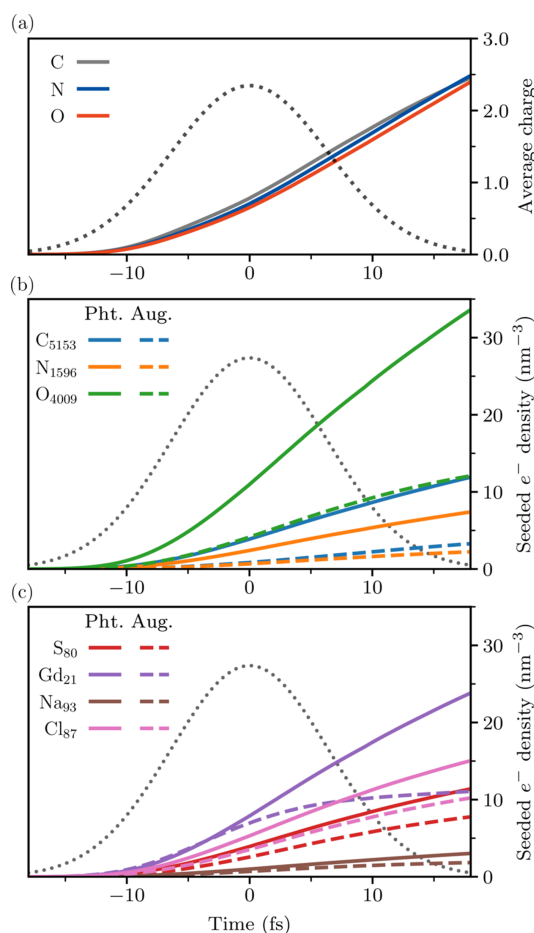
Note that the density of the cascades seeded by Gd Auger electrons [Fig. 4(b)] sees negligible growth after $t \simeq 0$ as a result of the depletion of the O- and N-shells, observable in Fig. 3(c). By $t = 5$ fs, the average Gd ion charge is above +20. In reality, the highly charged Gd ions would be replenished by bound electrons transported from neighboring atoms (Rudenko *et al.*, 2017), likely leading to a stronger damage seeding effect.

3.2. Examining unusual Gd charge measurements in experiments

The ionization of Gd ions in lysozyme.Gd under XFEL pulses has previously been investigated by Galli *et al.* (2015a) and Nass *et al.* (2020). The charge density of the Gd ions is inferred from the integrated electron density around the Gd site relative to a reference region of light atoms, which we refer to as the electron density ratio (EDR). Nominally,

$$\text{EDR} \propto \int dt \Phi(t) \frac{n_e^{\text{Gd}}(t)}{2n_e^{\text{C}}(t) + n_e^{\text{N}}(t) + n_e^{\text{O}}(t)}, \quad (5)$$

where $\Phi(t)$ is the normalized temporal pulse profile, and $n_e^X(t)$ is the average number of electrons bound to element X in the Gd site or light-atom reference region. Assuming the intensity-averaged light-atom ionization is small, the charge gained by Gd can be approximated. Notably, this measure indicated the ionization of Gd to be unexpectedly low in both studies.


Figure 4

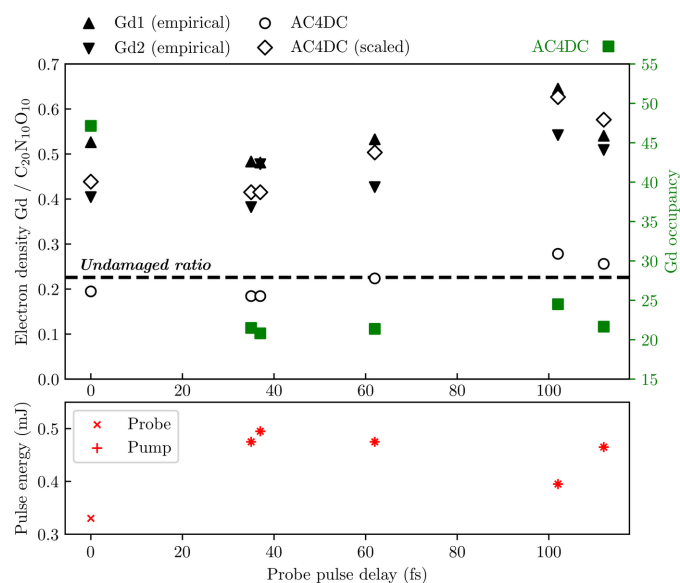
Contribution of each element to the global ionization in the lysozyme.Gd target. Panel (a) plots the average charge of each light-atom species. Panels (b) and (c) trace the collective free-electron density of the EII cascades seeded by each light-atom species and heavy-atom species, respectively. The traces in (b) and (c) correspond to all cascades initiated by photoionization (solid) or Auger decay (dashed) within the population of the denoted atomic species (subscripts in the legend give the population per unit cell). Note that these traces sum to the total free-electron density of the system, minus the negligible hydrogen photoelectron density. It can be seen that both photo- and Auger electrons ejected by heavy elements are significant to the global ionization dynamics. At all times, the heavy-element cascades (c) constitute the majority of the free-electron density (58% at $t = 0$). Secondary ionization of heavy elements was ignored in these simulations. Pulse parameters match those given in Fig. 2.

Fig. 5 compares EDR values obtained from X-ray-pump–X-ray-probe experiments by Nass *et al.* (2020) (using a light-atom reference region of $\text{C}_{20}\text{N}_{10}\text{O}_{10}$) with predictions from simulations of these experiments by *AC4DC* using the solvated lysozyme.Gd target. The simulated EDRs accurately reproduce the features of the experimental data, but rescaled by a factor of about two.

For the four EDR values where the total pulse energy has a range 0.95–0.99 mJ (35 fs, 37 fs, 62 fs and 112 fs), the EDR rises with increasing probe delay. We argue that this is consistent with the picture of the dynamics described in Section 3.1, which highlights that secondary ionization – the only ionization during the ‘dead time’ between pump and probe – affects the light atoms far more strongly than the

heavy atoms. This implies that the EDR does not gauge the electron density of the heavy atoms but instead serves as a rough indicator of the ratio of secondary ionization to primary ionization. Under this perspective, the EDRs at 0 fs and 102 fs – which correspond to lower pulse energies – are raised due to the greater significance of secondary ionization at lower fluences.

It is difficult to diagnose the cause of the factor of two systematic offset between the predicted and observed EDRs in Fig. 5. A likely contributing factor is the assumption of a homogeneous spatial profile in the *AC4DC* model. In the experiment, the beam was nominally focused to a $\sim 0.2 \mu\text{m}$ FWHM Gaussian, smaller than the $\sim 0.5 \mu\text{m}$ lysozyme.Gd nanocrystals (Nass *et al.*, 2020). Photons scattering outside the focus will sample a less-ionized region where, repeating the reasoning from earlier, secondary ionization will be relatively strong and the EDR will be higher. Prior work has highlighted that such photons can make up a substantial fraction of detected photons (Murphy *et al.*, 2014; Galli *et al.*, 2015a; Nass *et al.*, 2015), suggesting this correction could be quite strong.


Figure 5

Electron density ratio of Gd ions to light ions (EDR) in lysozyme.Gd as measured in experiments by Nass *et al.* (2020) and in corresponding *AC4DC* simulations. The pulse energy corresponds to individual pump or probe pulses. All data points correspond to a pump pulse at $t = 0$ and a probe pulse at some delay, except for the ‘single pulse’ at the $t = 0$ data point. Square points correspond to the probe-intensity-averaged Gd occupancy in the simulations (right axis). The dashed line shows the notional EDR of the undamaged target. The triangles correspond to the EDR observed by Nass *et al.* (2020) for the two Gd atoms in the asymmetric unit (Gd1 and Gd2). The simulated EDRs (circles) were computed as the probe-intensity-averaged occupancy of a Gd ion over that of a set of light atoms ($\text{C}_{20}\text{N}_{10}\text{O}_{10}$), as in the considered study (Nass *et al.*, 2020). The diamonds show the simulated EDRs scaled by a constant factor. The relative changes in the empirical EDR between probe pulse delays are remarkably similar to those predicted by *AC4DC*. In contrast, the EDRs are incommensurate with the Gd occupancy. The fluence of each pulse was modeled as the average within the nominal $0.2 \mu\text{m}$ focus of the experiment, and with the nominal 15 fs FWHM Gaussian temporal profiles. Simulations started 18 fs before the pump pulse, and ended 18 fs after the probe pulse. Theoretical values were measured by integrating over a 30 fs timespan, centered on the probe pulse.

An additional cause may be bound-to-bound electron transport, which *AC4DC* neglects. As alluded to in Section 3.1, electron transport was previously shown to slow the charge gain of iodine in iodobenzene (Rudenko *et al.*, 2017), a molecule with fewer light atoms than the Gd complex. Lastly, it is possible that the experimental measurements of the light-atom reference region did not capture all bound electrons. We note that the observed EDRs suggest implausibly high ionization rates: the higher EDR observed by the $t = 0$ pulse corresponds to an *intensity-averaged* light-atom charge of at least +3.8, even assuming negligible Gd ionization. For reference, the considered work predicted the average light-atom charge to remain below +1.6 at all times.

Using a similar methodology to Nass *et al.* (2020), Galli *et al.* (2015a) also attempted to measure the charge of Gd in two cases – ‘high fluence’ and ‘low fluence’, with fluences peaking at 7.8 and 0.13×10^{12} photons μm^{-2} , respectively. The work reports only the difference in intensity-averaged Gd charge between these cases, termed the Gd ‘ionization contrast’, as measured by computing the difference in the EDRs and assuming negligible light-atom ionization. A Gd ionization contrast of 8.8–12 electrons per Gd was observed; however, theoretical modeling using the *XATOM* toolkit predicted that the Gd charge difference ought to be 25 under a spatially homogeneous intensity. We repeated this modeling using *AC4DC*, finding an even more extreme 33.1. This apparent discrepancy between theory and experiment is substantially reduced by accounting for light-atom ionization lowering the denominator in equation (5) (see Fig. S4 of the supporting information) – we obtain a Gd ionization contrast of 21.4 after making this correction. Modeling the cascades seeded by the salt (namely Cl) and Gd ions in the crystal was crucial to seeing this level of agreement, due to their significant contribution to light-atom ionization. Without doing so, the simulated Gd ionization contrast only falls to 25.2 (note that the actual Gd charge difference remains at 33.1 in either case).

4. Damage by different trace elements

In this section, we present the results of simulations designed to examine how the energy-level structure of a trace, heavy element in a biological target affects electronic damage. Instead of modeling a specific protein, we consider a set of targets with densities of $\sim 1.2 \text{ g cm}^{-3}$ and atomic ratios of $\text{C}_{613}\text{N}_{193}\text{O}_{185}\text{X}_{10}$, with the ‘dopant’ X swapped out for various elements. The ionization of hydrogen was ignored for these simulations; testing showed this to be a minor approximation, increasing the average light-atom charge by ~ 0.1 or less.

4.1. Photon and primary electron energy

Fig. 6(a) shows that the choice of dopant substantially impacts global ionization. At all energies, the intensity-averaged carbon charge varies widely across the targets. Most traces show decreasing damage with increasing photon energy, consistent with prior work (Neutze *et al.*, 2000); however, the traces for the targets doped with Fe^{2+} ($Z = 26$), Zn^{2+} ($Z = 30$)

and Se ($Z = 34$) buck this trend, with the transition over the K -edge of the dopant increasing the average charge by roughly a factor of two or more in each case. This reflects the substantial contribution of the K -shell to the primary ionization of these dopants; the $1s$ orbitals dominate the photoabsorption cross-sections of their atoms and are replenished on a subfemtosecond timescale by Auger cycling (see Section 3.1).

