X-ray atomic scattering factors of low-*Z* ions with a core hole

Stefan P. Hau-Riege*
Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, California 94551, USA
(Received 14 June 2007; revised manuscript received 24 July 2007; published 15 October 2007)

Short and intense x-ray pulses may be used for atomic-resolution diffraction imaging of single biological molecules. One of the dominant damage mechanisms is atomic ionization, resulting in a large fraction of atoms with core holes. We calculated the atomic scattering factor of atoms with atomic charge numbers between 3 and 10 in different ionization states with and without a core hole. Our results show that orbital occupation and the change of the orbitals upon core ionization (core relaxation) have a significant impact on the diffraction pattern.

DOI: 10.1103/PhysRevA.76.042511 PACS number(s): 33.90.+h, 87.15.Aa, 82.53.Ps, 87.59.−e

Ultrafast coherent diffraction imaging of isolated biological moleculas is anticipated to be one of the premier applications of x-ray free electron lasers (XFELs) [1]. In a likely scenario, molecules are injected into the x-ray beam and two-dimensional projections of diffraction patterns are recorded. Patterns that belong to classes of similar orientation are averaged to improve the signal-to-noise ratio, and the class averages are assembled into a three-dimensional diffraction pattern. Finally, the electron density is reconstructed using phase-retrieval algorithms, from which atomic positions may be inferred.

Large x-ray fluences are needed to improve the signal-to-noise ratio due to the low scattering strength of single molecules. The associated increase in damage is overcome by using pulses that are shorter than the relevant damage mechanism. For x-ray energies of 8–12keV, as required for atomic-resolution imaging [2], one of the main damage processes is atomic ionization. For low-*Z* materials, ionization damage is initiated primarily by photoionization of the 1*s* (*K*) shell. A large fraction of the excited ions relax through Auger decay within 5–10 fs [3], emitting electrons of a few hundred eV energy. Since the time scale for atomic deexcitation is comparable to the XFEL pulse length, a large fraction of atoms have single or multiple core holes. It is anticipated that this nonequilibrium ionization state of the molecule has a profound effect on its diffraction pattern. In the simplest approximation, chemical bonding is neglected and the far-field molecular diffraction pattern is taken as the sum of the products of the atomic structure factors for each atomic state and the Fourier transform of the atomic positions.

In this paper we present calculations of the atomic structure factor of atoms in different ionization states with and without a core hole. We begin by considering the ionization dynamics of molecules in biomolecular imaging using XFELs, and we estimate the fraction of atoms that have single or double core holes. We then calculate the scattering factor of light atoms in different ionization states by calculating the electronic wave functions, extracting the electron densities and calculating the elastic scattering amplitude using the nonrelativistic form-factor approximation [4]. For atoms with atomic charge numbers *Z* between 3 and 10, we provide fitting parameters for a simple algebraic expression for the atomic form factor, similar to results reported in [5] for ground-state ions. Finally, we discuss the consequences of these results for biomolecular imaging using XFELs.

We used a continuum dynamics model as described in Ref. [6] to calculate the damage process in a protein molecule of diameter 80 Å. In this model we assume that the sample is a spherically symmetric continuum of matter, with a density of 1.35 g/cm^3 and a composition of $H_{51.6}C_{30.8}N_{8.2}O_{9.4}$. This is similar to the anthrax lethal factor, a particular protein of current interest [7]. Other protein molecules have similar composition. The main damage processes are ionization and Coulomb-force driven atomic motion. The model contains approximate descriptions of the dominant physical processes, including photoionization [8], Auger decay [3], trapping of electrons, Debye shielding, nonuniform collisional ionization [9], and three-body recombination. The contribution of shake-up and shake-off processes associated with photoionization and Auger decay for low-*Z* materials is small and is neglected [10]. We focus the analysis in this section on carbon since carbon is the dominant x-ray-interacting constituent of biomolecules. Figure 1 shows the average number of electrons in the *K* ($\bar{Z}_K$) and *L* ($\bar{Z}_L$) shells of carbon. In these simulations, we assumed a fluence of 3×10^{12} photons/(100 nm)^2 and pulse lengths between 1 and 100 fs. Damage is initiated by *K*-shell photoionization, leading to a decrease of $\bar{Z}_K$ with time. Most photoelectrons es-

