Supporting Information for:

Prospect of Retrieving Vibrational Wave function by Single-Object Scattering Sampling

Hosung Ki,[†,‡,⊥] Kyung Hwan Kim,[†,‡,⊥] Jeongho Kim,[§] Jae Hyuk Lee,[†] Joonghan Kim,[∥] and Hyotcherl Ihee*[,†,‡]

[†] Department of Chemistry, KAIST, Daejeon 305-701, Republic of Korea

[‡] Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 305-701, Republic of Korea

[§] Department of Chemistry, Inha University, Incheon 402-751, Republic of Korea

[∥] Department of Chemistry, The Catholic University of Korea, Bucheon 420-743, Republic of Korea

Contents:

A. Assumptions

For simplicity of the SOSS simulation presented in this work, we made the following assumptions: (1) independent atom model, which assumes that the electron density distribution of an iodine molecule can be expressed as a sum of two atomic electron density distribution; (2) the redistribution of electron density induced by ionization is very fast that the electron density distribution is static (or near equilibrium) during the sampling step size (0.1 fs in this work) in the simulation of ionization kinetics; and (3) the atomic positions in the molecule are static as well during the sampling step size. Besides, we did not consider any noise effect other than Poisson noise although other types of noise may severely reduce the quality of reconstructed wave function.

B. Characteristics of X-ray pulse

In the simulations performed in this work, the temporal intensity profile of X-ray pulse was assumed to be Gaussian that has the temporal duration of w (FWHM in fs unit) and is centered at w. The simulation was performed in the time range from zero to $(2w - dt)$ with the time step of dt, where dt was set to be 0.1 fs, thus covering nearly the entire temporal profile of the pulse. It was assumed that X-ray is circularly polarized, perfectly monochromatic at the wavelength of 1 Å, and focused on a circular spot with 3 nm diameter.

C. Ionization model

To obtain the vibrational wave function using SOSS, it is critical to extract a correct bond length from each single-molecule scattering image with the spatial resolution of smaller than a few picometers, which correspond to the length scale of typical molecular vibrations. The challenge lies in that the strong femtosecond X-ray pulse used for the SOSS experiment can easily destroy the molecule by ionization and Coulomb explosion[1-2] and thus affects the resultant scattering pattern. In fact, according to our simulation of the interactions between I_2 molecules and X-ray pulses of various pulse durations, the femtosecond X-ray pulse induces severe structural change of the I_2 molecule (see Figure S3).

To account for this effect of sample damage on the scattering pattern, we included a model describing the ionization induced by X-ray radiation in our simulation. Although it is possible to perform an elaborate simulation considering the dynamic change of the scattering cross section in all of the electronic states, such simulation would take an immense amount of

calculation time.[3] Instead, we used a simplified model for the interaction between the photon and electrons. Using the ionization model, we estimated the change in the number of electrons in an atom as a function of time and reflected this information to the dynamic form factors used to calculate the single-molecule scattering image.

A simplified model was used to describe ionization induced by X-ray radiation as schematically shown in Figure S1. In this simplified model, we consider only three energy levels: 1) K-shell, which cannot be ionized, 2) L_1-shell, the only level that can be directly ionized by the absorption of X-ray photons, and 3) higher shells, which is a single energy level approximating all other shells with higher energies than the L_1-shell. We assumed that only the L_1-shell is ionized directly by the absorption of X-ray photons based on the fact that the L_1-shell has the highest absorption cross section among the states in the energy range of our interest.[4] This assumption was made for a simple description of ionization, but might not be the accurate one. Still, it fits the need of this work that is focused on examining the feasibility of the SOSS experiment rather than the exactness of the ionization model itself. Upon ionization of the L_1-shell, a hole generated in the L_1-shell relaxes to the higher shells (or an electron in the higher shells relaxes to the L_1-shell) and induces the ejection of another electron in the higher shells via an Auger process. In other words, the higher shells serve as a sink for the relaxation of L_1-holes as well as a source of electrons for the Auger process. We assumed that the L_1-hole relaxes to higher shells with a time constant of 0.3 fs, which is deduced from the inverse of its natural line width.[5] In reality, as more electrons are ejected via many steps of ionization processes, the increased positive charge of the atom may hinder additional ionization. In this model, however, the influence of positive charge is ignored. Therefore, this model is not likely to underestimate the rate of ionization and represents the worst-case scenario in terms of the number of ejected electrons.

Generally, the electrons ejected off an atom are trapped by positive nuclear charge of other atoms in a molecule. Consequently, secondary ionization occurs by collision of the electron with the electron-trapping atom, stimulating sequential ionization processes in the entire molecule.[6] However, in a diatomic molecule, the overall positive charge generated by the ionization processes is much smaller than in large molecules such as proteins. As a result, an electron ejected from one atom has much smaller probability of being trapped by the other atom to cause secondary ionization. Therefore, we did not consider trapped electrons and secondary ionization in our ionization model.

Based on this ionization model, we performed a series of simulations at various time points as follows. At each time point of the simulation, the ionization probability of the L_1-shell was calculated using the following equation:

$$p = n_p \times (1 - \exp(-\sigma / (\pi r_X^2))) \tag{S1}$$

where p is the probability of ionization of the L_1-shell electron at each time step, n_p is number of incident photons at each time step, σ is the absorption cross-section of electron in the L_1-shell, and r_X is the radius of X-ray focal spot. We assumed that the absorption cross section σ of the L_1-shell electron is a constant, 2.5×10^{-4} Å^2,[4] and not affected by the extent of ionization. Each electron in the L_1-shell is ejected from the atom by the absorption of an X-ray photon with the probability p. Once a hole is generated by the ionization of the L_1 shell, it relaxes to higher shells (that is, decay of L_1-hole) with the probability of (dt / 0.3 fs) as long as there is any electron in the higher shells to be exchanged with the hole. When the hole decays to the higher shells, another electron in the higher shells is ionized via an Auger process. The simulation of these ionization processes (that is, ionization of the L_1 shell and subsequent ionization of the higher shells) were repeated while changing the time points through the temporal profile of X-ray pulse, providing the time-dependent electron density of each atom.

D. Calculation of scattering pattern

For the calculation of the scattering patterns, we defined the detector parameters as follows. First, we assumed that the detector is of spherical shape. The source of scattering (that is, the sample) is positioned at the center of the detector sphere. Second, the pixels are symmetric in the horizontal and vertical directions and, therefore, the separation angle ($d\varphi$) between two adjacent pixels in the azimuthal angle (φ) direction is the same to the separation angle ($d\theta$) in the polar angle (θ) direction. Third, the maximum values of both azimuthal and polar angles were set to be 60°, which corresponds to the maximum $q = 6.28$ Å^{-1}. Finally, the size of the entire array of pixels was set to be 30 × 30 (that is, 30 pixels in the azimuthal direction and 30 pixels in the polar direction).

We followed a simple independent atom model[7-8] to calculate the X-ray scattering patterns. With the independent atom model and the assumption of circular polarization of the incident X-ray beam, we used the following equation for calculating scattering images:[9]

$$I(\mathbf{u},\Omega) = 1/2(1+\cos^2 2\theta)\Omega r_e^2 \int I(t) \mid \sum_j f_j(t)\exp\{i\Delta\mathbf{k}(\mathbf{u})\cdot\mathbf{x}_j(t)\mid^2 \, dt \tag{S2}$$

where $\mathbf{u}$ is the position vector of a detector pixel, Ω is the solid angle of the detector pixel, θ is a half of the scattering angle corresponding to the pixel, r_e is the classical radius of an electron, $I(t)$ is the instant photon flux at time t, $f_j(t)$ is the instant form factor of the jth atom at time t, $\Delta\mathbf{k}(\mathbf{u})$ is change in the scattering vector, and $\mathbf{x}_j(t)$ is the position vector of the jth atom at time t. This equation calculates the number of elastically scattered photons, and incoherent scattering is ignored in this calculation. To calculate the scattering intensity detected at each pixel, we considered the intensity only at the center of the pixel and assumed that the intensity is equal at all different positions within a pixel. The instant form factor of an atom was calculated only with the electrons left in an atom at an instant. For example, if an iodine atom (I, $Z = 53$) loses an electron, its instant form factor was assumed to be the same as that of neutral tellurium (Te, $Z = 52$). The number of electrons left in each iodine atom at an instant (that is, instant electron density) was calculated from the simulation using the ionization model described in Section C.