Perhaps more surprisingly, the intensity-averaged carbon charge *continues* to increase up to $\sim 2 \text{ keV}$ above an inner-shell absorption edge, despite a decreasing photoabsorption cross-section. This is most clearly seen in Fig. 6(b). Since the separation from the edge corresponds to the energy of the photoelectrons emitted from the shell, such a result might be explained if the photoelectrons of $\sim 2 \text{ keV}$ are unusually damaging. Thus to isolate the source of the effect, it is necessary to quantify the damage that a single primary electron emission induces. We add an artificial, low-density primary-electron peak to the continuum in the simulation of

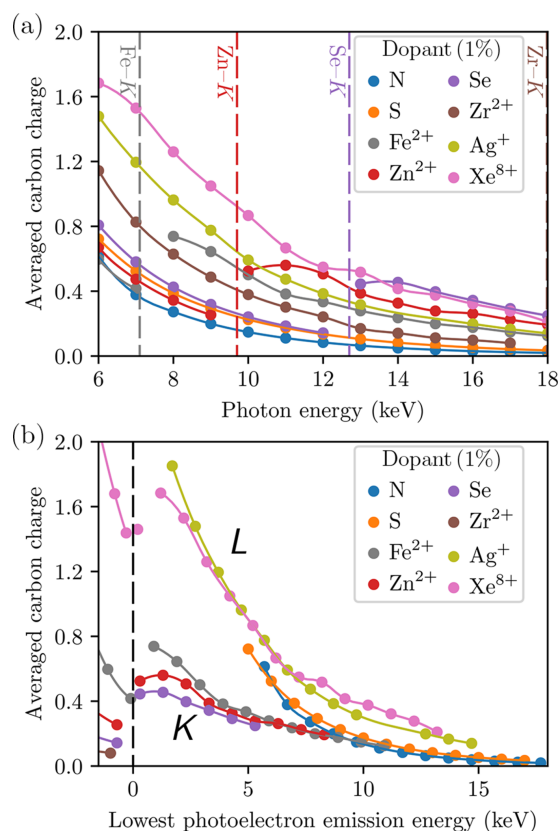
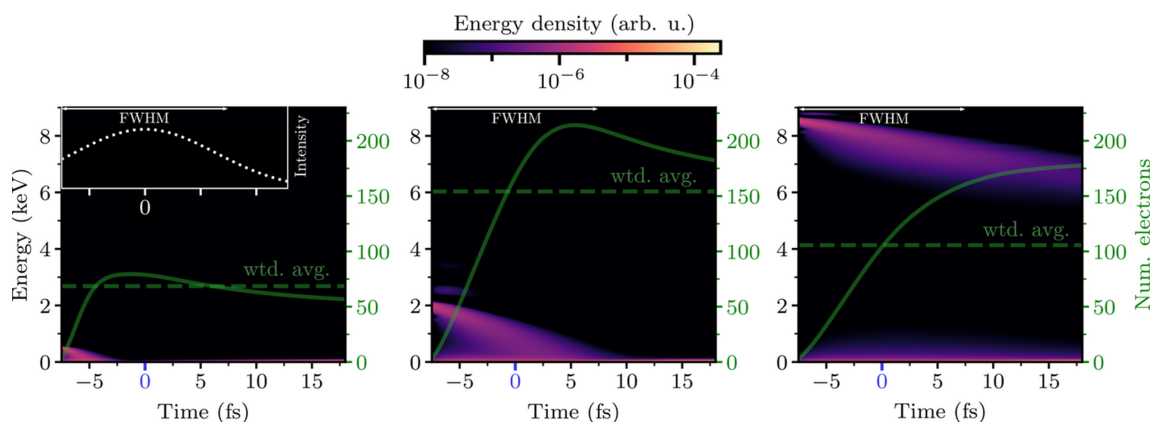


Figure 6 Relationship between light-atom ionization and the energy of photoelectrons ejected by trace heavy elements. Each trace corresponds to a target of composition $\text{C}_{613}\text{N}_{193}\text{O}_{185}\text{X}_{10}$, where X is the dopant denoted in the legend. Panel (a) shows the intensity-averaged charge of carbon against the photon energy of the pulse. The points are shifted in (b) to align with the separation between the deepest ionizable shell (DIS) of the dopants at 18 keV and the photon energy, representing the lowest-energy photoelectrons emitted by the DIS (when positive). While the severity of ionization varies considerably between targets at any given photon energy, the traces in (b) form two distinct groups according to the shell number of the DIS (annotated). The interpolating lines are included as a guide for the eye. Photon fluence was fixed at 10^{12} photons μm^{-2} , using a 15 fs FWHM Gaussian profile. Simulations were run from -18 fs to $+18$ fs relative to the pulse peak.


Figure 7

Evolution of electron-ionization cascades under plasma conditions in a light-atom target. Each plot shows a cascade density initiated by a free electron of energy (a) 0.5 keV, (b) 2 keV and (c) 8.5 keV. Each cascade occurs during a full dynamical simulation of the $C_{613}N_{203}O_{185}$ target under the 15 fs FWHM 9 keV pulse, in equivalent simulation conditions to Fig. 6, with the cascade initiated when the pulse first reaches the FWHM intensity (-7.5 fs). The right axis of each plot shows the density of the cascade, normalized to correspond to the number of electrons. Dashed horizontal lines correspond to the intensity-averaged electron count. The 500 eV cascade (oxygen Auger electron) deposits a small amount of energy quickly, while the 8.5 keV cascade (oxygen photoelectron) is largely outrun by the pulse. The 2 keV cascade is the most damaging, depositing its energy over a timescale similar to the 15 fs FWHM of the XFEL pulse. Each cascade was tracked by injecting a negligibly small electron density into the system at $t = -7.5$ fs. At late times, the cascade densities are reduced by recombination.

the CNO ($C_{613}N_{203}O_{185}$) when the pulse first reaches the FWHM intensity, then track the resulting electron cascade as in Section 3.1 [see equation (4)].

Fig. 7 confirms that the intensity-averaged ionization caused by cascades initiated by 2 keV primary electrons is higher than those caused by 0.5 keV and by 8.5 keV primary electrons. This observation is consistent with the local maximum of ionization near the photoelectron emission energy of 2 keV in Fig. 6(b).

To understand why the intermediate-energy cascade is more damaging, one can consider a free electron of energy E interacting with a gas of hydrogenic atoms, with electrons bound by energy B . As E falls, the EII cross-section grows, causing an increased ionization rate down to $E \simeq B$ (Kim & Rudd, 1994; Suno & Kato, 2006). This creates a trade-off: higher-energy cascades progress more slowly, but have the potential to free more electrons. Indeed, this effect can be seen in comparisons of the evolution of electron cascades at various energies in neutral targets presented in prior work (Chapman *et al.*, 2014; Coleman *et al.*, 2015). When counting all ionization events over a fixed duration, there will be a maximally ionizing cascade energy (Fig. 8 gives a qualitative sketch of this idea). This implies the existence of a ‘maximally damaging’ cascade energy E_{\max} , loosely corresponding to the energy that maximizes the intensity-averaged ionization. Under conditions that induce significant atomic disorder during the pulse, E_{\max} is likely lowered by Bragg gating (Barty *et al.*, 2012; Chapman *et al.*, 2014; Coleman *et al.*, 2015). Due to the higher photoabsorption cross-section near the edge, the most damaging photons eject primary electrons with energies below E_{\max} ; however, simulations that control for the rate of primary ionization show the difference to be slight (about 0.5 keV, see Appendix B).

Fig. 6(b) plots the average carbon charge against the lowest photoelectron energy (LPE) at emission instead of the photon energy. The traces form two distinct groups: the ‘K-group’,

consisting of targets doped by an element whose ‘deepest ionizable shell’ (DIS) in the considered energy range is the K-shell (N, S, Fe, Zn, Se), and the ‘L-group’, where it is the L-shell (Ag, Xe). The split in groups again appears to be a result of Auger cycling. As the DIS of each heavy dopant is maintained near maximum occupancy by decay processes, the production rate of ionizing electrons by Ag^+ and Xe^{8+} is scaled by a factor of about four relative to the other dopants at an equivalent LPE, due to the higher electron capacity of their L-shells.

In general, the variation in electronic damage across the targets with respect to photon frequency is predominantly controlled by the LPE of the dopants and the shell number of

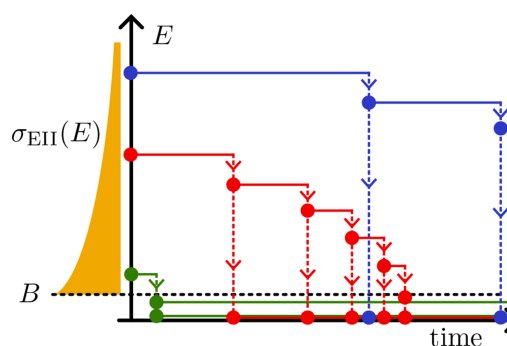

Figure 8

Diagram illustrating how intermediate-energy free electrons can be more damaging on sufficiently short timescales. Three free electrons, represented as dots, are shown initially with low (green), intermediate (red) and high (blue) energies incident on a gas of bound electrons; binding energy B . The horizontal lines represent the mean time between electron-impact ionization (EII) events, which is smaller for lower energy free electrons. Each branching event represents the most likely EII process: an electron of energy E ‘only just’ ionizes an atom, leaving one electron at zero energy and another at $E - B$. Because the EII cross-section (σ_{EII} , indicated by the yellow trace) is lower for faster free electrons, there exists a window of time where, on average, the intermediate-energy electron cascade will have ionized more atoms than the high- and low-energy cascades.

the DIS. However, the traces of the N-, S- and Fe-doped targets deviate from the rest of the *K*-group in Fig. 6(b) for photon energies at or below 8 keV. In these cases, the photoelectrons ejected from the light atoms are of a low enough energy to contribute significantly to their secondary ionization (relative to the dopants). Since the Zn and Se *K*-edges are above 8 keV, their traces show no deviation. It is unsurprising that the *L*-group does not contain such outliers, as the primary electron contribution of the *L*-group dopants is much more significant.

4.2. Damage landscape

Repeating the simulations shown in Fig. 6 while ignoring heavy-atom secondary ionization had little effect on the traces, as predicted by the analysis in Section 3.1. Since this substantially reduces the number of configurations to be processed in the EII and TBR calculations, additional simulations were performed using this approximation to map the landscape of damage (Netze *et al.*, 2000) for each target (Fig. 9).

Contour plots of the charge accumulated by the carbon atoms, as functions of photon energy and fluence [Fig. 9(a)] or pulse width [Fig. 9(b)], show significant differences due to the dopants. The effect of sulfur on the dynamics is relatively consistent; across the parameter space of Fig. 9, the S-doped target experiences levels of ionization that the pure CNO target only sees at 20–40% higher intensities. The impact of the heavier dopants is more variable due to their deep absorption edges, which are distinctly visible in these plots – the most extreme case is the Se-doped target, where the elevation in damage from increasing the photon energy from

12.5 keV to 13.5 keV is similar to the elevation in damage from tripling the pulse fluence [Fig. 9(a)], or from increasing the pulse width by an order of magnitude [Fig. 9(b)].

Inspection of Fig. 9 shows that the ionization in the doped targets consistently sees a local maximum around an LPE of 2 keV, suggesting that this may serve as an approximate value for E_{\max} in this regime (see Section 4.1). Both the slow Auger electrons and fast photoelectrons ejected by light atoms are well away from this value for E_{\max} in typical hard X-ray SFX experiments, while the primary electrons ejected by heavy atoms are almost always closer. Indeed, this is the case in the lysozyme.Gd scenario considered in Section 3. For example, the ~ 4 keV photoelectron peak of Cl, observable in Fig. 2(a), is closer to 2 keV than the >6.5 keV light-atom photoelectrons and thus more damaging, and the same is true of the Auger electrons, which are typically ~ 2.3 keV for Cl and <500 eV for light atoms. The tendency of heavy atoms to eject electrons with ‘intermediate’ energies closer to E_{\max} thus contributes to the strength of their effect.