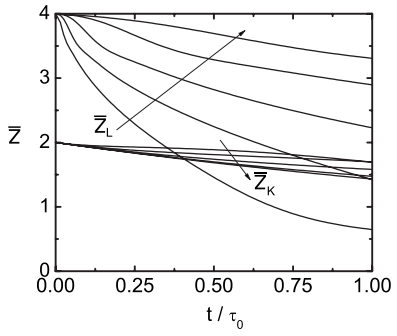

FIG. 1. Average number of electrons in the *K* and *L* shells of carbon as a function of time for different pulse lengths. The arrows indicate lines of decreasing pulse length τ_0=100,30,10,3,1 fs.

*hauriege1@llnl.gov

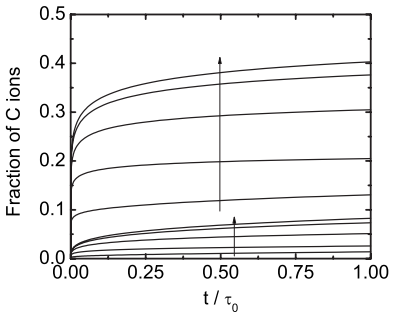

FIG. 2. Fraction of carbon ions with single core holes (upper set of five curves) and double core holes (lower set of five curves). The arrows indicate lines of decreasing pulse length $\tau_0 = 100, 30, 10, 3, 1$ fs.

cape and induce a few secondary ionization events on their way out of the molecule. Photoelectrons are electrostatically trapped later during the pulse when a sufficient charge has accumulated and the kinetic energy of the photoelectrons is smaller than the absolute value of its potential energy, which can only occur for large molecules. Once photoelectrons are trapped, they lead to rapid ionization of the atoms until this process is balanced by its reverse process (three-body recombination). Subsequently the atoms relax through Auger decay, emitting Auger electrons with an energy of a few hundred eV that are typically trapped and lead to secondary ionization events. The number of valence electrons (characterized by $\bar{Z}_L$) quickly decreases with time primarily due to electron-impact ionization and, to a lesser extent, due to Auger decay. Figure 2 shows the fraction of carbon ions with single and double core holes in the K shell that still have electrons remaining in the L shell. Since the pulse lengths can be larger or smaller than the Auger lifetime of carbon of about 10 fs, the fraction of atoms with core holes varies with pulse lengths. In this example, trapping of the photoelectrons does not occur. Figure 2 shows that shorter pulses lead to a larger fraction of atoms with core holes. The fraction of carbon atoms with two core holes is only between 1% and 7% at the end of the pulse, so that their effect on the molecular diffraction pattern is expected to be small. On the other hand, the fraction of carbon atoms with a single core hole quickly increases in time and reaches its maximum between 12% and 40% at the end of the pulse. For such a large fraction of atoms it is important to consider the effect of the modified electron density on the scattering factor. In the following we calculate these scattering factors by first calculating the wave function of ground-state ions and ions with a single core hole, extracting the electron density, and finally calculating the atomic form factor.