When calculating the scattering intensity, we considered the instant interatomic distance between two iodine atoms, which was simulated based on the instant electron density obtained from the ionization simulation described above. The interaction between the two atoms was described by a potential as a function of interatomic distance, and the instant interatomic force was calculated by differentiation of the potential. The potential is described by a sum of Coulomb potential and Morse potential, of which the latter is as follows:

$$V(r) = D_e(1-\exp(-a(r-r_e)))^2 \tag{S3}$$

where $a = 1.9$ Å^{-1} $r_e = 2.67$ Å, $D_e = 150.03$ kJ/mol. The instant interatomic distance between the iodine atoms was calculated at every time point of the simulation using the interatomic force obtained from the potential.

In this work, we calculated (instead of measuring) experimental scattering patterns from molecules of various bond lengths and molecular orientations. To do so, a set of parameters for bond length (r) and molecular orientation (θ, φ) were randomly generated following a well-defined reference probability density function, $P(r, \theta, \varphi)$. For a set of simulation, we calculated a total of 10,000 experimental scattering patterns.

In fact, Eq. S2 calculates the expectation value for the number of photons scattered to the detector pixel. When simulating the experimental scattering patterns, since the number of photons measured by a detector pixel must be an integer, this expectation value was converted to an integer using probability mass function based on Poisson distribution as follows:

$$f(X = k; \overline{X} = \lambda) = \lambda^k \cdot e^{-\lambda} \Big/ k! \tag{S4}$$

This equation is a probability mass function in terms of stochastic variable X, which has the expectation value of λ. However, when calculating the scattering patterns constituting the reference library, we simply used the expectation value for each scattering pattern without converting to an integer. Considering that each scattering pattern in the reference library is an average of 50 scattering patterns (to account for the random characters of ionization and scattering phenomena), it is desirable to use the expectation value, which is an average number of photons scattered to the detector pixel, rather than an integer.

E. Making reference library and mapping

The key step of SOSS is to extract structural parameters from a single-shot, single-molecule scattering image measured experimentally. There already exist many methods that can convert the scattering pattern in q-space to the molecular structure in real space, for example, genetic algorithm[10] and phase retrieval.[11-13] However, phase retrieval method has rather low spatial resolution and is not suitable for extracting atomic-scale structural information required for obtaining vibrational wave function. Also, femtosecond X-ray pulse of high photon flux still causes the sample damage by ionization, especially severe for small molecules. Therefore, it is uncertain whether the conventional phase-retrieval methods can be directly applied to SOSS.

To circumvent these problems and retrieve correct structural parameters from a single-shot, single-molecule scattering image, we adopted an approach of correlation mapping as mentioned in the main text. In this method, a reference library is constructed, which is a collection of scattering images calculated for all the possible structures and orientations of the molecule, and an experimental scattering image is mapped into the reference library using a correlation function.

When making a reference library, we calculated scattering images by considering the dynamic form factor and the dynamic bond length of the I_2 molecule, as mentioned in Section D, so that we can account for the effect of ionization. In particular, the dynamic form factor was used to account for ionization induced by X-ray radiation, and the dynamic bond length was used to address Coulomb repulsion between ionized atoms. Because both ionization and scattering are quantum mechanical phenomena and thus occur randomly, even molecules with the exactly same bond lengths and angular orientations can still give rise to slightly different single-shot, single-molecule scattering images. However, if we know the exact kinetics of ionization, a statistically-averaged reference scattering image can be precisely calculated.[14]

Each element of the reference library was prepared by (1) repeatedly calculating single-shot, single-molecule scattering images for a molecule with a given set of bond length and angular orientation (r, θ, φ) and (2) averaging a total of 50 calculated scattering images to obtain the most probable scattering image. The 50 single-molecule images are slightly different from each other because the ionization rate of atoms and the Poisson noise of scattered photons are different for each image due to random nature of ionization and scattering events. A reference library was constructed by calculating the averaged single-shot, single-molecule scattering images (that is, most probable scattering image) while varying (r, θ, φ). When making a one-dimensional reference library spanning r-space, the bond length r was changed from 2.3 to 3.0 Å with 0.01 Å step size while θ and φ were fixed at 0. When making a two-dimensional reference library spanning (r, φ)-space, the bond length r was changed from 2.3 to 3.0 Å with 0.02 Å step size and φ was varied from 0 to 360° with the step size of 1° while θ was fixed at 30°. When making another two-dimensional reference library spanning (r, θ)-space, the bond length r was changed from 2.3 to 3.0 Å with 0.02 Å step size and θ was varied from 0 to 180° with the step size of 1° while φ was fixed at 0.

We can derive a structural parameter of the I_2 molecule, I–I bond length, by mapping each experimental scattering image into the reference library. In principle, the mapping procedure is aimed at finding, in the reference library, the image that is most similar to the experimental one. To find such image, we defined a correlation value, c, between a pair of experimental image and an element of the reference library as follows:

$$c = (\sum_{j=1}^{N} I_j \cdot I'_j) / ((\sum_{j=1}^{N} I_j^2) \cdot (\sum_{j=1}^{N} I'_j{}^2))^{1/2}$$
(S5)

where I_j is the intensity of the jth pixel of the experimental scattering image, I'_j is the intensity of the jth pixel of the reference image, and N is the total number of pixels in a scattering image. As the two scattering images become more similar to each other, the value of c becomes higher until it reaches the maximum of 1. The bond length of the I_2 molecule captured in the experimental image is determined from the reference image that gives the highest c value.

F. Effect of X-ray pulse parameters on the sample damage

One of the biggest challenges for the realization of the SOSS experiment is that strong femtosecond X-ray pulse can easily destroy the molecule by ionization and Coulomb explosion, especially severe for small molecules. As a result, X-ray scattering pattern may be distorted significantly by the change of molecular structure during light-matter interaction between the X-ray pulse and the sample molecule. In this section, we investigated the extent of sample damage induced by X-ray pulse at various experimental conditions, particularly addressing the following issues: (1) How many electrons are ejected by ionization? (2) How much is the change of the interatomic distances induced by ionization? (3) How many photons are scattered to the detector? These three issues are essential for determining the optimum condition for the SOSS experiment. The first two issues are related to the structural change of the sample molecule induced by X-ray radiation, which should be minimized (or corrected) for the success of SOSS. The third issue determines the signal-to-noise ratio of the scattering pattern, which is critical for the analysis of single-shot, single-molecule scattering patterns. As a quantitative measure to address the third issue, total scattering intensity, which is a sum of the number of photons scattered to all the detector pixels, was calculated while

varying several X-ray pulse parameters. To calculate the total scattering intensity, scattering patterns were calculated following the procedure described in Section D, but without converting an expected number of scattered photons derived from Eq. S2 to an integer. The mean total scattering intensity was obtained by averaging the total scattering intensities of 50 calculated scattering patterns for each condition. We note that the 50 scattering patterns calculated with the same X-ray pulse parameters were slightly different from each other due to the difference in the dynamic form factor. The mean total scattering intensities are shown in Table S1.

In general, the extent of sample damage strongly depends on the duration and the intensity of X-ray pulse. To find suitable experimental conditions for the SOSS experiment, we simulated the sample damage for an iodine molecule with four different pulse durations and four different pulse intensities. Following the procedures of the ionization simulation and the calculation of instant interatomic distance described in Section C and D, we calculated the instant electron density and the instant interatomic distance. Also, we calculated and integrated the scattering pattern to obtain the mean total scattering intensity.