Fig. 9 allows for the magnitude of damage in targets containing different heavy elements to be compared on the keV energy scale. This picture is only accurate in a coarse-grained sense, as resonance-related phenomena near absorption edges that modulate the fine structure of the damage landscape (Kanter *et al.*, 2011; Rudek *et al.*, 2012; Rudek *et al.*, 2018; Ho *et al.*, 2020) are not modeled here. In preparation of these diagrams, we were careful to avoid any photon energies close to resonance. We neglected ionization potential depression (IPD) (Vinko *et al.*, 2014; Vinko *et al.*, 2015; Ciriosta *et al.*, 2016; Royle *et al.*, 2017), which a short order of magnitude calculation shows to be limited to ~ 50 eV at biomolecular densities.

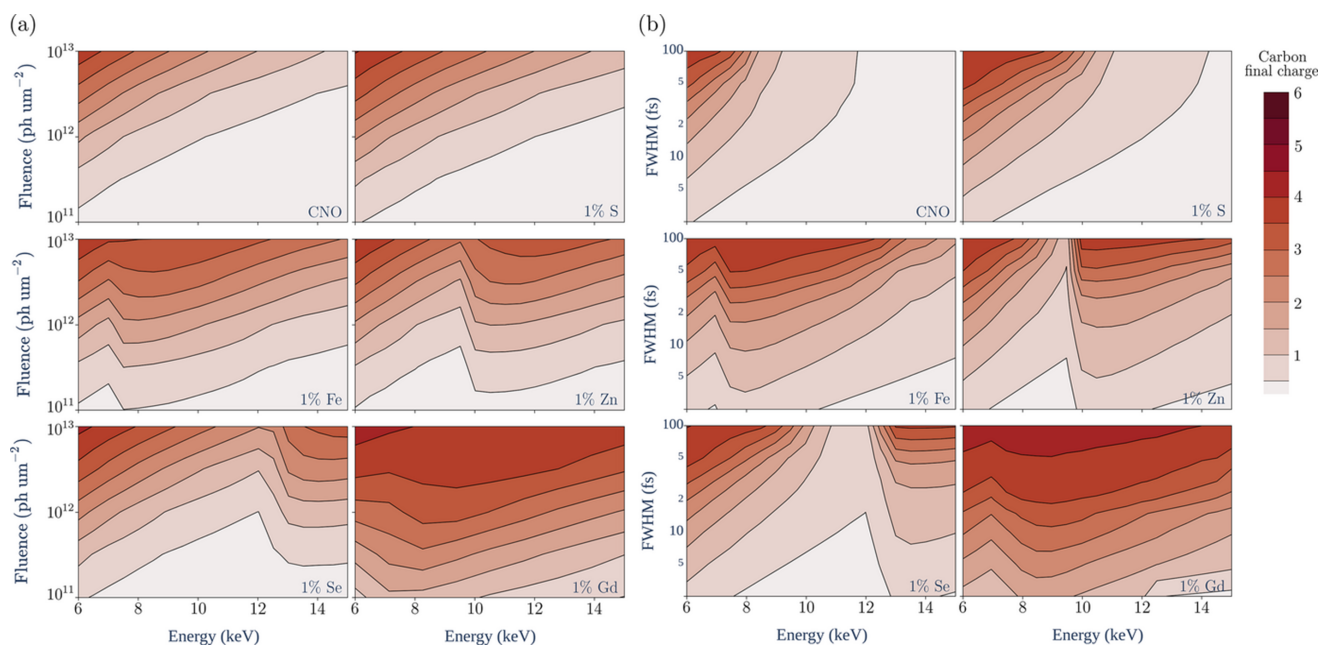


Figure 9 Effect of chemical composition on the electronic damage landscape. (a) 15 fs FWHM Gaussian pulses. (b) Gaussian pulses of 10^{12} photons μm^{-2} fluence. Each plot maps the average charge of carbons at the end of the illumination (+1.2 FWHM) to a target of composition $\text{C}_{613}\text{N}_{193}\text{O}_{185}\text{X}_{10}$, where *X* is the element denoted in the lower-right corner. Each plot uses (a) 133 or (b) 190 data points. The sharp features in the plots for targets doped by heavier elements correspond to the dopant absorption edges.

5. Discussion

The results presented here suggest that the absorbed dose and evolution of the electronic structure are strongly affected by the presence of heavy elements. For all targets considered, a high-energy regime exists in which the general principle that ‘higher photon energy = less damage’ remains true. However, for certain targets, the maximum photon energies available from XFEL sources are too low to enter this regime (Fig. 9). This is a departure from conventional crystallography, where higher energies are broadly beneficial. More positively, this suggests that SFX experiments may be able to control the severity of damage through the choice of X-ray frequency or solvent composition. For example, excluding salt from the solvent in the simulation of non-derivative lysozyme (Fig. 2) reduced the average carbon charge at $t = 10$ fs from +1.53 to +1.28.

Elements much heavier than sulfur, such as Se and Gd, are often introduced to targets for anomalous phasing, generally in tandem with an X-ray energy just above their ionization edges (Son *et al.*, 2013; Galli *et al.*, 2015a; Hunter *et al.*, 2016; Gorel *et al.*, 2017). It has been argued that difficulties in applying the technique to SFX might be mitigated by using increased heavy atom concentrations (Hunter *et al.*, 2016; Gorel *et al.*, 2017), but the results of the present work suggest that this is a trade-off in the ultrafast regime, boosting the anomalous signal at the cost of additional damage to the light atom structure. For example, consider the effect of phasing with the Se *K*-edge on global damage. For the Se-doped target, Fig. 9(b) shows that adjusting the X-ray energy from 14 keV to 12 keV is as effective at reducing the ionization of light atoms as compressing the pulse width from 50 fs to 5 fs. Alternatively, this cost might be reduced if phasing is performed with an X-ray frequency well above the absorption edge, as has been done in native phasing experiments (Nakane *et al.*, 2015; Galli *et al.*, 2015c).

There is preliminary evidence that these considerations are relevant in practice. Nass *et al.* attributed unusual noise in the scattering profile of lysozyme.Gd nanocrystals, absent from thaumatin nanocrystals, to increased radiation damage induced by Gd (Nass *et al.*, 2020). Our modeling supports this inference, as it found the addition of just Gd to the target (alongside C, N, O and S) increases the pump-pulse intensity-averaged charge of C, N and O by 25.2%, 26.9% and 26.7%, respectively. However, Fig. 9 suggests that, at the fluence of this experiment [nominally 3.5×10^{12} photons μm^{-2} , though likely lower in reality (Nagler *et al.*, 2017; Nass *et al.*, 2020)], the 7.1 keV photon energy used (just low enough to avoid ionizing the L_3 -edge of the Gd ions) would be close to the optimal choice for minimizing damage in this case. Indeed, for the 1.75×10^{12} photons μm^{-2} fluence pulse simulated in Section 3, increasing the photon energy to 9 keV caused the intensity-averaged charge of protein light atoms to rise from +0.72 to +0.74 for lysozyme.Gd in 0.1 M Gd. For reference, it fell from +0.57 to +0.34 for lysozyme in water.

The obvious limitation of the presented model is its zero-dimensional treatment. Target substructures such as metal

cofactors often have order 10 nm separations, and the large-scale distribution of heavy atoms is generally non-uniform due to the differing compositions of the protein and its aqueous environment. Naively, this suggests that heavy atoms produce a ‘sphere’ of electronic damage in their local region, with the distance spanned dependent on photon energy. However, whether such heterogeneous correlations actually occur is complicated by the non-uniformity in the large-scale solvent-protein structure of real targets. A model fit for exploring this possibility will likely need to break the crystal symmetry and account for spatial variation in electron density on the global scale of the target. Additionally, the effect of heavy atoms suggests a significance to the mother liquor composition in conjunction with electron transfer across the crystal boundary. It is likely, for example, that the high-energy electrons originating in the mother liquor replace those of the crystal to an extent dependent on the crystal size (Caleman *et al.*, 2011a).

Finally, we remark on the impact of the pulse profile on these dynamics. Previous works have generally modeled the temporal intensity profile as either square (Hau-Riege, 2013; Royle *et al.*, 2017) or Gaussian (Royle *et al.*, 2017; Ren *et al.*, 2023). The ionization dynamics proved to be sensitive to this choice (see Fig. S2 of the supporting information). Repeating the lysozyme.Gd simulation (Fig. 2) with a square pulse of the same fluence, energy and FWHM resulted in light-element intensity-averaged charges that were 19% lower than under the Gaussian pulse. This can be attributed to an outsized effect by the earlier EII cascades (Jönsson *et al.*, 2015) induced by the Gaussian pulse during its leading tail. Such cascades have a long period to ionize the target before the bulk of the elastic X-ray scattering. This indicates a necessity for modeling of the dynamics under more realistic SASE pulse profile statistics.

6. Conclusions

The zero-dimensional non-Maxwellian model employed in this study suggests that a significant amount of damage to biological macromolecules under XFEL illumination is seeded by heavy atoms, with even the presence of native sulfur atoms significantly affecting the damage-induced loss of coherence in the scattered wavefield. This result might appear surprising given that such targets only contain heavy elements in trace quantities; however, closer inspection shows this outcome to reasonably follow from two key points: (i) heavier species emit photo- and Auger electrons at much higher rates, considerably boosting the number of secondary ionization cascades instigated within the light atom bulk; (ii) relative to the 10 fs timescale on which the structural signal is captured in XFEL experiments, light-atom-sourced electron avalanches will either have a very low energy and thus dissipate prematurely or have a very high energy and thus a small EII rate; in contrast, avalanches initiated by heavy atoms, with energies between these two extremes, more severely degrade the captured structural signal. The non-Maxwellian treatment of the electron continuum was necessary to capture this latter point.

The addition of heavy atoms to the environment of proteins – such as potassium and sodium ions in the mother liquor – is routine in protein crystallography; however, the results of this work suggest that on the short timescales of XFEL pulses their use becomes a trade-off for additional ionization. Judicious choices to reduce the number of low-intermediate energy primary electron emissions may thus improve experimental outcomes where damage is a concern, or where controlling for damage across pulse parameters is necessary. Specific to *de novo* refinement, anomalous phasing methodologies that allow for weaker anomalous signals would see a reduction in damage-induced noise, suggesting a strength for native phasing over artificial introduction of heavier elements such as Gd. Further, the production of primary electrons near the maximally ionizing energy can be avoided entirely with careful choice of photon frequency.

This work has restricted its broader analysis of the electronic damage landscape to targets where heavy elements make up 1% of the atomic population, but this is far from sufficient to generalize the influence of heavy atoms across the varied ratios seen in real targets, including targets containing multiple heavy elements. However, it is evident that experimental differences generally considered marginal in traditional crystallography can substantially affect the amount of radiation damage suffered by targets in the ultrafast regime. This complexity emphasizes the value of using theoretical modeling to inform SFX experimental design – a role it is already fulfilling (Dickerson *et al.*, 2020) – particularly as a tool for gauging the viability of successful refinement in the high-intensity regime. The zero-dimensional framework employed in *AC4DC* can capably examine the complete electronic damage dynamics across a large number of candidate pulse parameterizations without significant investment of computational resources. For studies concerned with the ions' motions, the simulation may be integrated within a hybrid plasma-MD framework (Kozlov *et al.*, 2020a; Dawod *et al.*, 2024), where delegation of the ultrafast electron dynamics to a zero-dimensional model makes simulating the molecular dynamics of 10–100 nm scale structures feasible.