We use the multiconfigurational self-consistent-field (MCSCF) theory to calculate the electronic wave function. MCSCF theory is a generalization of the Hartree-Fock (HF) theory to systems dominated by more than one electronic configuration and has been proven to be very useful for describing excited states of molecules [11]. In HF theory, the electronic wave function is approximated by a single configuration of spin orbitals and the energy is optimized with respect to variations of these spin orbitals. In the MCSCF approximation, the wave function is written as a linear combination of configurations of spin orbitals, whose expansion coefficients are optimized simultaneously with the one-electron functions (the molecular orbitals). Since core-hole states are embedded in electronic continua with an infinite number of states with lower energy, they are subject to variational collapse of (i) the orbitals and (ii) the electronic configuration. We use a two-step second-order optimization procedure that prevents the orbital collapse as described in Refs. [12–14]. We use the concept of restricted active space (RAS) [15] to avoid variational collapse in the configurational space by imposing occupancy restrictions for the core orbitals. For a RAS wave function, the orbitals are divided into inactive, active, and secondary orbitals. No restrictions are placed on the occupation of the active orbitals, the inactive orbitals are always doubly occupied, and the secondary orbitals are always unoccupied, subject to the condition that the total number of electrons be conserved. The active orbitals are further subdivided into RAS1, RAS2, and RAS3 spaces. A lower limit is placed on the allowed number of electrons in RAS1 orbitals, and an upper limit is placed on the allowed number of electrons in RAS3 orbitals. No constraints are placed on the occupations of RAS2 orbitals. For single-core-hole calculations, RAS1 contains only the core orbital occupied by exactly one electron, and RAS2 and RAS3 are used for complete and restricted electron distributions.

Molecular orbitals are generated by expansion in a finite set of atomic basis functions. We have employed a standard ANO basis set [16] comprising $14s9p4d3$ which is sufficiently flexible for core hole optimization problems [14]. The calculations were performed using the SIRIUS program system that is part of DALTON [23].

To calculate the electron density ρ from the N-electron wave function Φ, we consider the first-order reduced density matrix γ, which can be written as

$$\gamma(\vec{x},\vec{x}') = N \int \Phi(\vec{x},\vec{x}_2, \ldots ,\vec{x}_N)\Phi^*(\vec{x}',\vec{x}_2, \ldots ,\vec{x}_N)d\vec{x}_2 \cdots d\vec{x}_N. \tag{1}$$

The coordinates $\vec{x}$ represent collectively the spatial coordinates $\vec{r}$ and the spin coordinate σ of an electron. Since γ is Hermitian, it is possible to define an orthonormal basis $\{\eta_i\}$ in which γ is diagonal. It can then be written as

$$\gamma(\vec{x},\vec{x}') = \sum_i \lambda_i \eta_i(\vec{x})\eta_i^*(\vec{x}'). \tag{2}$$

λ_i is called the occupation number of the natural spin orbital η_i in the wave function Φ [17]. The diagonal element of γ is the density of the electrons, $\rho(\vec{x}) = \gamma(\vec{x},\vec{x})$. The electron density as a function of the spatial coordinates only is obtained by integration over the spin coordinate, $\rho(\vec{r}) = \int \gamma(\vec{x},\vec{x})d\sigma$.

Figure 3 shows the radial dependence of the electron density broken down into the natural orbitals $1s$, $2s$, and $2p$ for neutral carbon and for a carbon ion with a single $1s$ core hole. It can be seen that the electron density of the $1s$ orbital is reduced and the $2s$ and $2p$ orbitals are contracted toward the center of the atom due to core relaxation. We performed similar calculations using a simpler HF code [18] which is in

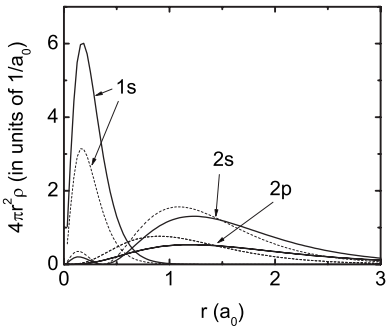

FIG. 3. Electron density of the $1s$, $2s$, and $2p$ subshells for neutral carbon (solid lines) and for carbon with a single core hole (dashed lines). a_0 is the Bohr radius.