Effect of X-ray pulse duration

First, the sample damage was simulated with four different pulse durations: 1 fs, 3 fs, 5 fs, and 10 fs. The number of photons per pulse was set to be a constant, 10^{11}. The simulation results are shown in Figure S2, Figure S3, and Table S1. With 1-fs pulse, as can be seen in Figure S2, the atoms are not fully ionized and there remain about 30 electrons in the atom even after the interaction with the X-ray pulse is complete. At the intensity maximum of the X-ray pulse, about 75 % of electrons are left bound to the nucleus, ensuring that enough number of photons will be scattered to the detector. Also, as can be seen in Figure S3, the bond length of the iodine molecule stays nearly unchanged during the measurement. At the temporal center of the pulse, the bond length becomes larger by less than 0.01 Å, and by less than 0.05 Å even at the end of the pulse. This bond elongation is comparable to a picometer, which is the spatial resolution required for retrieving vibrational wave functions of I_2. Therefore, the sample damage induced by ionization is negligible with 1-fs pulse.

With 3-fs pulse, there are at least 15 electrons left at the temporal center of the pulse as can be seen in Figure S2. Therefore, we expect that some photons are scattered to the detector but the number of scattered photons would be much less than in the case of 1 fs. As can be seen in Figure S3, the change in bond length is not severe until the temporal center of the

pulse (t = 3 fs), but the bond length increases by larger than an angstrom at the end of the pulse (t = 6 fs). Therefore, the sample molecule is damaged to some extent in this condition.

In the cases of much longer pulses, 5 fs and 10 fs, the atoms are fully ionized (except two electrons in the K-shell) even before reaching the center of the X-ray pulse. Due to the fast ionization, most electrons escape the molecule before enough scattering occurs, resulting in relatively small number of photons scattered to the detector. Also, bond elongation occurs on angstrom length scale, reflecting considerable damage of the sample. Thus, we cannot expect to have the spatial resolution of a few picometers in these conditions.

Effect of X-ray intensity

The extent of sample damage was simulated with four different X-ray intensities (10^{10}, 10^{11}, 10^{12}, and 10^{13} photons per pulse) using the same pulse duration (1 fs). The result is shown in Figure S4, Figure S5, and Table S1. As can be seen in Figures S4 and S5, we cannot see any dependence of instant electron density and bond length on the X-ray intensity. We note that this independence is not a real situation but an artifact which originates from an assumption used in our ionization model. As already mentioned in Section C, we assumed that the L_1 shell is the only electronic level that can be directly ionized by X-ray absorption. But, according to this simulation, it turned out that ionization of the L_1 shell is already saturated even with the intensity of 10^{10} photons per pulse because the ionization probability of the L_1 shell is calculated to be ~100 % in all the steps of simulation. Therefore, in this work, the extent of sample damage does not increase with X-ray intensity. In the real experiment, the ionization rate would increase with the X-ray intensity because the upper electronic shells, L_2, L_3, or other higher shells can be also ionized. These shells have smaller cross sections of ionization than the L_1 shell and should not be saturated with the intensity of 10^{10} photons per pulse.

G. Effect of experimental conditions (X-ray parameters, angular orientation of molecules) on the instrument response

Effect of X-ray pulse intensity

First, we examined the effect of X-ray intensity on PDF reconstruction. To do so, we varied the X-ray intensity from 10^{10} to 10^{13} photons per pulse and checked the quality of reconstructed PDF at each condition. We used a delta function as the reference PDF and

assumed that an I_2 molecule is perfectly aligned in one direction (that is, the first configuration of orientation described in the section above). To calculate the scattering patterns, the X-ray pulse duration of 1 fs was used as determined from the previous section. The reconstructed PDFs for various X-ray intensities are shown in Figure S6. Each of the reconstructed PDFs, $P^R(r)$, has a Gaussian distribution and we found that the width of distribution is strongly dependent on the X-ray intensity. When comparing with the results on the ground-state and the first excited-state vibrational PDFs of I_2 shown in the main text, we can see that ρ and Δ values for the three different reference PDFs (that is, the delta function, the ground-state and the first excited-state vibrational PDFs) are about the same. For the delta function, the ρ value decreases from 0.247 Å to 0.009 Å as the X-ray intensity increases from 10^{10} to 10^{13} photons per pulse as listed in Table S2. As mentioned in the main text, the large ρ value is caused by the lack of the number of photons scattered to the detector. Overall, the increase of X-ray intensity by a factor of ten leads to the decrease of the width of the reconstructed PDF by a factor of three. Thus, to retrieve a vibrational wave function of I_2, it is desirable to use an intense X-ray pulse with the intensity of at least more than 10^{12} photons per pulse (assuming that all the photons are focused into the 3-nm diameter used in our simulations).

Effect of X-ray pulse duration

In order to test the dependence of PDF reconstruction on the temporal duration of X-ray pulse, we repeated the same simulation as shown in Figure S6 using 3-fs X-ray pulse. The reconstructed PDFs using 3-fs X-ray pulse are shown in Figure S7. First of all, we note that ρ and Δ are much larger with the 3-fs pulse than with the 1-fs pulse. The deterioration of spatial resolution with the 3-fs pulse can be attributed to the lack of the number of scattered photons. With the 3-fs pulse, the number of scattered photons is about four times smaller than with the 1-fs pulse, leading to the larger ρ value. Secondly, we can see that the reconstructed PDF is much noisier with the 3-fs pulse than with the 1-fs pulse, and the signal-to-noise ratio does not improve even with the increase of X-ray intensity. The noise introduced with the use of 3-fs pulse is ascribed to the increased ionization. With the increase of pulse duration, ionization occurs more significantly because the time for the interaction between the X-ray pulse and the molecule increases. As a result, the I_2 molecule suffers more severe structural change due to repulsion between atomic nuclei, resulting in the distortion of single-molecule scattering image. As more ionization occurs, the bond length extracted from each single-molecule

scattering image becomes more random, thus making the reconstructed PDF noisier. Thus, SOSS requires an extremely short X-ray pulse in order to obtain a reconstructed PDF that closely resembles the reference PDF.

Effect of angular orientation of molecules

When a scattering pattern is obtained from an ensemble of randomly oriented molecules, an isotropic pattern is obtained because of rotational averaging of the contributions from various orientations of molecules. In contrast, an anisotropic scattering pattern (with respect to the azimuthal angle) is obtained from aligned molecules and contains more structural information than an isotropic scattering pattern.[15] In contrast to the ensemble measurements, a single-shot, single-molecule scattering pattern should be intrinsically anisotropic regardless of the molecular orientation because X-ray is scattered off only one molecule. To examine the effect of molecular orientation on the PDF of bond length reconstructed from the SOSS analysis, we considered three configurations of angular orientation for the I_2 molecule.

In the first configuration, we assumed that the I_2 molecule is perfectly aligned in one direction. In other words, all the molecules were set to be parallel to the y-axis shown in Figure 1a of the main text, and the reconstructed and reference PDFs are dependent only on I–I bond length, r.

In the second configuration, only φ was varied while θ was fixed at 30°. In this case, the reconstructed and reference PDFs are defined with respect to the variables r and φ as follows:

$$P(r, \varphi) = P(r) \times P(\varphi) \tag{S6}$$

where $P(\varphi)$ has a uniform distribution over φ. $P^R(r, \varphi)$ can be integrated over φ to obtain $P^R(r)$, the PDF of bond length.

In the third configuration, we varied only θ following the probability distribution of $\cos^2\theta$ and fixed φ at 0°. In this case, the PDF can be written as follows:

$$P(r, \theta) = a \times P(r) \times \cos^2\theta \tag{S7}$$

where a is a normalization constant.

For each of the above three configurations of molecular orientation, the experimental images were simulated and mapped to the reference library to reconstruct the PDF of I–I bond length. By comparing ρ and Δ values extracted from the PDFs for the three configurations, we investigated how the angular orientation of the molecule affects the retrieval of the bond-length distribution. In this case, we used a delta function as the reference PDF and calculated the scattering pattern using 1-fs X-ray pulse. Four different X-ray intensities varying from 10^{10} to 10^{13} photons per pulse were used.

