PERSPECTIVES FOR IMAGING SINGLE PROTEIN MOLECULES WITH THE PRESENT DESIGN OF THE EUROPEAN XFEL

Svitozar Serkez, Vitali Kocharyan, Evgeni Saldin, Igor Zagorodnov
Deutsches Elektronen-Synchrotron (DESY), Hamburg, Germany
Gianluca Geloni
European XFEL GmbH, Hamburg, Germany
Oleksandr Yefanov
Center for Free-Electron Laser Science, Hamburg, Germany

Abstract

European XFEL aims to support imaging and structure determination of biological specimens between less than 0.1 microns and 1 micron size with working photon energies between 3 keV and 16 keV. This wide operation range is a cause for challenges to the focusing optics. A long propagation distance of about 900 m between x-ray source and sample leads to a large lateral photon beam size at the optics. Due to the large divergence of nominal X-ray pulses with durations shorter than 10 fs, one suffers diffraction from mirror apertures, leading to a 100-fold decrease in fluence at photon energies around 4 keV, which seem ideal for imaging of single biomolecules. Moreover, the nominal SASE1 is very far from the level required for single particle imaging. Here we show how it may be possible to optimize the SPB instrument for single biomolecule imaging with minimal additional costs and time, achieving diffraction without destruction at near-atomic resolution with $10^{13}$ photons in a 4 fs pulse at 4 keV photon energy and in a 100 nm focus, corresponding to a fluence of $10^{23}$ ph/cm$^2$. This result is exemplified using the RNA Pol II molecule as a case study.

INTRODUCTION AND REQUIREMENTS

Imaging of single molecules at near-atomic resolution is expected to result in a significant advance in structural biology. One could obtain structural information of large macro-molecular assemblies that cannot crystallize, like membrane proteins. In this contribution we study possibilities and opportunities in this field of science, which will be enabled by applying advanced FEL techniques to the SPB (Single Particle and Biomolecule) instrument to be installed in the European XFEL baseline\(^1\) [1,2]. In order to perform single molecule imaging, a straightforward “diffraction before destruction” method has been proposed [3]–[6]. A great number of single molecules with the same structure are injected into vacuum and interact with ultrashort X-ray pulses, before being completely destroyed. A sufficient number of diffraction patterns is recorded, with unknown orientation. Next, the relative orientations of the different images is determined, so that a 3D diffraction pattern can be assembled in the reciprocal space [7]–[11]. The 3D electron density of the molecule is obtained from the 3D diffraction pattern with the help of a phase retrieval method. An important parameter of the problem is the number of scattered photons per effective Shannon pixel. For biological material, the photon count per shot per pixel of solid angle $\Omega_p$, averaged over shells of wavenumber $q$ is proportional to the wavelength $\lambda^2$ [13]. Lower photon energies result in a stronger diffraction signal, but a limit is dictated by the resolution that one needs to achieve, a balance being in the range between 3 keV and 5 keV. The FWHM focal spot size should be roughly between 5 and 10 times larger than the sample size to grant good photon beam quality within the interaction area. For example, a spot size of 100 nm is good for sample sizes of 10 - 20 nm [14]. We find therefore that a biomolecule of around 15 nm diameter, with $N_{\text{atom}} \sim 30000$, requires a pulse fluence of about $10^{13}$ph/$(100$ nm$)^2$, for an average of $\langle N_p \rangle \sim 1.5$ photons per Shannon pixel at a photon energy of 4 keV. This signal level is higher than what is required by usual methods of pattern orientation determination. Photons have to be delivered in extremely short X-ray pulses to limit radiation-induced changes during the exposure. Estimations indicate that an X-ray pulse duration shorter than about 4 fs is needed [15]–[19]. The key metric for optimizing a photon source for single biomolecule imaging is then the peak power. Ideally, the peak power in our case of interest should be more than 1 TW. For example, we note that $10^{13}$ photons at 4 keV correspond to an energy of about 6 mJ which yields, in 4 fs, a peak power of about 1.5 TW. It is worthwhile to mention that 1 TW at 4 keV gives the same signal per Shannon pixel as 27 TW at 12 keV (assuming a fixed pulse duration).