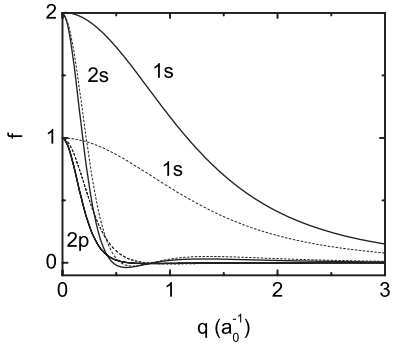

FIG. 4. Atomic scattering factor the $1s$, $2s$, and $2p$ subshells in neutral carbon (solid lines) and in carbon with a single core hole (dashed lines).

principle less precise than MCSCF since it uses only a single configuration of spin orbitals. We found, however, that the HF results agree within 2% with the MCSCF results.

We now consider the case of coherent elastic photon scattering from isolated atoms or ions. In this case, no energy is transferred to the atomic system and its internal state does not change. The photon energy is assumed to be sufficiently low so that nuclear Thompson and nuclear resonance scattering can be neglected. For light elements, this is the case for photon energies below 100 keV [4]. In this case the elastic scattering amplitude is dominated by elastic scattering from bound electrons (Rayleigh scattering). The amplitude for Rayleigh scattering in the nonrelativistic form-factor approximation is $-r_0(\vec{\epsilon}_i\cdot\vec{\epsilon}_f)f(\vec{q})$ [19], where $\vec{\epsilon}_i$ and $\vec{\epsilon}_f$ are the polarization vectors of the incoming and scattered photons, respectively, r_0 is the classical electron radius, and $\vec{q}$ is the wave-vector transfer. The form-factor approximation is valid in the limit of photon energies much larger then electron binding energies (typically less than 1 keV in light atoms) and small momentum transfer. In this case the electrons can be considered as being loosely bound to the atom. The atomic form factor $f(\vec{q})$ is given by the Fourier transform of the electron charge number density,

$$f(\vec{q}) = \int \rho(\vec{r})e^{-i\vec{q}\cdot\vec{r}}d\vec{r}. \tag{3}$$

If $\rho(\vec{r})$ is spherically symmetric, which we assume in the following calculations, then f can be written as

$$f(q) = 4\pi\int_0^\infty \rho(r)r^2\,\mathrm{sinc}(qr)dr \tag{4}$$

with $\mathrm{sinc}(x)\equiv\sin(x)/x$. In the form-factor approximation, the bound electrons are treated as a continuous charge distribution with Thompson-like scattering from each charge element. In the special case of forward scattering, $f(0)=N$, where N is the number of electrons in the atom. We neglect many-electron correlation effects, usually referred to as the independent particle approximation, which is a good approximation when the photon energy is larger than the photoeffect threshold. For experiments at XFELs we are primarily interested in photon energies between 8 and 14 keV, so that Eq. (3) adequately describes the amplitude of the Rayleigh scattering off isolated light atoms and ions.

$f(\vec{q})$ can be written as the sum of the atomic form factors f_n of the occupied natural orbitals n. Figure 4 shows f_n for neutral carbon and carbon with a single core hole. $f_n(0)$ equals the natural orbital occupation number. With increasing q, f_n drops once $1/q$ is smaller than the size of the subshell. The electrons of all subshells contribute similarly to the atomic scattering factor for small q, whereas for larger q only the inner subshells contribute significantly. It can also be seen that core ionization does not only affect f_{1s} but also f_{2s} and f_{2p} due to core relaxation. This effect is most pronounced around $0.25a_0^{-1}$, which falls into a regime that is important for biomolecular imaging using XFELs.

Figure 5 shows the total atomic scattering factor f of carbon in different ionization states with and without a core hole. This figure indicates the importance of the detailed subshell occupancy for $q \geqslant 0.1a_0^{-1}$. Figure 5 also shows the effect of core relaxation on f. Calculations using the Hartree-Fock method without core relaxation typically lead to smaller values of f than more precise MCSCF calculations that take core relaxation into account.