The results of PDF reconstruction for various configurations of angular orientation and various X-ray intensities are listed in Table S2. The reconstructed PDFs for three different configurations of angular orientation (that is, perfect alignment of the I_2 molecule in one direction; variation of φ with θ fixed at a constant; and variation of θ with φ fixed at a constant) with the X-ray intensity of 10^{13} photons per pulse are shown in Figure S8. When the I_2 molecule rotates freely along φ direction with a fixed θ value, ρ is 0.022 Å and Δ is 0.001 Å. When the I_2 molecule rotates freely along θ direction with φ fixed at a constant, ρ is 0.001 Å and Δ is zero. At a glance, the ρ values for the cases of variable φ (0.022 Å) and variable θ (0.001 Å) look different from that for the case of perfect alignment of the I_2 molecule in one direction (0.009 Å). However, as can be seen in Figure S8, only two distinct bond lengths have nonzero occurrences in the PDFs for all the three orientational configurations, and the ρ value, which is determined by fitting only these two points, is likely to have a large error. Thus, the relatively large difference between the ρ values for the three orientational configurations is simply an artifact, and the actual spatial resolution of PDF reconstruction is indistinguishably good for all the three orientational configurations. For different X-ray intensities (10^{10}, 10^{11}, 10^{12} photons/pulse), the effect of angular orientation is found to be negligible as well. This result indicates that the spatial resolution for retrieving the vibrational wave function is independent of the molecular orientation.

H. Further applications of SOSS

Besides the vibrational wave function investigated in this work, we can also apply the SOSS method to retrieving the internal rotation wave function of a small molecule containing a pair of heavy atoms. For example, the rotational wave functions of 1,2-diiodoethane ($C_2H_4I_2$) are illustrated in Figure S9. A single-shot, single-molecule scattering pattern of the $C_2H_4I_2$ molecule sensitively changes depending on the conformation around the carbon-

carbon bond. Therefore, by constructing a reference library with respect to suitable structural parameters (including the torsional angle of carbon-carbon bond) and mapping the experimental scattering images, we can retrieve the rotational wave function around the carbon-carbon bond.

The molecular system (I_2) proposed in this study and the motion of interest (vibration) is a rather extreme example suitable for SOSS. For molecules consisting of more than two atoms, the method presented in this work (that is, making a reference library and correlation mapping) must be ineffective because much larger amount of computational resource would be required and the mapping consistency will be much poorer for polyatomic molecules than for I_2 molecules. Therefore, a new method needs to be developed to effectively retrieve an instant molecular structure from a scattering pattern of a polyatomic molecule. Nevertheless, the application of the SOSS method can be further extended to biological macromolecules if a specific structural parameter is amplified instead of attempting to capture the overall 3D structure. For example, if a biomolecule is labeled with a pair of metal nanoparticle probes at specific locations via mutagenesis and chemical coupling, the scattering from the molecule will be dominated by the contribution of those heavy-scattering probes.[16-17] In other words, we selectively probe only a specific structural parameter (in this case, the distance between two nanoparticles) and discard all other detailed structural information contained in the scattering pattern. In this simplified case where a nanoparticle is approximated as a large atom, we can apply the same methodology used for I_2 with similar amount of computing resources. By extracting the occurrence distribution of the distance between the two probes using SOSS, we can monitor the conformational fluctuation of the biomolecule.

# I. Comparison with ultrafast vibrational spectroscopy

The SOSS method based on X-ray scattering is closely related to ultrafast vibrational spectroscopy, specifically ultrafast infrared (IR) spectroscopy (IR pump-probe spectroscopy and two-dimensional IR echo spectroscopy), because both techniques can provide information on the structures of rapidly fluctuating molecules in the electronic ground state at equilibrium. However, in principle, the physical quantities measured by the two techniques (and thus their applications) are different from each other. In SOSS, a femtosecond X-ray pulse captures the instantaneous structure of a rapidly fluctuating molecule in the form of a scattering pattern. From the measured X-ray scattering pattern, structural parameters such as

bond lengths and bond angles are retrieved. By sampling many scattering patterns from single molecules, a distribution (that is, PDF) of a structural parameter of interest can be obtained. In other words, SOSS probes the structural fluctuation of molecules by random sampling of individual single molecules using only one X-ray pulse. In ultrafast IR spectroscopy, a series of ultrashort IR pulses tag the initial molecular structures by the frequency of a vibrational mode and then monitors the evolution of that frequency with respect to time. From the evolution dynamics of the vibrational frequency and comparison with theoretical models, the information on the local molecular structure around the probed molecular vibration can be obtained indirectly. In this manner, ultrafast IR spectroscopy probes the structural fluctuation of molecules by keeping track of the evolution of the vibration frequencies of optically tagged molecules using multiple laser pulses. Considering that geometric parameters directly associated with the molecular structure are obtained, SOSS can be regarded as more direct structural method than ultrafast IR spectroscopy. However, at the current stage, due to the limitation of algorithms and computational resources used for extracting the molecular structure from the measured X-ray scattering pattern, the application of SOSS will be limited to only simple molecules isolated from the environment as examined in this work.

In terms of the target system, SOSS is designed for probing the structural variation of a single molecule by making use of ultra-intense X-ray pulses available at XFEL. In contrast, ultrafast IR spectroscopy has been applied only for an ensemble of molecules and it would be difficult to extend the technique to the single-molecule level due to low absorption cross section of the vibrational transition and/or the lack of ultra-intense IR source. The biggest advantage of ultrafast IR spectroscopy is the non-destructive nature of the measurement because low-energy photons at IR frequencies are used. On the other hand, high-energy X-ray photons destroy the molecules even with ultrashort interaction, imposing significant problems on the experimental scheme and the data analysis of the SOSS experiment.

Both techniques can be extended to more sophisticated time-resolved experiment that probes the structural change of molecules undergoing non-equilibrium transitions among multiple electronic states, that is, a photochemical reaction. Such extension can be achieved by combining with an extra visible laser pulse that can excite the molecule to excited electronic states. For ultrafast vibrational spectroscopy, such pump-probe scheme has been already achieved as exemplified by time-resolved IR (visible pump–IR probe) spectroscopy[18] and transient 2D-IR spectroscopy.[19] In the same manner, SOSS has the potential of being