APPENDIX A

Numerical method

A1. Cubic spline representation of the free-electron energy distribution

The collision kernel on the right-hand side of the Boltzmann equation, equation (1), takes the form of a sum over distinct processes,

$$\begin{aligned} \mathcal{Q}[P_{\xi}, f](\epsilon) = & \mathcal{Q}^{\text{Photo}}[P_{\xi}](\epsilon) + \mathcal{Q}^{\text{Auger}}[P_{\xi}](\epsilon) + \mathcal{Q}^{\text{EII}}[P_{\xi}, f](\epsilon) \\ & + \mathcal{Q}^{\text{TBR}}[P_{\xi}, f](\epsilon) + \mathcal{Q}^{\text{EE}}[f](\epsilon). \end{aligned} \quad (6)$$

The solution of these coupled rate equations presents several numerical challenges:

(1) The primary-ionization terms $\mathcal{Q}^{\text{Photo}}$ and $\mathcal{Q}^{\text{Auger}}$ are essentially Dirac delta-like source terms; such singularities often destabilize numerical PDE solutions.

(2) The number of possible electron configurations scales factorially with the number of electrons in a given atom.

(3) The three-body recombination term \mathcal{Q}^{TBR} is quadratic in the free-electron energy distribution f , leading to worst-case N^3 complexity if the representation of f is N -dimensional.

(4) The electron–electron recombination term \mathcal{Q}^{EE} depends on derivatives of f .

We seek a representation for f that (i) is inherently smooth and at least once differentiable, (ii) is capable of representing strongly peaked ionization functions without Gibbs-like phenomena, and (iii) admits a computationally efficient representation of \mathcal{Q}^{TBR} .

The standard approach (Morgan & Penetrante, 1990) to non-Maxwellian plasma simulation solves the Boltzmann equation using finite differencing of $f(\epsilon, t)$. We take a more general approach, expanding f with respect to a time-invariant basis $\mathcal{B} = \{\phi_i(\epsilon), i = 1, \dots, N\}$, contracted with time-varying expansion coefficients $c^i(t)$ and multiplied by an explicit weight function $w(\epsilon)$,

$$f(\epsilon, t) = w(\epsilon) \sum_i c^i(t) \phi_i(\epsilon). \quad (7)$$

This expansion must be valid both at early times, when f resembles a sum of Dirac delta-functions at the photoelectron and Auger electron energies, and at later times, when it resembles a Maxwell Boltzmann distribution $f_T(\epsilon) \simeq \sqrt{\epsilon} \exp(-\epsilon/k_B T)$. Typical orthogonal bases for $L^2(\mathbb{R})$ (e.g. Legendre polynomials, Fourier sines) are simple to differentiate and suitable for describing the smooth features of a Maxwellian, but suffer from Gibbs' phenomena in the vicinity of the strongly peaked atomic lines. The other typical basis choice – the orthogonal rectangles of finite-element analysis – works well when describing sharp peaks, but is susceptible to numerical instabilities when calculating derivatives (Kho, 1985).

Order- k B-splines ($k = 3$ in this work) are a good compromise between the two approaches. A collection of N splines of order k is defined over a grid of non-decreasing 'knot points' that collectively form the *knot vector* $\mathbf{T} = \{t_j | j = 0, \dots, N + k - 1, t_k \geq t_{k-1}\}$ (de Boor, 1978). An expansion f constructed from order- k B-splines is automatically polynomial away from the knots, and at a knot point of multiplicity m has $k - 1 - m$ continuous derivatives. The knot vector is constructed such that the first and last knots have multiplicity k , while all interior points are distinct. With this choice of knot vector, ϕ_k form a partition of unity – $\forall x \in [t_0, t_N + k - 1], \sum_n \phi_n(x) = 1$.

Such functions are not mutually orthogonal, possessing a non-trivial overlap matrix $S_{ij} := \int d\epsilon w(\epsilon) \phi_i(\epsilon) \phi_j(\epsilon)$. This basis permits an approximate representation of a narrow primary electron peak by a C^1 piecewise polynomial, ensuring that the distribution has a well defined derivative everywhere. Equation (1) may then be recast in a finite form by integrating both sides against a test basis element ϕ_j ,

$$\int d\epsilon \phi_j(\epsilon) \frac{\partial f(\epsilon, t)}{\partial t} = \int d\epsilon \phi_j(\epsilon) \mathcal{Q}[\mathbf{P}, f](\epsilon, t), \quad (8)$$

$$\Rightarrow S_{ij} \frac{dc^i(t)}{dt} = Q_j[\mathbf{P}, f](\epsilon, t), \quad (9)$$

which has the added benefit of rendering the Q tensors sparse.

In our approximation scheme, we assume:

- (1) purely (semi)classical collisional dynamics,
- (2) independent, decoupled atoms,
- (3) spatial homogeneity.

Assumption (2) is equivalent to discarding second-order and higher correlations in the electrons' distribution function.

We separated the time-dependent energy distribution of free electrons $f(\epsilon, t)$ from the discrete probability distribution $P_\xi(t)$ capturing the classical populations of each atomic state ξ . The Boltzmann and master equations (1) and (2) then generate a deterministic time evolution for the classical energy-state distribution of the electrons.

Further numerical methods employed included: (i) an asynchronous implicit-explicit (IMEX) method, stepping the stiff but computationally cheap free-electron interactions with much shorter steps than the bound-bound and bound-free contributions; and (ii) adaptive time steps to avoid divergence in $f(\epsilon, t)$.

A2. Explicit form of numerical couplings

We adopt the following notation:

a : index for atom types, *e.g.* C, N, O,

ξ, η : indices for electron configurations of a particular atom type.

$\Gamma_{\xi \rightarrow \eta}$: transition rate from configuration ξ to configuration η . Square brackets indicate an f -dependent decay rate.

$B_{\xi\eta}$: difference in total binding energy between electron configurations, $B_\xi - B_\eta$.

A2.1. Electron–electron interactions

We model the Coulomb-mediated relaxation of the electron gas using a standard Fokker–Planck kernel, specifically the form quoted by Kho (1985),

$$Q^{ee}[f](\epsilon) = -\frac{\partial J^{ee}[f](\epsilon)}{\partial \epsilon}, \quad (10)$$

where the electron current J^{ee} has the form

$$J^{ee}(\epsilon) = \alpha \left[F(\epsilon, t) \left(\frac{f}{2\epsilon} - \frac{\partial f}{\partial \epsilon} \right) - G(\epsilon, t) f \right], \quad (11)$$

$$F(\epsilon, t) = 2\epsilon^{-1/2} \int_0^\epsilon x f(x, t) dx + 2\epsilon \int_\epsilon^\infty x^{-1/2} f(x, t) dx, \quad (12)$$

$$G(\epsilon, t) = 3\epsilon^{-1/2} \int_0^\epsilon f(x, t) dx, \quad (13)$$

$$\alpha = \frac{2}{3} \pi e^4 (2/m)^{1/2} \ln \Lambda. \quad (14)$$

Note that the electron–electron interaction strength requires an estimate of the Coulomb logarithm $\ln \Lambda = \int d\chi/\chi$ (Morgan & Penetrante, 1990; Royle *et al.*, 2017; Oxenius, 1986). In our

case, we cut the integral off using the Debye length and radius of closest approach $r_0 = e^2/4\pi\epsilon_0 k_B T$, giving $\Lambda = n4\pi\lambda_D^3$ (Oxenius, 1986). The temperature T here represents the notional temperature of the thermalized free electrons, estimated by a dynamical fit to the Maxwellian portion of the distribution.

A2.2. Photoionization and Auger decay

Both processes produce primary electron emission spectra with very narrow linewidths. For numerical stability, we artificially broaden these features into a linear combination of the two basis splines nearest the desired energy, with coefficients chosen such that

$$\int d\epsilon \Delta_E(\epsilon) = 1, \quad \int d\epsilon \Delta_E(\epsilon) \epsilon = E. \quad (15)$$

This approach avoids introducing spurious negative-density parts to the distribution, as would be the case if a naive expansion of $\delta(E - \epsilon)$ were used. This should be a mild approximation in the XFEL regime, as the semi-empirical estimate of Kho (1985) indicates that such broadening occurs on a subfemtosecond timescale,

$$Q_a^{\text{Pht}}[P_a](\epsilon, t) = \sum_{\xi \in C_a} P_\xi(t) \sum_{\eta \neq \xi} \Delta_{\hbar\omega - B_{\xi\eta}}(\epsilon) \Gamma_{\xi \rightarrow \eta}^{\text{Pht}}, \quad (16)$$

$$Q_a^{\text{Auger}}[P_a](\epsilon, t) = \sum_{\xi \in C_a} P_\xi(t) \sum_{\eta \neq \xi} \Delta_{B_{\xi\eta}}(\epsilon) \Gamma_{\xi \rightarrow \eta}^{\text{Auger}}. \quad (17)$$

A2.3. Bound–free interactions

Expressions for Q_a^{EII} and Q_a^{TBR} were derived according to the work of Leonov *et al.* (2014), to which we direct the reader.

A3. Dynamic grid implementation

The cubic spline treatment of the plasma code has previously been shown to well approximate the evolution of relatively static Boltzmann-governed systems with as few as ten energy grid points/knots (Khurana & Thachuk, 2012). However, this requirement grew by an order of magnitude when the initial distribution of particles was modestly (relative to the energy scale of the present work) displaced from equilibrium with a thermal bath. The unbound electron continuum of a biological target under XFEL illumination sees an acute difference between its state at early times and its partially equilibrated state near the end of the pulse. Moreover, these systems see two complications not present in the systems considered by Khurana & Thachuk: (i) the narrow, low-energy peak of the Maxwellian at early times is difficult to fit due to the rigid energy conservation condition of $d^2f/d\epsilon^2 \simeq 0$; (ii) the thermal bath is substituted for sharp, high-energy emission profiles, which, like the Maxwellian, shift significantly over the course of the pulse as they relax through electron interactions (Hau-Riege, 2013).

An adaptive grid was implemented to address these issues. A set of static low-density regions spans the full energy range

and a set of *dynamic* high-density regions spans (i) the Maxwellian and (ii) up to four of the most dominant high-energy peaks. We define a set of energy ranges (regions), each with an associated knot density function $\xi(\epsilon)$ that is only non-zero within the region. The local knot density of the full grid $\Xi(\epsilon)$ is then defined as

$$\Xi(\epsilon) = \max(\xi_1, \xi_2, \dots, \xi_{n-1}, \xi_n). \quad (18)$$

The high-density regions of the electron continuum are dynamically updated throughout the simulation to follow their respective features. A high density of knots supports the sharp features in the continuum at early times, then redistributes to continue to support the growing and shifting peaks as the electron population equilibrates. Testing showed accurate fitting of the continuum at early times to be particularly crucial, so a partial run of the simulation with a ‘guess grid’ is used to obtain the initial grid for the actual simulation. This flexible approach drastically reduced the number of knots necessary to achieve convergence in the evolution of the ionic states.

The spline basis is transformed with a 64-point Gaussian quadrature. Transformations in the simulations of this work were all performed an order of ten times. Testing found the convergence to be independent of the associated error for >100 transformations. The photoelectron peaks were identified dynamically based on their maximum energy density relative to the transition energy region, without regard for the prior basis.