We have performed similar calculations for other low-Z elements with $3 \leqslant Z \leqslant 10$. For computational convenience, the calculated form factors have been fitted by the analytical expression

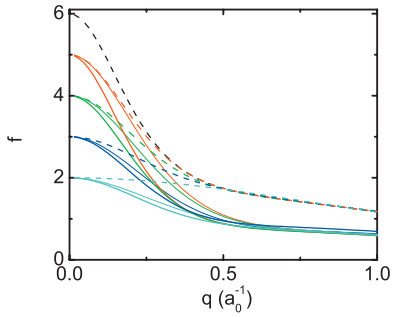

FIG. 5. (Color) Total atomic scattering factors of carbon with $N=6$ (black), $N=5$ (red), $N=4$ (green), $N=3$ (blue), and $N=2$ (light blue). The dashed lines correspond to ions in the ground state, the thick solid lines to ions with a core hole without core relaxation, and the thin solid lines to ions with a core hole with core relaxation.

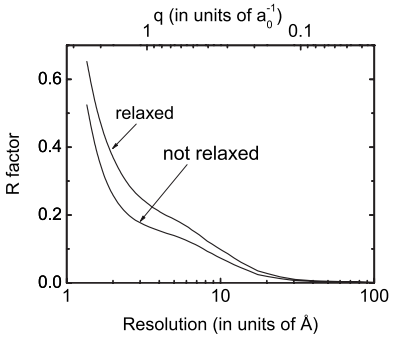

FIG. 6. Residual R factor as a function of resolution length for a protein molecule with and without core relaxation. For atoms with two core holes, atomic form factors without core relaxation were used.

$$f(q) = c + \sum_{i=1}^{5} a_i + e^{-b_i q^2}, \tag{5}$$

following the procedure described in [5]. The fitting parameters for low-Z elements in their various ionization states with and without a core hole are available from EPAPS [24].

For biomolecular diffraction imaging using short x-ray pulses, it is anticipated that the diffraction pattern is recorded up to a scattering angle θ of about 22° [1], where 2θ is the angle between the wave vectors of the incoming and the scattered light. For 8-keV photons, as will be provided by the LCLS [20], this corresponds to wave-vector transfers of up to $q = 4\pi \sin(\theta)/\lambda = 1.6a_0^{-1}$. For 12-keV photons, as will be provided by the Euro XFEL [21], this corresponds to $q = 4.7a_0^{-1}$. For $0.1a_0^{-1} \leqslant q \leqslant 0.6a_0^{-1}$, the atomic scattering factor strongly differs for core-hole ions with and without core relaxation. To demonstrate the significance of this effect on imaging, we have calculated the time-integrated intensity of the scattered x rays of a molecule undergoing damage using atomic scattering factors (i) without and (ii) with core relaxation. We use the residual R factor [1] to quantify the degree of image degradation for each of these two cases. Figure 6 shows the R factor as a function of resolution length. R generally increases at finer resolutions. If core relaxation is taken into account, significantly larger R factors are obtained than without core relaxation. In these calculations we neglect chemical bonding and the deviation of the atomic scattering factor from spherical symmetry. The errors associated with these effects may further affect the diffraction pattern [22].

In summary, we have calculated the atomic scattering factors of low-Z ions with and without a core hole and fitted simple algebraic expressions to the form factors. In the context of biomolecular diffraction imaging using short x-ray pulses, we have found that core relaxation has a significant effect on the diffraction pattern of molecules.