extended to time-resolved SOSS experiment, where a visible laser pulse initiates a reaction and SOSS monitors the progress of the reaction.
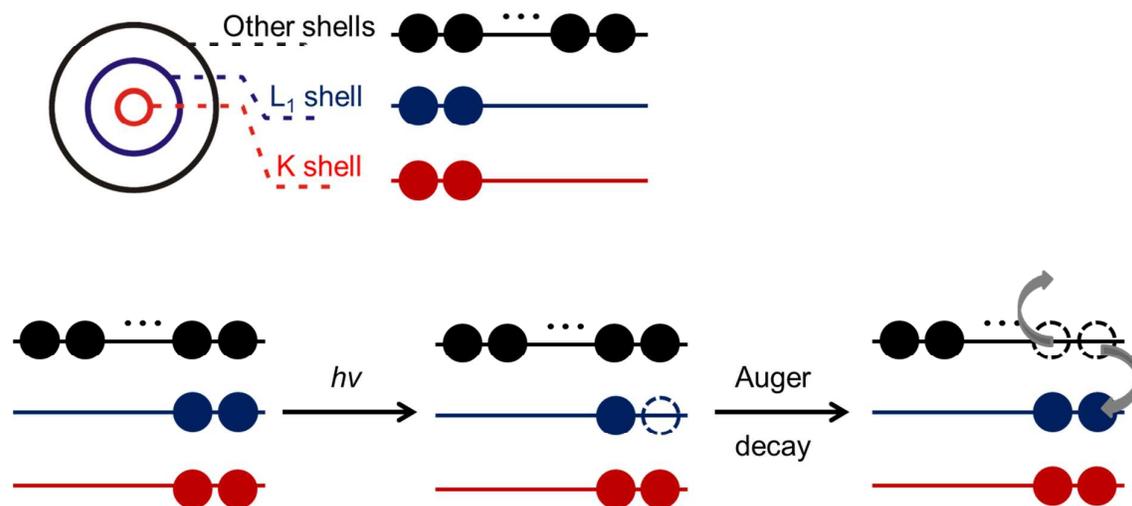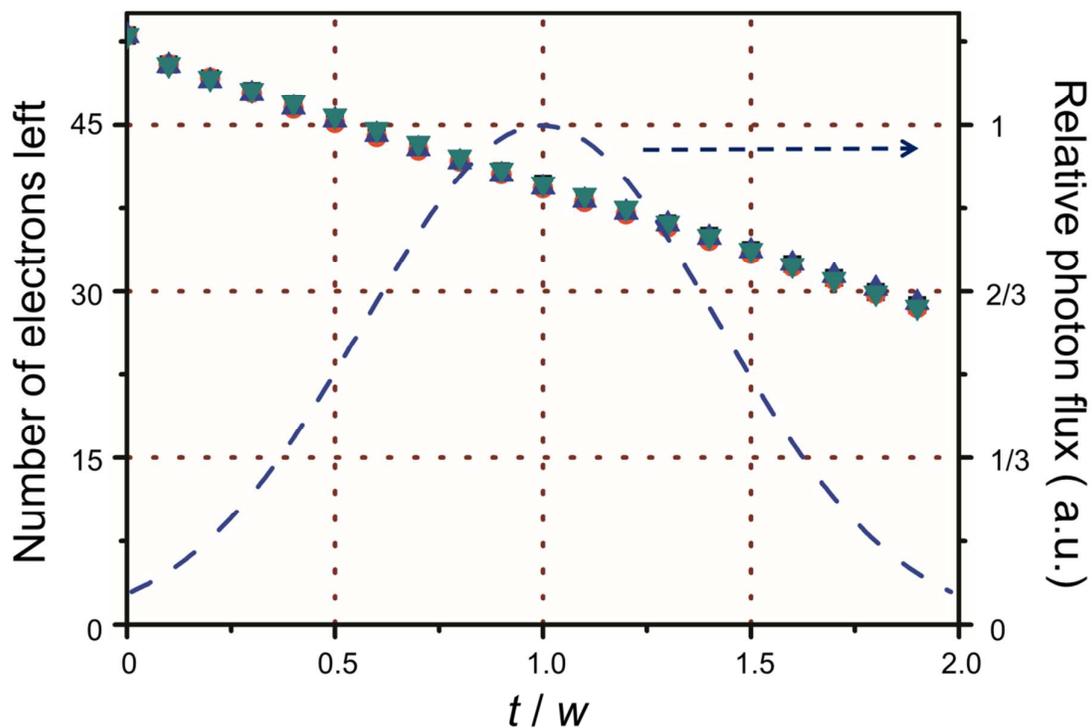
Figure S1. The model of photo-induced ionization used in this work. The two atoms of an I_2 molecule are considered to be independent of each other with any possible interactions between them being ignored. In this simplified model, only three energy levels are considered for the structure of electronic shell of an iodine atom. In this figure, a filled circle means an electron and an empty dashed circle represents a hole. In this ionization model, absorption of X-ray photon occurs only in L_1-shell, resulting in the ionization of two electrons via Auger process.
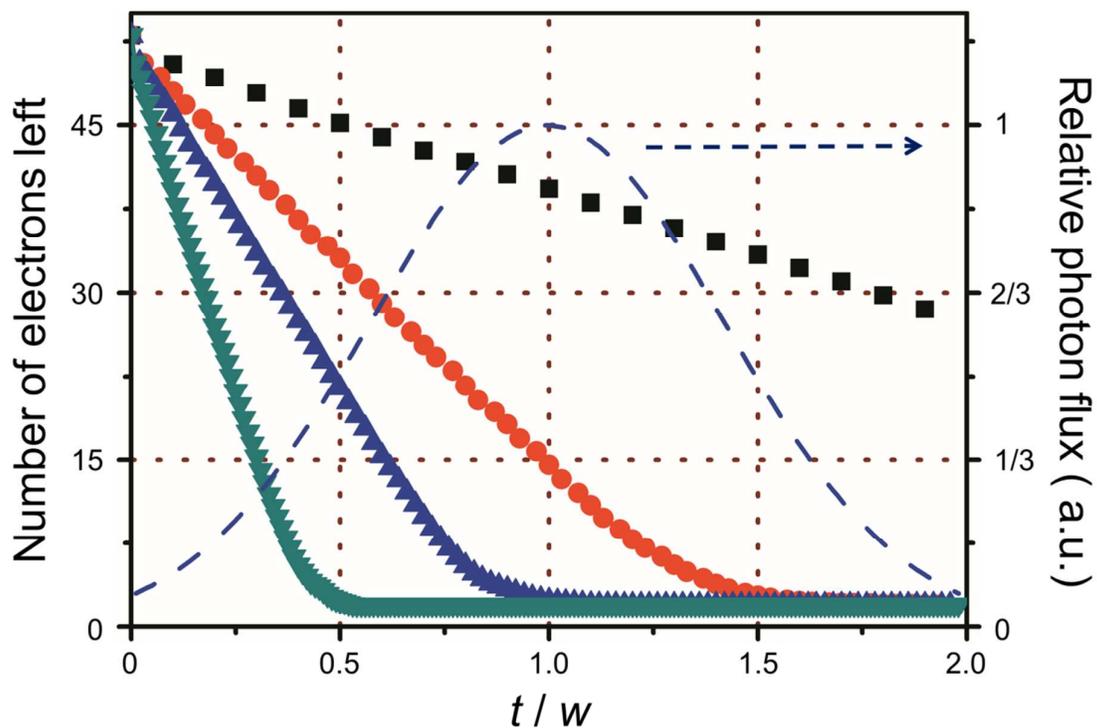
Figure S2. Effect of pulse duration on the instant electron density of an iodine atom. The number of electrons bound to an iodine atom during the interaction of the X-ray pulse and the I_2 molecule was calculated with four different pulse durations: 1 fs (black squares), 3 fs (red circles), 5 fs (blue triangles), and 10 fs (green reverse triangles). A total of 50 simulations were performed and averaged. The step size of the simulation was set to be 0.1 fs. The time axis in the bottom was normalized by the duration of X-ray pulse, w, so that the rate of ionization can be compared relative to the temporal profile of X-ray pulse. The blue dashed line is the normalized temporal profile of the X-ray pulse with its intensity scaling with the label on the right-side axis.
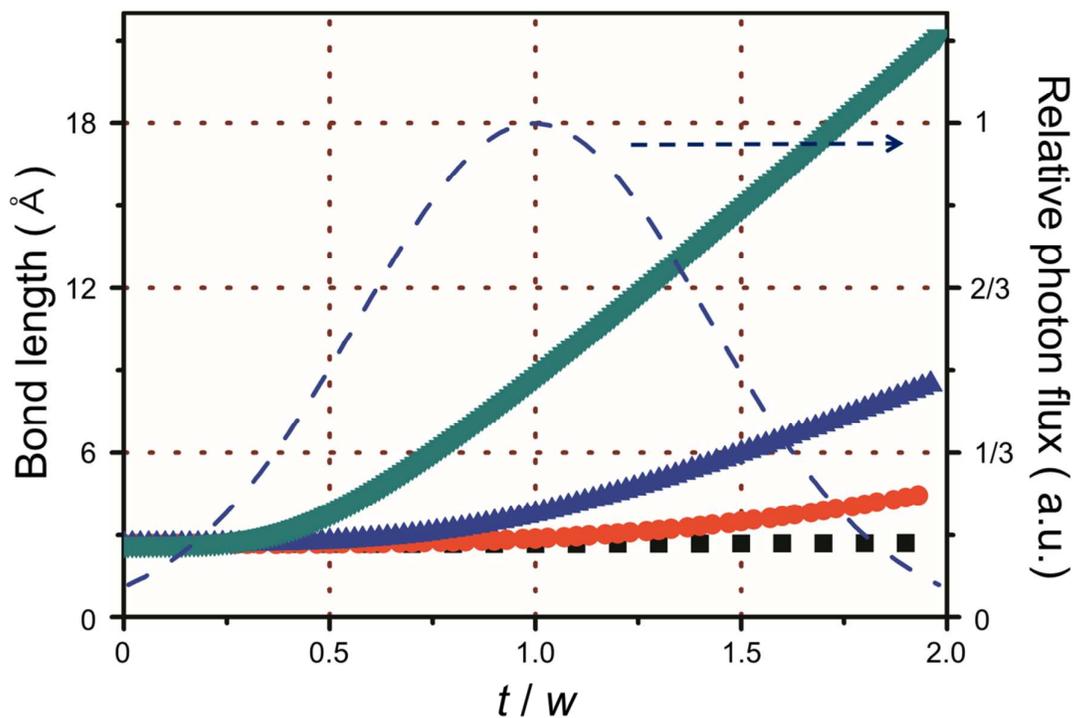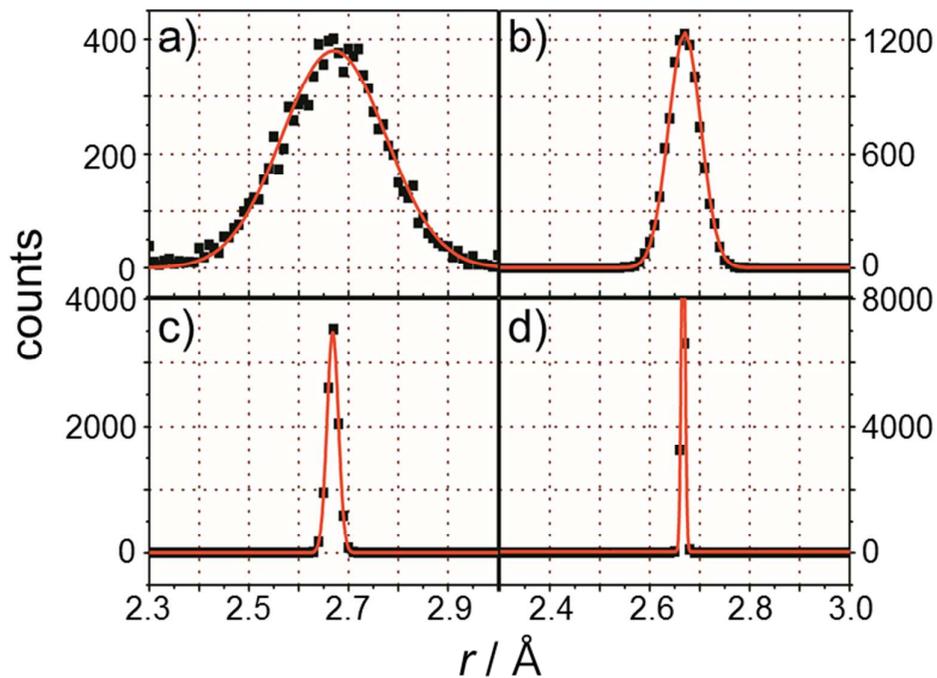
Figure S3. Effect of pulse duration on the instant interatomic distance of an iodine molecule. The instant interatomic distance between two iodine atoms during the interaction of X-ray pulse and the iodine molecule was calculated with four different pulse durations: 1 fs (black squares), 3 fs (red circles), 5 fs (blue triangles), and 10 fs (green reverse triangles). A total of 50 simulations were performed and averaged. The step size of the simulation was set to be 0.1 fs. The time axis in the bottom was normalized by the duration of X-ray pulse, w, so that the rate of ionization can be compared relative to the temporal profile of X-ray pulse. The blue dashed line is the normalized temporal profile of the X-ray pulse with its intensity scaling with the label on the right-side axis.