APPENDIX B

Influence of cascade energy with controlled seeding rate

Increasing the X-ray energy further above an absorption edge raises the energy of the seeded EII cascades. At the same time, the number of these cascades decreases due to the reduced

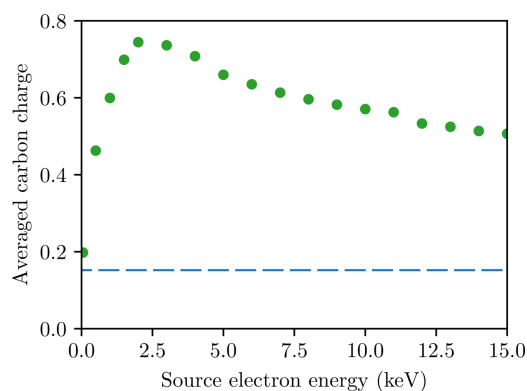


Figure 10 Relationship between the energy of electrons seeded within a solid light-atom target, $C_{613}N_{203}O_{185}$, and the intensity-averaged charge of carbon. Each dot corresponds to a simulation of the light-atom control subjected to a 10^{12} 10 keV photons μm^{-2} , 15.0 fs FWHM pulse, but with an artificial injection of free electrons at the energy denoted by the horizontal axis. Electrons were injected at the same rate across all simulations, proportional to the intensity. Specifically, electrons were added at three times the photoionization rate of the undamaged target at the given intensity. The dashed line indicates the averaged carbon charge when no electrons are injected.

absorption cross-section at higher energies. To isolate the relationship between the energy of a primary electron and the damage caused by the resulting cascade, *AC4DC* was modified to model the ionization of the light-atom control with additional electrons injected at a rate independent of their energy. Each system was simulated under a 15.0 fs FWHM Gaussian pulse with a fluence of 10^{12} 10 keV photons μm^{-2} . The injection rate was fixed at three times the photoionization rate of the *undamaged* target for a given incident intensity, coarsely corresponding to the ionization rate of the Zn-doped target by its Zn ions under the same pulse conditions (Zn was chosen due to the distinct local maximum it produces, as shown in Fig. 6). The resulting relationship between the electron source energy and carbon ionization (Fig. 10) shows strong similarities to the relationship between LPE and carbon ionization in the Zn-doped target shown in Fig. 6(b), including a similar position and magnitude for the local maximum.

APPENDIX C

Model comparisons

C1. Amorphous carbon – *ddcMD*

Hau-Riege (2013) used the MD code *ddcMD* to simulate the ionization of amorphous carbon of density 1 g cm^{-3} under a 2 mJ, 6.0 keV, 10 fs square pulse focused to a 100 nm FWHM focal size. Modeling with *AC4DC* was performed taking a homogeneous intensity equal to the average in the 100 nm focus. This corresponds to a photon fluence of 2.6×10^{14} photons μm^{-2} , an order of magnitude above the maximum fluence considered elsewhere in this work. *AC4DC* saw excellent agreement with *ddcMD* in both the evolution of the carbon ionic states and the free-electron energy distribution. Fig. 11 shows the Maxwellian fits to the thermal electrons at several snapshots in time predicted by *ddcMD*, overlaid on the corresponding normalized electron energy density distributions predicted by *AC4DC*.

C2. Aluminium sheet – *PICLS*

A system corresponding to an infinite plane $1 \mu\text{m}$ thick aluminium foil sheet was simulated by Royle *et al.* (2017) using

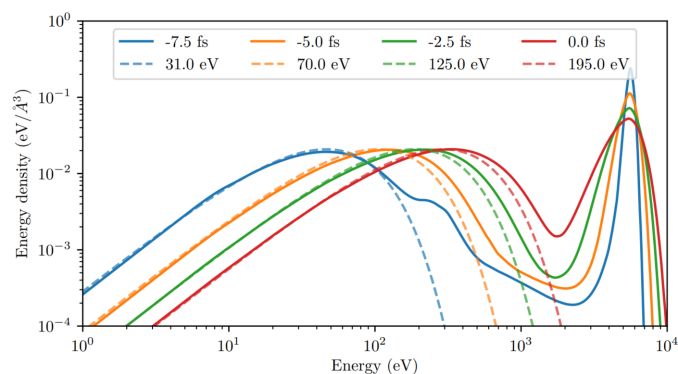
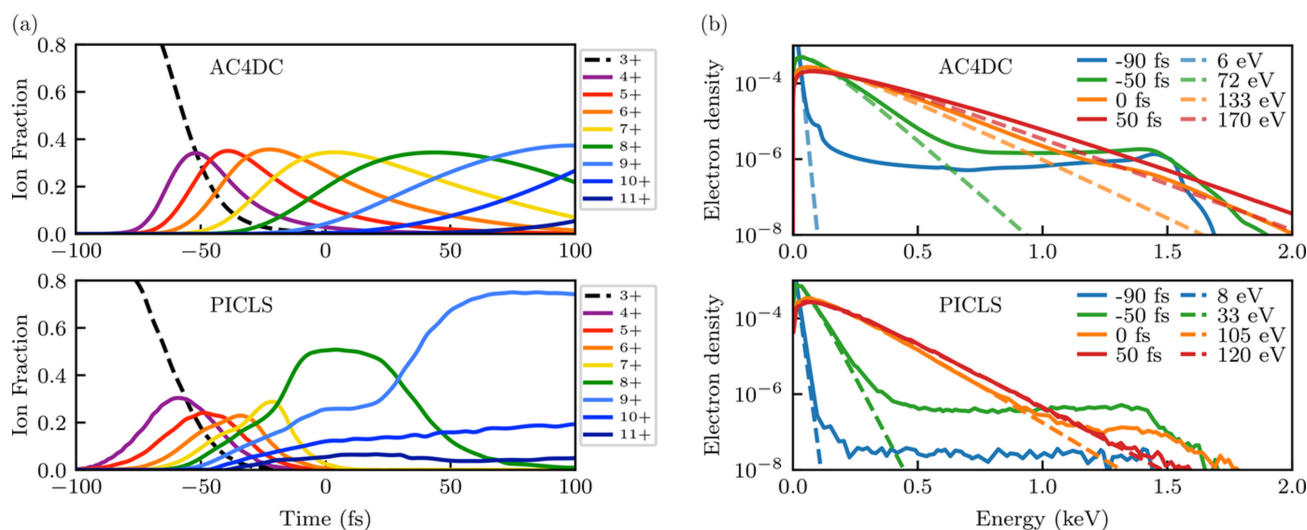


Figure 11 Snapshots of the normalized electron energy density distribution in amorphous carbon. Solid lines are predictions by *AC4DC*, while the dashed lines are the Maxwellian fits by Hau-Riege (2013) based on *ddcMD*.

**Figure 12**

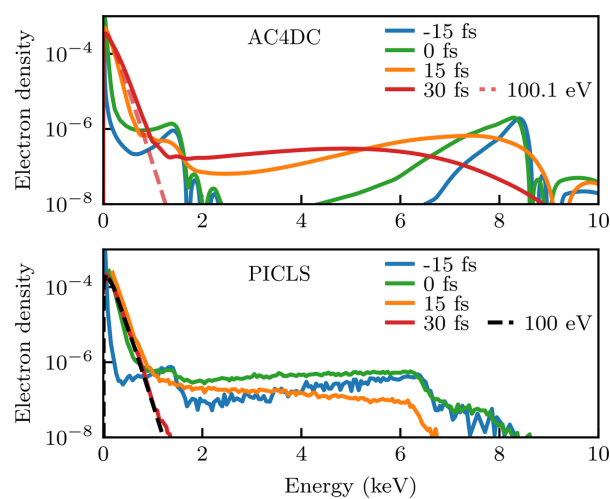
Predictions by *AC4DC* (top) and *PICLS* (Royle *et al.*, 2017) (bottom) for the (a) ion fractions and (b) normalized electron density of an XFEL-driven solid Al plasma. Maxwellian fits to low-energy electrons are shown as dashed lines. Both simulations model peak intensities of $1.36 \times 10^{17} \text{ W cm}^{-2}$. In the *PICLS* simulations, the spatial pulse profile is approximately homogeneous.

the particle-in-cell code *PICLS* under varying conditions. The code tracks individual electrons in space, and determines atomic transitions and impact ionization using Monte Carlo modeling. TBR is incorporated through enforcing an equation of state, which, as highlighted by Royle *et al.* (2017), is less accurate than modeling TBR cross-sections. Unlike *AC4DC*, *PICLS* models ionization potential depression (IPD). IPD is a stronger effect in denser matter, and so would be expected to play a greater role in solid Al than in biological matter. As the EII rate is increased by IPD, it was expected that *AC4DC* would underestimate the rate of secondary ionization.

We consider two scenarios from this study corresponding to quite different pulse parameters. First, a pulse of 1.70 keV photons with a 80 fs FWHM Gaussian profile and peak intensity of $1.36 \times 10^{17} \text{ W cm}^{-2}$. Second, a pulse of 10.0 keV photons with a 20 fs FWHM Gaussian profile and peak intensity of $1 \times 10^{19} \text{ W cm}^{-2}$. Note that, for the first pulse, the photoelectrons are ejected with a very low energy (<150 eV), so the cascades seeded by the 1.4 keV Auger electrons dominate the secondary ionization dynamics.

For the 1.7 keV pulse, the predictions for the evolution of the ion fractions by *AC4DC* and *PICLS*, shown in Fig. 12(a), see relatively good agreement, similar to that observed in the considered study with another collisional-radiative code, *SCFLY*. In comparison with *PICLS*, the ionization rate predicted by *AC4DC* is overall slightly lower. Inspection of the electron density distributions shown in Fig. 12(b) indicates that the electrons thermalize more slowly in the *AC4DC* simulation. This is consistent with a higher rate of collisional-ionization in the *PICLS* model. The different level of thermalization between the models limits the physical meaning of a comparison between temperatures. Similar can be said of the electron density distributions for the 10 keV pulse.

For the 10 keV pulse, the models again make similar predictions for the evolution of the free-electron energy distribution, as can be observed in Fig. 13. There is a

**Figure 13**

Normalized electron densities of Al plasma from *AC4DC* (top) and *PICLS* (bottom) (Royle *et al.*, 2017) simulations under a 10 keV 20 fs FWHM Gaussian pulse. Times are denoted relative to the pulse peak. Both simulations model peak intensities of $1 \times 10^{19} \text{ W cm}^{-2}$. Here, *PICLS* models the radial intensity profile with a super-Gaussian approximately equivalent to a square profile. Dashed lines show the Maxwellian fits to the low-energy electrons at $t = 30 \text{ fs}$ (<2 keV for the *AC4DC* trace).

surprisingly good match to the temperature of the thermal electrons at $t = +30 \text{ fs}$. However, this is again a limited comparison as the low-energy electrons have not completely relaxed to a Maxwellian in the *AC4DC* simulation, while they are nearly completely thermalized in the *PICLS* model. The low-energy electrons remained unrelaxed at $t = +100 \text{ fs}$ in the *AC4DC* simulation (with a best-fit Maxwellian of 213.5 eV).

C3. Glycine crystal – *XMDYN*

Abdullah *et al.* (2017) used the Monte Carlo MD code *XMDYN* to model radiation damage in a glycine ($\text{H}_5\text{C}_2\text{NO}_2$) crystal. The study modeled pulses of 10 keV photons with 10 fs

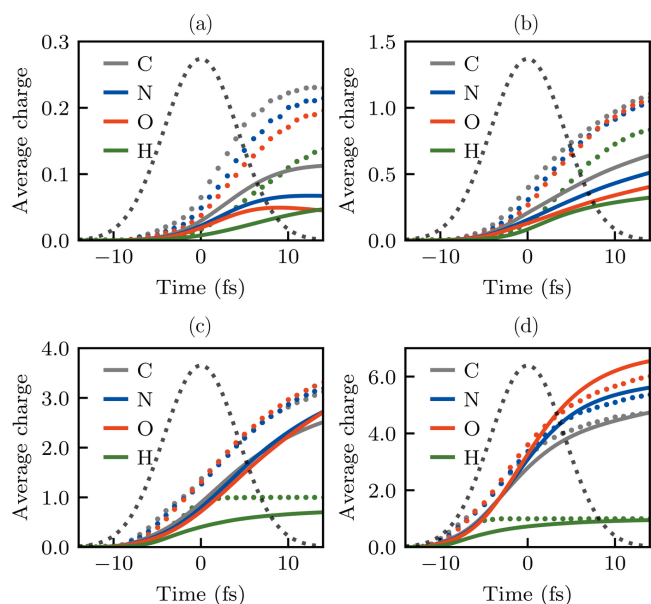


Figure 14 Predictions by *XMDYN* (Abdullah *et al.*, 2017) (dots) and *AC4DC* (lines) for the average charges in the glycine crystal under different peak intensities: (a) $1.5 \times 10^{18} \text{ W cm}^{-2}$, (b) $1.5 \times 10^{19} \text{ W cm}^{-2}$, (c) $1.5 \times 10^{20} \text{ W cm}^{-2}$, (d) $1.5 \times 10^{21} \text{ W cm}^{-2}$.