The author would like to thank Hans Agren, Abraham Szoke, and Mau Chen for helpful discussions. This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

[1] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, Nature (London) **406**, 752 (2000).
[2] S. Hau-Riege, R. London, G. Huldt, and H. N. Chapman, Phys. Rev. E **71**, 061919 (2005).
[3] E. J. McGuire, Phys. Rev. **185**, 1 (1969).
[4] L. Kissel, B. Zhou, S. Roy, S. S. Gupta, and R. Pratt, Acta Crystallogr., Sect. A: Found. Crystallogr. **51**, 271 (1995).
[5] D. Waasmaier and A. Kirfel, Acta Crystallogr., Sect. A: Found. Crystallogr. **51**, 416 (1995).
[6] S. P. Hau-Riege, R. A. London, and A. Szoke, Phys. Rev. E **69**, 051906 (2004).
[7] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, and P. Bourne, Nucleic Acids Res. **28**, 235 (2000).
[8] D. A. Verner, D. G. Yakovlev, I. M. Band, and M. B. Trzhaskovskaya, At. Data Nucl. Data Tables **55**, 233 (1993).
[9] M. A. Lennon, K. L. Bell, H. B. Gilbody, J. G. Hughes, A. E. Kingston, M. J. Murray, and F. J. Smith, J. Phys. Chem. Ref. Data **17**, 1285 (1988).
[10] *Ionization and Transition Probabilities*, Vol. I of *Atomic Inner-Shell Processes*, edited by B. Crasemann (Academic Press, New York, 1975).
[11] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, New York, 2000).
[12] H. J. A. Jensen, P. Jorgensen, and H. Agren, J. Chem. Phys. **87**, 451 (1987).
[13] A. N. de Brito, S. Svensson, N. Correia, and H. Agren, J. Chem. Phys. **95**, 2965 (1991).
[14] H. Agren and H. J. A. Jensen, Chem. Phys. **172**, 45 (1993).
[15] J. Olsen, B. O. Roos, P. Jorgensen, and H. J. A. Jensen, J. Chem. Phys. **89**, 2185 (1988).
[16] P. O. Widmark, P. A. Malmquist, and B. O. Roos, Theor. Chim. Acta **77**, 291 (1990).
[17] P.-O. Loewdin, Phys. Rev. **97**, 1474 (1955).
[18] F. Herman and S. Skillman, *Atomic Structure Calculations* (Prentice-Hall, Englewood Cliffs, NJ, 1963).
[19] W. Franz, Z. Phys. **95**, 652 (1935).
[20] J. Arthur, K. Bane, V. Bharadwaj, G. Bowden, R. Boyce, R. Carr, J. Clendenin, W. Corbett, M. Cornacchia, T. Cremer, P. Emma, A. Fasso, C. Field, A. Fisher, R. Gould, R. Hettel, J. Humphrey, K. Ko, T. Kotseroglou, Z. Li, D. Martin, B. McSwain, R. Miller, C. Ng, H.-D. Nuhn, D. Palmer, M. Pietryka, S. Rokni, R. Ruland, J. Schmerge, J. Sheppard, R. Tatchyn, V. Vylet, D. Walz, R. Warnock, H. Winick, M. Woodley, A. Yeremian, R. Yotam, F. Zimmermann, C. Pellegrini, J. Rosenzweig, L. Bertolini, K. van Bibber, L. Griffith, M. Libkind, R. Moore, E. T. Scharlemann, S. Schriber, R. Sheffield, W. Fawley, K. Halback, D. Humphries, K.-J. Kim, S. Lidia, R. Schlueter, M. Xie, A. Freund, D. Meyerhofer, L. Serafini, T. Limberg (unpublished).

[21] F. Richard, J. R. Schneider, D. Trines, and A. Wagner (unpublished).

[22] M. Pirenne, *The Diffraction of X Rays and Electrons by Free Molecules* (Cambridge University Press, London, 1946).

[23] DALTON, a molecular electronic structure program, release 2.0, 2005; see http://www.kjemi.uio.no/software/dalton/dalton.html

[24] See EPAPS Document No. E-PLRAAN-76-121709 for fitting parameters for the atomic scattering scattering factors for $3 \leqslant Z \leqslant 10$. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.