Figure S4. Effect of X-ray intensity on the instant electron density of an iodine atom. The number of electrons bound to the iodine atom during the interaction of X-ray pulse and the iodine molecule was calculated with four different intensities of X-ray pulse: 10^{10} photons/pulse (black squares), 10^{11} photons/pulse (red circles), 10^{12} photons/pulse (blue triangles), 10^{13} photons/pulse (green reverse triangles). A total of 50 simulations were performed and averaged. The step size of the simulation was set to be 0.1 fs. The time axis in the bottom was normalized by the duration of X-ray pulse, w, as in Figure S2. The blue dashed line is the normalized temporal profile of the X-ray pulse with its intensity scaling with the label on the right-side axis.
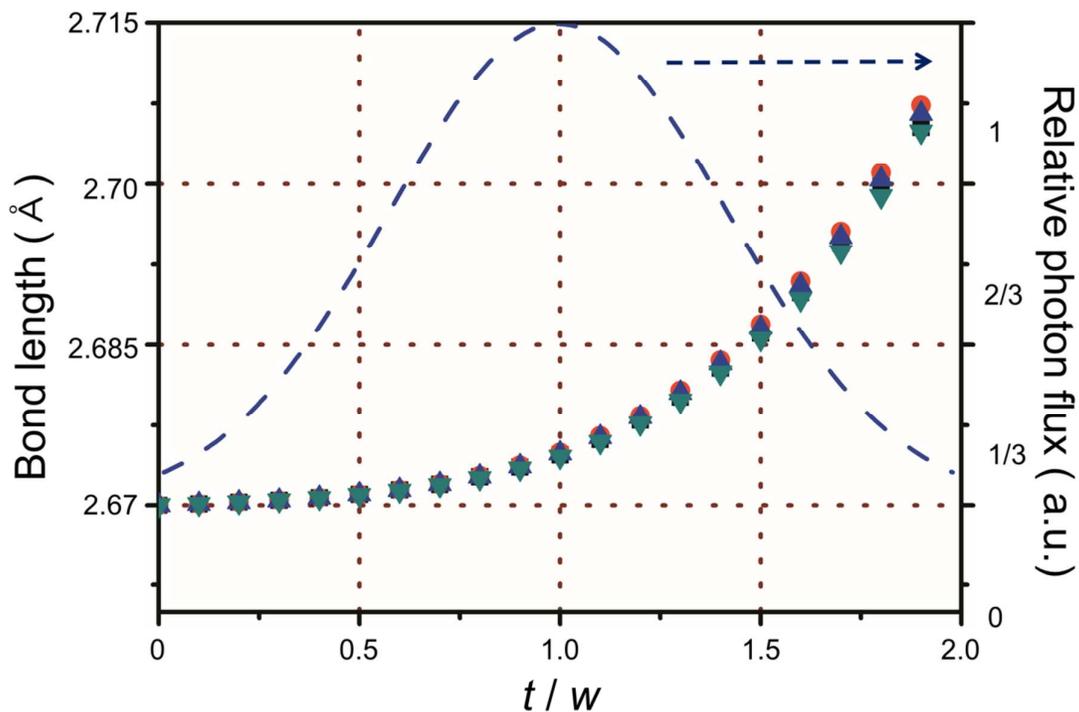
Figure S5. Effect of X-ray intensity on the instant interatomic distance of an iodine molecule. The instant interatomic distance between two iodine atoms during the interaction of the X-ray pulse and the iodine molecule was calculated with four different intensities of X-ray pulse: 10^{10} photons/pulse (black squares), 10^{11} photons/pulse (red circles), 10^{12} photons/pulse (blue triangles), 10^{13} photons/pulse (green reverse triangles). A total of 50 simulations were performed and averaged. The step size of the simulation was set to be 0.1 fs. The time axis in the bottom was normalized by the duration of X-ray pulse, w, as in Figure S2. The blue dashed line is the normalized temporal profile of the X-ray pulse with its intensity scaling with the label on the right-side axis.