FWHM Gaussian temporal profiles, at four different peak intensities, between 1.5×10^{18} and $1.5 \times 10^{21} \text{ W cm}^{-2}$. The simulation volume contained 1050 atoms, under periodic boundary conditions. The results of these simulations are shown in Fig. 14. The evolution of charges by element in equivalent *AC4DC* simulations, also shown in Fig. 14, was in good agreement with this modeling for the two higher fluence cases. In the two lower fluence cases, *XMDYN* predicts a higher level of ionization.

C4. Silicon crystal – non-Maxwellian plasma code

We modeled the conditions that a silicon crystal was subjected to in simulations performed by Leonov *et al.* (2014). The considered study used an atomic code that was shown to provide good agreement with the Hartree–Fock method employed in *AC4DC*. The physical model for the plasma dynamics is very similar to *AC4DC*, though the numerical approaches differ significantly. *AC4DC* was used to model the 4 keV and 8 keV pulses considered in this study; for reference, the Si *K*-edge is at 1.8 keV. *AC4DC* predicted a much higher rate of ionization and differing electron density continuums. Results for the 8 keV pulse are shown in Fig. 15.

C5. Lysozyme.Gd – *RADDOSE-XFEL*

RADDOSE-XFEL (Dickerson *et al.*, 2020) is a module of *RADDOSE-3D* (Bury *et al.*, 2018) designed to measure the dose absorbed in XFEL experiments by tracking collisional ionization cascades. Unlike *AC4DC*, the code does not model high-density, strong-ionization effects such as TBR and the depletion of atomic orbitals. Since primary ionization drives the damage-seeding effect described in this study, the code was

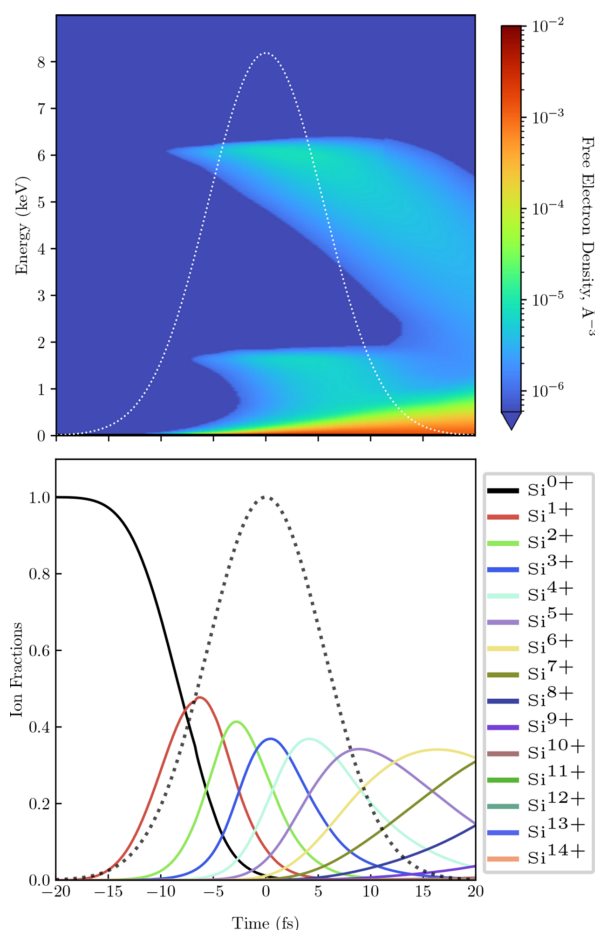


Figure 15 Electron energy density (top) and ion fractions (bottom) in Si under an 8 keV 13 fs FWHM Gaussian pulse of fluence $1.6 \times 10^5 \text{ J cm}^{-2}$. The dotted lines show the pulse profile.

expected to reflect an outsized impact by heavy elements on the absorbed dose.

Table 1 summarizes the damage predictions of *RADDOSE-XFEL* and *RADDOSE-3D* applied to targets equivalent to the lysozyme.Gd target and its toy variants described in Section 3. Values are scaled based on a sample of 2×10^7 individually simulated photons, as described by Dickerson *et al.* (2020). The doses predicted by *RADDOSE-3D* are consistently higher than those predicted by *RADDOSE-XFEL* because *RADDOSE-3D* assumes that the energies of photons are instantaneously absorbed after photoionization (Dickerson *et al.*, 2020). The influence of heavy atoms is underestimated by this approximation; adding sulfur (light control to lysozyme, water) and salt (lysozyme.Gd, 0.1 M Gd to lysozyme.Gd, 0.1 M Gd 10% NaCl) magnifies the dose by a much larger factor when the more accurate *RADDOSE-XFEL* model is used. Notably, accounting for sulfur increases the absorbed dose and average number of non-H ionizations by 49% and 45%, respectively. This effect is stronger than predicted by the *AC4DC* simulations, likely because *RADDOSE-XFEL* does not model the ionic states and so does not capture the suppression of EII that occurs as the target becomes increasingly ionized.

Table 1Predictions of *RADDOSE-XFEL* and *RADDOSE-3D*.

The target column gives the solvent. Note that the right column shows the average number of ionizations for non-H atoms only. Pulse parameters and targets are equivalent to those given in Section 3.1.

Protein	Solvent (35.1% v/v)	<i>RADDOSE-3D</i> dose (MGy)	<i>RADDOSE-XFEL</i> dose (MGy)	<i>RADDOSE-XFEL</i> average non-H ionizations
Light-atom control	Water	1931.4	325.7	1.278
Lysozyme	Water	2338.8	483.6	1.856
Lysozyme.Gd	0.1 M Gd	2923.2	1088.6†	1.859†
Lysozyme.Gd	0.1 M Gd 10% NaCl	3476.6	1304.8†	2.779†

† *RADDOSE-XFEL* assumes instantaneous dose absorption for some Gd ionization processes.

RADDOSE-XFEL predicts that accounting for Gd (lysozyme, water to lysozyme.Gd, 0.1 M Gd) does not increase ionization but increases the dose by an amount comparable with that predicted by *RADDOSE-3D*. This result should be taken as unphysical. *RADDOSE-XFEL* models instantaneous energy deposition similar to *RADDOSE-3D* for many ionization processes of Gd. Indeed, Fig. 4 indicates that the contribution of cascades seeded by Gd to global ionization should be similar to those seeded by Cl in the 10% NaCl solvent.

Acknowledgements

We thank Associate Professor Nadia Zatsopin, Swinburne University of Technology, Hawthorn, Australia, for feedback on the discussions of SFX experimental methods, and Thipayawis Cheunchitra, University of Melbourne, Parkville, Australia, for inspiring examination of the damage landscape above deep absorption edges. Open access publishing facilitated by The University of Melbourne, as part of the Wiley–The University of Melbourne agreement via the Council of Australian University Librarians.

Data availability

Simulation input files for Section 3 may be found at <https://github.com/Phoelionix/AC4DC/tree/master/input/lysozyme>. Input files for Section 4 may be found at https://github.com/Phoelionix/AC4DC/tree/master/input/templates/doped_targets. Additional data used in this work are available upon reasonable request to a corresponding author.

References

- Abbey, B., Dilanian, R. A., Darmanin, C., Ryan, R. A., Putkunz, C. T., Martin, A. V., Wood, D., Streltsov, V., Jones, M. W. M., Gaffney, N., Hofmann, F., Williams, G. J., Boutet, S., Messerschmidt, M., Seibert, M. M., Williams, S., Curwood, E., Balaur, E., Peele, A. G., Nugent, K. A. & Quiney, H. M. (2016). *Sci. Adv.* **2**, e1601186.
- 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.
- Abdullah, M. M., Son, S.-K., Jurek, Z. & Santra, R. (2018). *IUCrJ* **5**, 699–705.

- 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. & Schlichting, I. (2014). *Nature* **505**, 244–247.
- Barends, T. R. M., Foucar, L., Shoeman, R. L., Bari, S., Epp, S. W., Hartmann, R., Hauser, G., Huth, M., Kieser, C., Lomb, L., Motomura, K., Nagaya, K., Schmidt, C., Strecker, R., Anielski, D., Boll, R., Erk, B., Fukuzawa, H., Hartmann, E., Hatsui, T., Holl, P., Inubushi, Y., Ishikawa, T., Kassemeyer, S., Kaiser, C., Koeck, F., Kunishima, N., Kurka, M., Rolles, D., Rudek, B., Rudenko, A., Sato, T., Schroeter, C.-D., Soltau, H., Strueder, L., Tanaka, T., Togashi, T., Tono, K., Ullrich, J., Yase, S., Wada, S., Yao, M., Yabashi, M., Ueda, K. & Schlichting, I. (2013). *Acta Cryst. D* **69**, 838–842.
- Barty, A., Caleman, C., Aquila, A., Timneanu, N., 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., Doak, R. B., 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., Hauser, G., Hirsemann, H., Holl, P., 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. (2012). *Nat. Photon.* **6**, 35–40.
- Bethe, H. A. & Salpeter, E. E. (1957). *Quantum Mechanics of One- and Two-Electron Atoms*. Berlin: Springer.
- Botha, S. & Fromme, P. (2023). *Structure* **31**, 1306–1319.
- Boutet, S., Lomb, L., Williams, G., Barends, T., Aquila, A., Doak, R. B., Weierstall, U., DePonte, D., Steinbrener, J., Shoeman, R., Messerschmidt, M., Barty, A., White, T., Kassemeyer, S., Kirian, R., Seibert, M., Montanez, P., Kenney, C., Herbst, R., Hart, P., Pines, J., Haller, G., Gruner, S., Phillip, H., Tate, M., Hromalik, M., Koerner, L., van Bakel, N., Morse, J., Ghonsalves, W., Arnlund, D., Bogan, M., Calemann, C., Fromme, R., Hampton, C., Hunter, M., Johansson, L., Katona, G., Kupitz, C., Liang, M., Martin, A., Nass, K., Redecke, L., Stellato, F., Timneanu, N., Wang, D., Zatsopin, N., Schafer, D., Defever, K., Neutze, R., Fromme, P., Spence, J., Chapman, H. & Schlichting, I. (2012b). *Hen egg-white lysozyme solved from 40 fs free-electron laser pulse data* <https://doi.org/10.2210/pdb4ET8/pdb>.
- Boutet, S., Lomb, L., Williams, G. J., Barends, T. R. M., Aquila, A., Doak, R. B., 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., Zatsopin, N. A., Schafer, D.,