Figure S6. Effect of X-ray intensity on the PDF reconstruction. The PDFs were reconstructed by SOSS using various X-ray intensities: a) 10^{10}, b) 10^{11}, c) 10^{12}, and d) 10^{13} photons per pulse. A total of 10,000 experimental scattering images were calculated for an I_2 molecule whose I–I bond length is governed by a reference PDF in the form of a delta function, $P(r) = \delta\,(r \square\,2.67\,\text{Å})$. The temporal duration of X-ray pulse was set to be 1 fs, and it was assumed that the I_2 molecule is perfectly aligned in one direction. Black squares represent the occurrence of scattering images corresponding to a certain bond length. Red solid lines are the fits of the black squares using Eq. (1) of the main text and represent the reconstructed PDF of bond length. Note that y-scales are different between the plots.
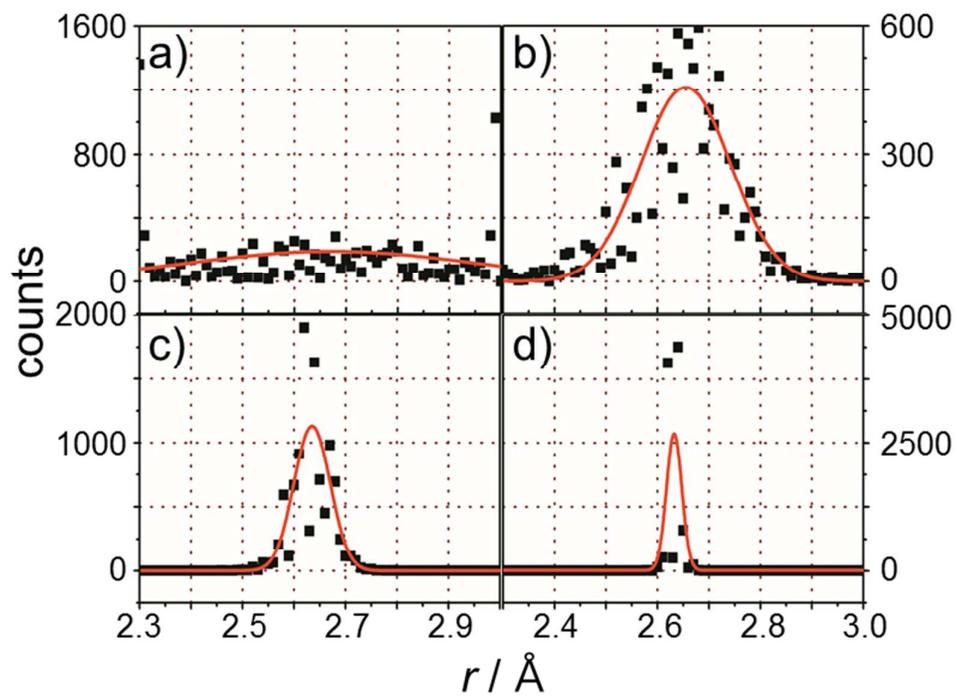
Figure S7. Effect of temporal duration of X-ray pulse on the PDF reconstruction. All other conditions are the same as in Figure S6 except that X-ray pulse duration is stretched to 3 fs. Note that the scales of y-axes are different between the plots.
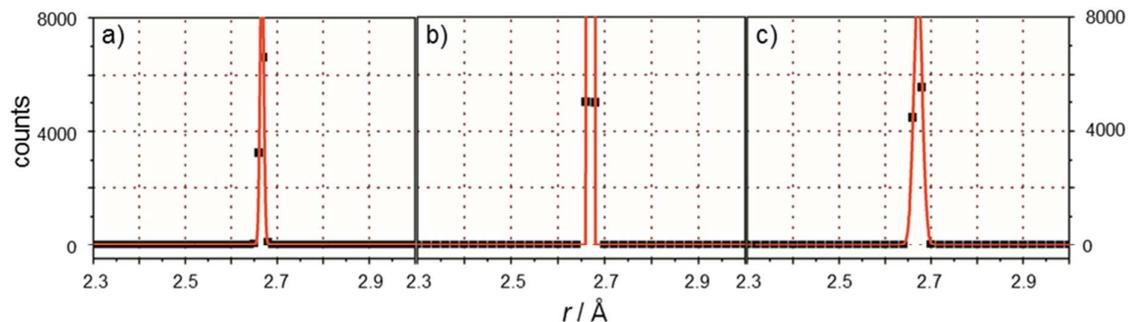
Figure S8. Effect of angular orientation on the PDF reconstruction. The panel a) is equivalent to Figure S6d, which is a reconstructed PDF obtained from perfectly aligned molecules. In panel b), a molecule is perfectly aligned in θ direction, but freely rotates along φ direction. A total of 10,000 scattering images were calculated with a reference PDF, $P(r, \varphi) = \delta (r \square 2.67$ Å$) \times P(\varphi)$, where $P(\varphi)$ is an uniform distribution along φ. In panel c), a molecule is fixed at a constant φ value, but freely rotates along θ direction. A total of 10,000 scattering images were calculated with a reference PDF, $P(r, \theta) = a \times \delta (r \square 2.67$ Å$) \times \cos^2 \theta$. The X-ray intensity of 10^{13} photons per pulse was used for all the configurations of orientation, and the pulse duration was set to be 1 fs. Black squares represent the occurrence of scattering images corresponding to a certain bond length. Red solid lines are the fits of the black squares using Eq. (1) of the main text and represent the reconstructed PDF of bond length.
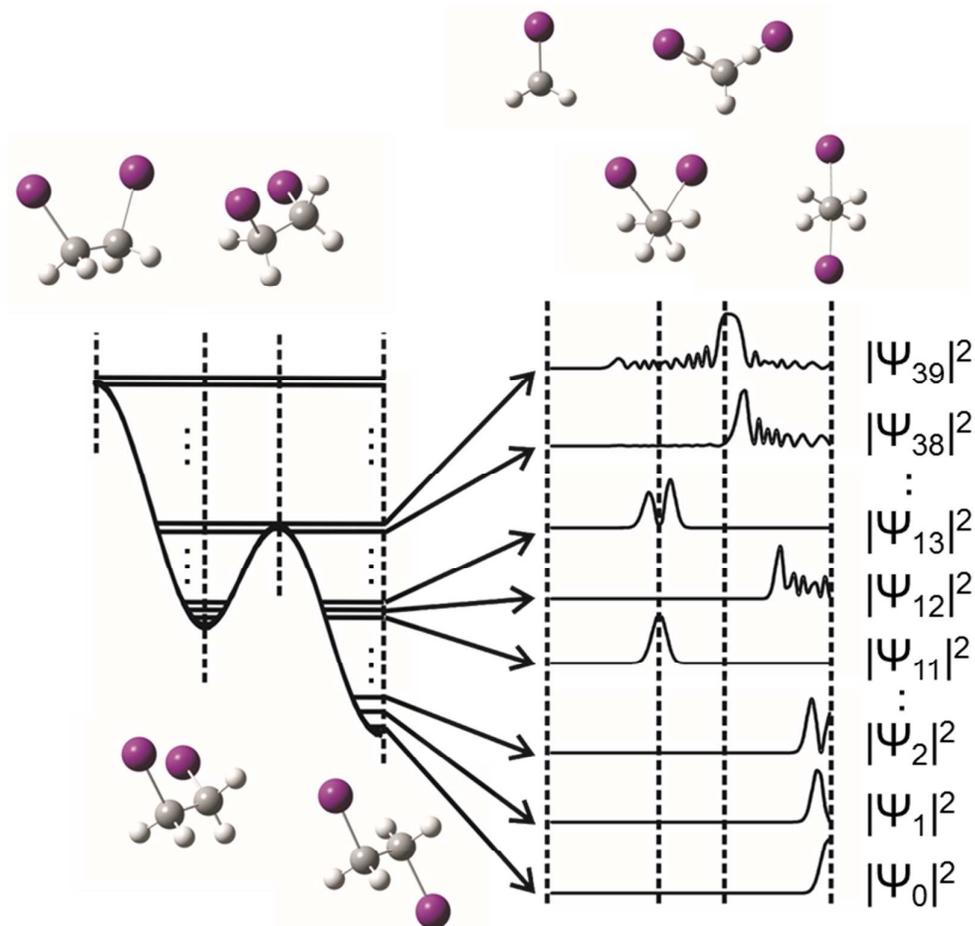
Figure S9. Rotational wave functions describing internal rotational motion around carbon-carbon bond of $C_2H_4I_2$. Only a half side of the potential energy surface (PES) and rotational wave functions are plotted since they have mirror symmetry. The PES was obtained from quantum chemical calculation[20] and the rotational wave function was calculated by using our own program based on a method described in the literature.[21] The molecular structures shown in top right represent Newman projections of four conformational states that correspond to the vertical dashed lines.

Table S1. Mean total scattering intensities measured by the detector with various durations and intensities of X-ray pulse. The total scattering intensity was calculated by summing the number of photons scattered to all the detector pixels. Numbers in parentheses are the total scattering intensity calculated without considering the ionization of the I_2 molecule.

pulse duration / photons per pulse	1 fs	3 fs	5 fs	10 fs
10^{10}	60.27 (105.581)	16.13 (106.801)	6.04 (106.99)	1.60 (107.106)
10^{11}	596.36 (1055.81)	152.72 (1068.01)	58.92 (1069.9)	16.41 (1071.06)
10^{12}	5926.68 (10558.1)	1664.34 (10680.1)	613.34 (10699)	168.02 (10710.6)
10^{13}	60343.2 (105581)	16314.6 (106801)	5934.46 (106990)	1649.55 (107106)

Table S2. Result of PDF reconstruction using SOSS with various experimental conditions. The parameters defining the broadening (ρ) and shift (Δ) of PDF induced by instrument response was calculated by fitting of the reconstructed PDF based on Eqs. (1), (2), and (3) of the main text. For comparison, simulation results shown in Table 1 of main text are shown together.

alignment (free variables)	reference PDF	pulse duration (fs, FWHM)	pulse intensity (photons / pulse)	peak shift, Δ (Å)	broadening, ρ (Å, FWHM)
r	$\delta(r \square 2.67$ Å$)$	1	10^{10}	0.001	0.247
r	$\delta(r \square 2.67$ Å$)$	1	10^{11}	−0.001	0.076
r	$\delta(r \square 2.67$ Å$)$	1	10^{12}	−0.002	0.027
r	$\delta(r \square 2.67$ Å$)$	1	10^{13}	−0.004	0.009
r, θ	$\delta(r \square 2.67$ Å$)$	1	10^{10}	0.001	0.230
r, θ	$\delta(r \square 2.67$ Å$)$	1	10^{11}	−0.001	0.071
r, θ	$\delta(r \square 2.67$ Å$)$	1	10^{12}	0.000	0.029
r, θ	$\delta(r \square 2.67$ Å$)$	1	10^{13}	0.001	0.022
r, φ	$\delta(r \square 2.67$ Å$)$	1	10^{10}	0.004	0.240
r, φ	$\delta(r \square 2.67$ Å$)$	1	10^{11}	−0.001	0.073
r, φ	$\delta(r \square 2.67$ Å$)$	1	10^{12}	0.001	0.028
r, φ	$\delta(r \square 2.67$ Å$)$	1	10^{13}	0.000	0.001
r	$\delta(r \square 2.67$ Å$)$	3	10^{10}	0.002	0.661
r	$\delta(r \square 2.67$ Å$)$	3	10^{11}	−0.016	0.206
r	$\delta(r \square 2.67$ Å$)$	3	10^{12}	−0.037	0.083
r	$\delta(r \square 2.67$ Å$)$	3	10^{13}	−0.038	0.035
r	Ground[a]	1	10^{10}	−0.001	0.260
r	Ground[a]	1	10^{11}	0	0.080
r	Ground[a]	1	10^{12}	−0.002	0.030
r	Ground[a]	1	10^{13}	−0.005	0.013

r	1st excited[b]	1	10^{10}	0.003	0.261
r	1st excited[b]	1	10^{11}	0	0.077
r	1st excited[b]	1	10^{12}	−0.002	0.029
r	1st excited[b]	1	10^{13}	−0.005	0.015

[a] vibrational PDF of a harmonic oscillator in the ground state. [b] vibrational PDF of a harmonic oscillator in the first excited state.

Reference

(1)   Li, J. J.; Wang, X.; Chen, Z. Y.; Zhou, J.; Mao, S. S.; Cao, J. M. Real-Time Probing of Ultrafast Residual Charge Dynamics. *Appl. Phys. Lett.* **2011,** *98*, 011501.

(2)   Stoian, R.; Rosenfeld, A.; Ashkenasi, D.; Hertel, I. V.; Bulgakova, N. M.; Campbell, E. E. B. Surface Charging and Impulsive Ion Ejection During Ultrashort Pulsed Laser Ablation. *Phys. Rev. Lett.* **2002,** *88*, 097603.

(3)   Son, S. K.; Chapman, H. N.; Santra, R. Multiwavelength Anomalous Diffraction at High X-Ray Intensity. *Phys. Rev. Lett.* **2011,** *107*, 218102.

(4)   Veigele, W. J. Photon Cross Sections from 0.1 Kev to 1 Mev for Elements Z = 1 to Z = 94. *At. Data* **1973,** *5*, 51-111.

(5)   Krause, M. O.; Oliver, J. H. Natural Widths of Atomic K-Levels and L-Levels,K-Alpha X-Ray-Lines and Several Kll Auger Lines. *J. Phys. Chem. Ref. Data* **1979,** *8*, 329-338.

(6)   Hau-Riege, S. P.; London, R. A.; Szoke, A. Dynamics of Biological Molecules Irradiated by Short X-Ray Pulses. *Phys. Rev. E* **2004,** *69*, 051906.

(7)   Ben-Nun, M.; Cao, J. S.; Wilson, K. R. Ultrafast X-Ray and Electron Diffraction: Theoretical Considerations. *J. Phys. Chem. A* **1997,** *101*, 8743-8761.

(8)   Azároff, L. V. *X-Ray Diffraction*; McGraw-Hill: New York; 1974.

(9)   Neutze, R.; Wouts, R.; van der Spoel, D.; Weckert, E.; Hajdu, J. Potential for Biomolecular Imaging with Femtosecond X-Ray Pulses. *Nature* **2000,** *406*, 752-757.

(10)   Habershon, S.; Zewail, A. H. Determining Molecular Structures and Conformations Directly from Electron Diffraction Using a Genetic Algorithm. *ChemPhysChem* **2006,** *7*, 353-362.

(11)   Marchesini, S.; He, H.; Chapman, H. N.; Hau-Riege, S. P.; Noy, A.; Howells, M. R.; Weierstall, U.; Spence, J. C. H. X-Ray Image Reconstruction from a Diffraction Pattern Alone. *Phys Rev B* **2003,** *68*, 140101.

(12)   Kim, S.; Kim, C.; Lee, S.; Marathe, S.; Noh, D. Y.; Kang, H. C.; Kim, S. S.; Sandy, A.; Narayanan, S. Coherent Hard X-Ray Diffractive Imaging of Nonisolated Objects Confined by an Aperture. *Phys Rev B* **2010,** *81*, 165437.

(13)   Nam, D.; Park, J.; Gallagher-Jones, M.; Kim, S.; Kohmura, Y.; Naitow, H.; Kunishima, N.; Yoshida, T.; Ishikawa, T.; Song, C. Imaging Fully Hydrated Whole Cells by Coherent X-Ray Diffraction Microscopy. *Phys. Rev. Lett.* **2013,** *110*, 098103.

(14)   Quiney, H. M.; Nugent, K. A. Biomolecular Imaging and Electronic Damage Using X-Ray Free-Electron Lasers. *Nat. Phys.* **2011,** *7*, 142-146.

(15)   Kim, J.; Kim, K. H.; Kim, J. G.; Kim, T. W.; Kim, Y.; Ihee, H. Anisotropic Picosecond X-Ray Solution Scattering from Photoselectively Aligned Protein Molecules. *J. Phys. Chem. Lett.* **2011,** *2*, 350-356.

(16)   Mathew-Fenn, R. S.; Das, R.; Harbury, P. A. B. Remeasuring the Double Helix. *Science* **2008,** *322*, 446-449.

(17)   Mathew-Fenn, R. S.; Das, R.; Silverman, J. A.; Walker, P. A.; Harbury, P. A. B. A Molecular Ruler for Measuring Quantitative Distance Distributions. *PLoS One* **2008,** *3*, e3226.

(18)   Lim, M. H.; Jackson, T. A.; Anfinrud, P. A. Ultrafast Rotation and Trapping of Carbon Monoxide Dissociated from Myoglobin. *Nat. Struct. Biol.* **1997,** *4*, 209-214.

(19)   Kolano, C.; Helbing, J.; Kozinski, M.; Sander, W.; Hamm, P. Watching Hydrogen-Bond Dynamics in a Beta-Turn by Transient Two-Dimensional Infrared Spectroscopy. *Nature* **2006,** *444*, 469-472.

(20)   Kim, J. Theoretical Studies of Photochemical Reactions and Spin-Orbit Effect on Ground and Excited States. Ph.D. Thesis, Korea Advanced Institute of Science and Technology, Korea, 2010.

(21)   Lewis, J. D.; Malloy, T. B.; Chao, T. H.; Laane, J. Periodic Potential Functions for Pseudorotation and Internal-Rotation. *J. Mol. Struct.* **1972,** *12*, 427-449.