- Defever, J., Neutze, R., Fromme, P., Spence, J. C. H., Chapman, H. N. & Schlichting, I. (2012a). *Science* **337**, 362–364.
- Bury, C. S., Brooks–Bartlett, J. C., Walsh, S. P. & Garman, E. F. (2018). *Protein Sci.* **27**, 217–228.
- Caleman, C., Bergh, M., Scott, H. A., Spence, J. C. H., Chapman, H. N. & Timneanu, N. (2011b). *J. Mod. Opt.* **58**, 1486–1497.
- Caleman, C., Huld, G., Maia, F. R. N. C., Ortiz, C., Parak, F. G., Hajdu, J., van der Spoel, D., Chapman, H. N. & Timneanu, N. (2011a). *ACS Nano* **5**, 139–146.
- Caleman, C., Jares Junior, F., Grånäs, O. & Martin, A. V. (2020). *Crystals* **10**, 585.
- Caleman, C., Timneanu, N., Martin, A. V., Jönsson, H. O., Aquila, A., Barty, A., Scott, H. A., White, T. A. & Chapman, H. N. (2015). *Opt. Express* **23**, 1213–1231.
- Campbell, J. & Papp, T. (2001). *At. Data Nucl. Data Tables* **77**, 1–56.
- Carrillo, M., Pandey, S., Sanchez, J., Noda, M., Poudyal, I., Aldama, L., Malla, T. N., Claesson, E., Wahlgren, W. Y., Feliz, D., Šraj, V., Maj, M., Castillon, L., Iwata, S., Nango, E., Tanaka, R., Tanaka, T., Fangia, L., Tono, K., Owada, S., Westenhoff, S., Stojković, E. A. & Schmidt, M. (2021). *Structure* **29**, 743–754.e4.
- Chapman, H. N. (2017). *Protein Crystallography: Methods and Protocols* edited by A. Wlodawer, Z. Dauter & M. Jaskolski, Vol. 1606, pp. 295–324. Humana Press Inc.
- Chapman, H. N., Caleman, C. & Timneanu, N. (2014). *Philos. Trans. R. Soc. B* **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., 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., 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., Gorkhovev, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature* **470**, 73–77.
- Circosta, O., Vinko, S. M., Barbrel, B., Rackstraw, D. S., Preston, T. R., Burian, T., Chalupský, J., Cho, B. I., Chung, H. K., Dakovski, G. L., Engelhorn, K., Hájková, V., Heimann, P., Holmes, M., Juha, L., Krzywinski, J., Lee, R. W., Toleikis, S., Turner, J. J., Zastra, U. & Wark, J. S. (2016). *Nat. Commun.* **7**, 11713.
- Dasgupta, M., Budday, D., de Oliveira, S. H. P., Madzelan, P., Marchany-Rivera, D., Seravalli, J., Hayes, B., Sierra, R. G., Boutet, S., Hunter, M. S., Alonso-Mori, R., Batyuk, A., Wierman, J., Lyubimov, A., Brewster, A. S., Sauter, N. K., Applegate, G. A., Tiwari, V. K., Berkowitz, D. B., Thompson, M. C., Cohen, A. E., Fraser, J. S., Wall, M. E., van den Bedem, H. & Wilson, M. A. (2019). *Proc. Natl Acad. Sci. USA* **116**, 25634–25640.
- Dawod, I., Cardoch, S., André, T., De Santis, E. E. J., Mancuso, A. P., Caleman, C. & Timneanu, N. (2024). *Chem. Phys.* **160**, 184112.
- de Boor, C. (1978). *A Practical Guide to Splines*. New York: Springer.
- de la Mora, E., Coquelle, N., Bury, C. S., Rosenthal, M., Holton, J. M., Carmichael, I., Garman, E. F., Burghammer, M., Colletier, J. P. & Weik, M. (2020). *Proc. Natl Acad. Sci. USA* **117**, 4142–4151.
- Dickerson, J. L., McCubbin, P. T. N. & Garman, E. F. (2020). *J. Appl. Cryst.* **53**, 549–560.
- Galli, L., Son, S., Klinge, M., Bajt, S., Barty, A., Bean, R., Betzel, C., Beyerlein, K. R., Caleman, C., Doak, R. B., Duszynski, M., Fleckenstein, H., Gati, C., Hunt, B., Kirian, R. A., Liang, M., Nanao, M. H., Nass, K., Oberthür, D., Redecke, L., Shoeman, R., Stellato, F., Yoon, C. H., White, T. A., Yefanov, O., Spence, J. & Chapman, H. N. (2015c). *Struct. Dyn.* **2**, 041703.
- 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. (2015a). *IUCrJ* **2**, 627–634.
- Galli, L., Son, S.-K., White, T. A., Santra, R., Chapman, H. N. & Nanao, M. H. (2015b). *J. Synchrotron Rad.* **22**, 249–255.
- Gisriel, C., Coe, J., Letrun, R., Yefanov, O. M., Luna-Chavez, C., Stander, N. E., Lisova, S., Mariani, V., Kuhn, M., Aplin, S., Grant, T. D., Dörner, K., Sato, T., Echelmeier, A., Cruz Villarreal, J., Hunter, M. S., Wiedorn, M. O., Knoska, J., Mazalova, V., Roy-Chowdhury, S., Yang, J. H., Jones, A., Bean, R., Bielecki, J., Kim, Y., Mills, G., Weinhausen, B., Meza, J. D., Al-Qudami, N., Bajt, S., Brehm, G., Botha, S., Boukhelef, D., Brockhauser, S., Bruce, B. D., Coleman, M. A., Danilevski, C., Discianno, E., Dobson, Z., Fangohr, H., Martin-Garcia, J. M., Gevorkov, Y., Hauf, S., Hosseinizadeh, A., Januschek, F., Ketawala, G. K., Kupitz, C., Maia, L., Manetti, M., Messerschmidt, M., Michelat, T., Mondal, J., Ourmazd, A., Previtali, G., Sarrou, I., Schön, S., Schwander, P., Shelby, M. L., Silenzi, A., Sztuk-Dambietz, J., Szuba, J., Turcato, M., White, T. A., Wrona, K., Xu, C., Abdellatif, M. H., Zook, J. D., Spence, J. C. H., Chapman, H. N., Barty, A., Kirian, R. A., Frank, M., Ros, A., Schmidt, M., Fromme, R., Mancuso, A. P., Fromme, P. & Zatsepin, N. A. (2019). *Nat. Commun.* **10**, 5021.
- Gorel, A., Motomura, K., Fukuzawa, H., Doak, R. B., Grünbein, M. L., Hilpert, M., Inoue, I., Kloos, M., Kováčová, G., Nango, E., Nass, K., Roome, C. M., Shoeman, R. L., Tanaka, R., Tono, K., Joti, Y., Yabashi, M., Iwata, S., Foucar, L., Ueda, K., Barends, T. R. M. & Schlichting, I. (2017). *Nat. Commun.* **8**, 1170.
- Hau-Riege, S. P. (2013). *Phys. Rev. E* **87**, 053102.
- Hau-Riege, S. P. & Bennion, B. J. (2015). *Phys. Rev. E* **91**, 022705.
- Ho, P. J., Daurer, B. J., Hantke, M. F., Bielecki, J., Al Haddad, A., Bucher, M., Doumy, G., Ferguson, K. R., Flückiger, L., Gorkhovev, T., Iwan, B., Knight, C., Moeller, S., Osipov, T., Ray, D., Southworth, S. H., Svenda, M., Timneanu, N., Ulmer, A., Walter, P., Hajdu, J., Young, L., Maia, F. R. N. C. & Bostedt, C. (2020). *Nat. Commun.* **11**, 167.
- Ho, P. J. & Knight, C. (2017). *J. Phys. B At. Mol. Opt. Phys.* **50**, 104003.
- Hull, J. A., Lee, C., Kim, J. K., Lim, S. W., Park, J., Park, S., Lee, S. J., Park, G., Eom, I., Kim, M., Hyun, H., Combs, J. E., Andring, J. T., Lomelino, C., Kim, C. U. & McKenna, R. (2024). *Acta Cryst.* **D80**, 194–202.
- Hunter, M., Yoon, C. H., DeMirici, H., Sierra, R. G., Dao, E. H., Ahmadi, R., Aksit, F., Aquila, A. L., Ciftci, H., Guillet, S., Hayes, M. J., Lane, T. J., Liang, M., Lundström, U., Koglin, J. E., Mgbam, P., Rao, Y., Zhang, L., Wakatsuki, S., Holton, J. M. & Boutet, S. (2016). *Nat. Commun.* **7**, 13388.
- Hunter, M. S. & Fromme, P. (2011). *Methods* **55**, 387–404.
- Jönsson, H. O., Timneanu, N., Östlin, C., Scott, H. A. & Caleman, C. (2015). *J. Synchrotron Rad.* **22**, 256–266.
- Jurek, Z., Son, S.-K., Ziája, B. & Santra, R. (2016). *J. Appl. Cryst.* **49**, 1048–1056.
- Kanter, E. P., Krässig, B., Li, Y., March, A. M., Ho, P., Rohringer, N., Santra, R., Southworth, S. H., DiMauro, L. F., Doumy, G., Roedig, C. A., Berrah, N., Fang, L., Hoener, M., Bucksbaum, P. H., Ghimire, S., Reis, D. A., Bozek, J. D., Bostedt, C., Messerschmidt, M. & Young, L. (2011). *Phys. Rev. Lett.* **107**, 233001.
- Kho, T. H. (1985). *Phys. Rev. A* **32**, 666–669.
- Khurana, S. & Thachuk, M. (2012). *J. Chem. Phys.* **136**, 094103.
- Kim, Y.-K. & Rudd, M. E. (1994). *Phys. Rev. A* **50**, 3954–3967.
- Kitamura, H. (2019). *J. Electron Spectrosc. Relat. Phenom.* **232**, 45–52.
- Kozlov, A., Gureyev, T. E., Paganin, D. M., Martin, A. V., Caleman, C. & Quiney, H. M. (2020b). *IUCrJ* **7**, 1114–1123.
- Kozlov, A., Martin, A. & Quiney, H. (2020a). *Crystals* **10**, 478.
- Kozlov, A. & Quiney, H. M. (2019). *Phys. Scr.* **94**, 075404.

- Le, H. P., Sherlock, M. & Scott, H. A. (2019). *Phys. Rev. E* **100**, 013202.
- Leonov, A., Ksenzov, D., Benediktovitch, A., Feranchuk, I. & Pietsch, U. (2014). *IUCrJ* **1**, 402–417.
- Martin, A. V. & Quiney, H. M. (2016). *J. Phys. B At. Mol. Opt. Phys.* **49**, 244001.
- Martin, J. & Sawyer, A. (2019). *BioTechniques* **66**, 167–170.
- Morgan, W. L. & Penetrante, B. M. (1990). *Comput. Phys. Commun.* **58**, 127–152.
- Murphy, B., Osipov, T., Jurek, Z., Fang, L., Son, S., Mucke, 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.
- Nagler, B., Aquila, A., Boutet, S., Galtier, E. C., Hashim, A. S., Hunter, M., Liang, M., Sakdinawat, A. E., Schroer, C. G., Schropp, A., Seaberg, M. H., Seiboth, F., van Driel, T., Xing, Z., Liu, Y. & Lee, H. J. (2017). *Sci. Rep.* **7**, 13698.
- Nakane, T., Song, C., Suzuki, M., Nango, E., Kobayashi, J., Masuda, T., Inoue, S., Mizohata, E., Nakatsu, T., Tanaka, T., Tanaka, R., Shimamura, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Iwata, S. & Sugahara, M. (2015). *Acta Cryst.* **D71**, 2519–2525.
- Nass, K. (2019). *Acta Cryst.* **D75**, 211–218.
- Nass, K., Bacellar, C., Cirelli, C., Dworkowski, F., Gevorkov, Y., James, D., Johnson, P. J. M., Kekilli, D., Knopp, G., Martiel, I., Ozerov, D., Tolstikova, A., Vera, L., Weinert, T., Yefanov, O., Standfuss, J., Reiche, S. & Milne, C. J. (2021). *IUCrJ* **8**, 905–920.
- Nass, K., Foucar, L., Barends, T. R. M., Hartmann, E., Botha, S., Shoeman, R. L., Doak, R. B., Alonso-Mori, R., Aquila, A., Bajt, S., Barty, A., Bean, R., Beyerlein, K. R., Bublitz, M., Drachmann, N., Gregersen, J., Jönsson, H. O., Kabsch, W., Kassemeyer, S., Koglin, J. E., Krumrey, M., Mattle, D., Messerschmidt, M., Nissen, P., Reinhard, L., Sitsel, O., Sokaras, D., Williams, G. J., Hau-Riege, S., Timneanu, N., Caleman, C., Chapman, H. N., Boutet, S. & Schlichting, I. (2015). *J. Synchrotron Rad.* **22**, 225–238.
- Nass, K., Gorel, A., Abdullah, M. M. V., Martin, A., Kloos, M., Marinelli, A., Aquila, A., Barends, T. R. M., Decker, F. J., Bruce Doak, R., Foucar, L., Hartmann, E., Hilpert, M., Hunter, M. S., Jurek, Z., Koglin, J. E., Kozlov, A., Lutman, A. A., Kovacs, G. N., Roome, C. M., Shoeman, R. L., Santra, R., Quiney, H. M., Ziaja, B., Boutet, S. & Schlichting, I. (2020). *Nat. Commun.* **11**, 1814.
- Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature* **406**, 752–757.
- Olmos, J. L. Jr, Pandey, S., Martin-Garcia, J. M., Calvey, G., Katz, A., Knoska, J., Kupitz, C., Hunter, M. S., Liang, M., Oberthuer, D., Yefanov, O., Wiedorn, M., Heyman, M., Holl, M., Pande, K., Barty, A., Miller, M. D., Stern, S., Roy-Chowdhury, S., Coe, J., Nagar-atnam, N., Zook, J., Verburg, J., Norwood, T., Poudyal, I., Xu, D., Koglin, J., Seaberg, M. H., Zhao, Y., Bajt, S., Grant, T., Mariani, V., Nelson, G., Subramanian, G., Bae, E., Fromme, R., Fung, R., Schwander, P., Frank, M., White, T. A., Weierstall, U., Zatsepin, N., Spence, J., Fromme, P., Chapman, H. N., Pollack, L., Tremblay, L., Ourmazd, A., Phillips, G. N. Jr & Schmidt, M. (2018). *BMC Biol.* **16**, 59.
- Oxenius, J. (1986). *Kinetic Theory of Particles and Photons: Theoretical Foundations of Non-LTE Plasma Spectroscopy*. Springer Series in Electronics and Photonics. Springer-Verlag.
- Pande, K., Hutchison, C. D. M., Groenhof, G., Aquila, A., Robinson, J. S., Tenboer, J., Basu, S., Boutet, S., DePonte, D. P., Liang, M., White, T. A., Zatsepin, N. A., Yefanov, O., Morozov, D., Oberthuer, D., Gati, C., Subramanian, G., James, D., Zhao, Y., Koralek, J., Brayshaw, J., Kupitz, C., Conrad, C., Roy-Chowdhury, S., Coe, J. D., Metz, M., Xavier, P. L., Grant, T. D., Koglin, J. E., Ketawala, G., Fromme, R., Šrajer, V., Henning, R., Spence, J. C. H., Ourmazd, A., Schwander, P., Weierstall, U., Frank, M., Fromme, P., Barty, A., Chapman, H. N., Moffat, K., van Thor, J. J. & Schmidt, M. (2016). *Science* **352**, 725–729.
- Pandey, S., Bean, R., Sato, T., Poudyal, I., Bielecki, J., Cruz Villarreal, J., Yefanov, O., Mariani, V., White, T. A., Kupitz, C., Hunter, M., Abdellatif, M. H., Bajt, S., Bondar, V., Echelmeier, A., Doppler, D., Emons, M., Frank, M., Fromme, R., Gevorkov, Y., Giovanetti, G., Jiang, M., Kim, D., Kim, Y., Kirkwood, H., Klimovskaia, A., Knoska, J., Koua, F. H. M., Letrun, R., Lisova, S., Maia, L., Mazalova, V., Meza, D., Michelat, T., Ourmazd, A., Palmer, G., Ramilli, M., Schubert, R., Schwander, P., Silenzi, A., Sztuk-Dambietz, J., Tolstikova, A., Chapman, H. N., Ros, A., Barty, A., Fromme, P., Mancuso, A. P. & Schmidt, M. (2020). *Nat. Methods* **17**, 73–78.
- Quiney, H. M. & Nugent, K. A. (2011). *Nat. Phys.* **7**, 142–146.
- Ren, S., Shi, Y., van den Berg, Q. Y., Kasim, M. F., Chung, H., Fernandez-Tello, E. V., Velarde, P., Wark, J. S. & Vinko, S. M. (2023). *Commun. Phys.* **6**, 99.
- Royle, R., Sentoku, Y., Mancini, R. C., Paraschiv, I. & Johzaki, T. (2017). *Phys. Rev. E* **95**, 063203.
- Rudek, B., Son, S., Foucar, L., Epp, S. W., Erk, B., Hartmann, R., Adolph, M., Andritschke, R., Aquila, A., Berrah, N., Bostedt, C., Bozek, J., Coppola, N., Filsinger, F., Gorke, H., Gorkhover, T., Graafsma, H., Gumprecht, L., Hartmann, A., Hauser, G., Herrmann, S., Hirsemann, H., Holl, P., Hömke, A., Journal, L., Kaiser, C., Kimmel, N., Krasniqi, F., Kühnel, K., Matysek, M., Messerschmidt, M., Miesner, D., Möller, T., Moshhammer, R., Nagaya, K., Nilsson, B., Potdevin, G., Pietschner, D., Reich, C., Rupp, D., Schaller, G., Schlichting, I., Schmidt, C., Schopper, F., Schorb, S., Schröter, C., Schulz, J., Simon, M., Soltau, H., Strüder, L., Ueda, K., Weidenspointner, G., Santra, R., Ullrich, J., Rudenko, A. & Rolles, D. (2012). *Nat. Photon.* **6**, 858–865.
- Rudek, B., Toyota, K., Foucar, L., Erk, B., Boll, R., Bomme, C., Correa, J., Carron, S., Boutet, S., Williams, G. J., Ferguson, K. R., Alonso-Mori, R., Koglin, J. E., Gorkhover, T., Bucher, M., Lehmann, C. S., Krässig, B., Southworth, S. H., Young, L., Bostedt, C., Ueda, K., Marchenko, T., Simon, M., Jurek, Z., Santra, R., Rudenko, A., Son, S. K. & Rolles, D. (2018). *Nat. Commun.* **9**, 4200.
- Rudenko, A., Inhester, L., Hanasaki, K., Li, X., Robotjazi, S. J., Erk, B., Boll, R., Toyota, K., Hao, Y., Vendrell, O., Bomme, C., Saveliev, E., Rudek, B., Foucar, L., Southworth, S. H., Lehmann, C. S., Kraessig, B., Marchenko, T., Simon, M., Ueda, K., Ferguson, K. R., Bucher, M., Gorkhover, T., Carron, S., Alonso-Mori, R., Koglin, J. E., Correa, J., Williams, G. J., Boutet, S., Young, L., Bostedt, C., Son, S. K., Santra, R. & Rolles, D. (2017). *Nature* **546**, 129–132.
- Schmidt, M. (2013). *Adv. Condens. Matter Phys.* **2013**, 1–10.
- Seibert, M. M., Ekeberg, T., Maia, F. R. N. C., Svenda, M., Andreasson, J., Jönsson, O., Odić, D., Iwan, B., Rocker, A., Westphal, D., Hantke, M., DePonte, D. P., Barty, A., White, T. A., Gumprecht, L., Coppola, N., Aquila, A., Liang, M., Schulz, J., Martin, A., Caleman, C., Stern, S., Abergel, C., Seltzer, V., Claverie, J., Bostedt, C., Bozek, J. D., Boutet, S., Miahnahri, A. A., Messerschmidt, M., Krzywinski, J., Williams, G., Hodgson, K. O., Bogan, M. J., Hampton, C. Y., Sierra, R. G., Starodub, D., Andersson, I., Bajt, S., Barthelmeß, M., Spence, J. C. H., Fromme, P., Weierstall, U., Kirian, R., Hunter, M., Doak, R. B., Marchesini, S., Hau-Riege, S. P., Frank, M., Shoeman, R. L., Lomb, L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Schmidt, C., Foucar, L., Kimmel, N., Holl, P., Rudek, B., Erk, B., Hömke, A., Reich, C., Pietschner, D., Weidenspointner, G., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Schlichting, I., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K., Andritschke, R., Schröter, C., Krasniqi, F., Bott, M., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B., Chapman, H. N. & Hajdu, J. (2011). *Nature* **470**, 78–81.
- Shi, Y., Ren, S., Chung, H., Wark, J. S. & Vinko, S. M. (2024). *Phys. Plasmas* **31**, 082305.
- Son, S.-K., Boll, R. & Santra, R. (2020). *Phys. Rev. Res.* **2**, 023053.

- Son, S.-K., Chapman, H. N. & Santra, R. (2013). *J. Phys. B At. Mol. Opt. Phys.* **46**, 164015.
- Son, S.-K., Chapman, H. N. & Santra, R. (2011). *Phys. Rev. Lett.* **107**, 218102.
- Spence, J. C. H. (2017). *Struct. Dyn.* **4**, 044027.
- Standfuss, J. (2019). *Curr. Opin. Struct. Biol.* **57**, 63–71.
- Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H. & Shen, J. R. (2015). *Nature* **517**, 99–103.
- Suga, M., Akita, F., Sugahara, M., Kubo, M., Nakajima, Y., Nakane, T., Yamashita, K., Umena, Y., Nakabayashi, M., Yamane, T., Nakano, T., Suzuki, M., Masuda, T., Inoue, S., Kimura, T., Nomura, T., Yonekura, S., Yu, L. J., Sakamoto, T., Motomura, T., Chen, J. H., Kato, Y., Noguchi, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Nango, E., Tanaka, R., Naitow, H., Matsuura, Y., Yamashita, A., Yamamoto, M., Nureki, O., Yabashi, M., Ishikawa, T., Iwata, S. & Shen, J. R. (2017). *Nature* **543**, 131–135.
- Suga, M., Shimada, A., Akita, F., Shen, J., Tosha, T. & Sugimoto, H. (2020). *Biochim. Biophys. Acta* **1864**, 129466.
- Suno, H. & Kato, T. (2006). *At. Data Nucl. Data Tables* **92**, 407–455.
- Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G. J., Koglin, J. E., Oberthuer, D., Heymann, M., Kupitz, C., Conrad, C., Coe, J., Roy-Chowdhury, S., Weierstall, U., James, D., Wang, D., Grant, T., Barty, A., Yefanov, O., Scales, J., Gati, C., Seuring, C., Srajer, V., Henning, R., Schwander, P., Fromme, R., Ourmazd, A., Moffat, K., Van Thor, J. J., Spence, J. C. H., Fromme, P., Chapman, H. N. & Schmidt, M. (2014). *Science* **346**, 1242–1246.
- Vinko, S. M., Ciricosta, O., Cho, B. I., Engelhorn, K., Chung, H., Brown, C. R. D., Burian, T., Chalupský, J., Falcone, R. W., Graves, C., Hájková, V., Higginbotham, A., Juha, L., Krzywinski, J., Lee, H. J., Messerschmidt, M., Murphy, C. D., Ping, Y., Scherz, A., Schlotter, W., Toilekis, S., Turner, J. J., Vysin, L., Wang, T., Wu, B., Zastra, U., Zhu, D., Lee, R. W., Heimann, P. A., Nagler, B. & Wark, J. S. (2012). *Nature* **482**, 59–62.
- Vinko, S. M., Ciricosta, O. & Wark, J. S. (2014). *Nat. Commun.* **5**, 3533.
- Vinko, S. M., Ciricosta, O., Preston, T. R., Rackstraw, D. S., Brown, C. R. D., Burian, T., Chalupský, J., Cho, B. I., Chung, H., Engelhorn, K., Falcone, R. W., Fiokovini, R., Hájková, V., Heimann, P. A., Juha, L., Lee, H. J., Lee, R. W., Messerschmidt, M., Nagler, B., Schlotter, W., Turner, J. J., Vysin, L., Zastra, U. & Wark, J. S. (2015). *Nat. Commun.* **6**, 6397.
- Zhao, F.-Z., Zhang, B., Yan, E. K., Sun, B., Wang, Z. J., He, J. H. & Yin, D. C. (2019). *FEBS J.* **286**, 4402–4417.
- Ziaja, B., Bekx, J. J., Masek, M., Medvedev, N., Lipp, V., Saxena, V. & Stransky, M. (2023). *Philos. Trans. R. Soc. A.* **381**, 20220216.