TW SOURCE FOR THE SPB LINE

The SPB instrument at the European XFEL will be located at the SASE1 undulator line [1,2]. Figure 1 shows this line from the injector up to the SASE1 undulator. Our scheme for an X-ray source suitable for the SPB instrument is heavily based on the use of a slotted spoiler foil in the last bunch compressor chicane, a method devised and ex-

\(^1\) A much more fleshed-out report can be found in [12], where the reader is also addressed to for a more complete list of references.
perimentially proved at the LCLS [20]-[22]. The last linac section before the third bunch compressor BC2 is set at an off-crest accelerating rf phase, so that a $\gamma - t$ bunch tilt is present at the center of BC2. A thin foil with a narrow slot at its center is placed in the beam path. Coulomb scattering of the electrons passing through the foil increases the emittance of most of the beam, but leaves a thin unspoiled slice, where the beam passes through the slit, thus allowing for an x-ray FEL pulse much shorter than the FWHM electron bunch duration. The minimum duration of the unspoiled slice of the electron bunch measured at the LCLS is about 3 fs. A design of a self-seeding setup based on the undulators system for the European XFEL is sketched in Fig. 2. We exploit a combination of a self-seeding scheme [23]-[42] with an undulator tapering technique [43]-[51] consisting in a slow reduction of the field strength of the undulator in order to preserve the resonance wavelength, while the kinetic energy of the electrons decreases due to the FEL process. Highly monochromatic pulses generated with the self-seeding technique make the tapering more efficient than in the SASE case. Here we study a scheme for generating 1 TW-level X-ray pulses in the SASE1 tapered undulator. We optimize our setup based on start-to-end simulations for 14 GeV electron beam with 1 nC charge compressed up to 10 kA peak current. In this way, the output power of the SASE1 undulator could be increased from the value of 100 GW in the SASE regime to about 1.5 TW at the photon energy range around 4 keV. For self-seeding we consider a single-crystal scheme, with a crystal identical to that installed at the LCLS, allowing for exploitation of different reflections. In Fig. 3 we show scattering geometry, amplitude and phase of the transmittance for the C(111) asymmetric Laue reflection at 4.1 keV. Shown that in our case of interest the HOMs are expected to preserve the radiation wavefront. Once the radiation pulse enters the experiment area, it is focused by a KB mirror system to about 100 nm size [2]. A baseline layout for KB system is shown in Fig. 4. Two elliptical mirrors with a 950 mm clear aperture along the mirror surface and a fixed incidence angle of 3.5 mrad are assumed in the vertical and horizontal direction in order to achieve high efficiency at high photon energies. Considering a 950 mm clear aperture and a 3.5 mrad reflection angle, one obtains a lateral aperture of 3.3 mm. However, the about 900 m-long propagation distance from source to sample leads to a large lateral beam size at the focusing optics. In fact, accepting 4$\sigma$ of the beam, for a photon energy of 4 keV the desired lateral aperture for the ultra-short pulse case is about 8 mm. As a result, due to the large divergence of a nominal X-ray pulse shorter than 10 fs, one suffers major diffraction effects from the KB mirror aperture, leading to about a hundred-fold decrease in fluence at photon energies around 4 keV. However, it is possible to obtain an X-ray source capable of producing X-ray pulses with smaller angular divergence of about 2 µrad and, simultaneously, a few fs duration, by taking advantage of a minimal modification in the accelerator complex, amounting to the introduction of a slotted foil in the last electron bunch compressor.

**OPTICS LAYOUT FOR THE SPB LINE**

The SPB optical layout [1,2] is sketched in Fig. 4. The first upstream optical element is a Horizontal Offset Mirror (HOM) pair with a clear aperture along the mirror surface of 800 mm [52]. For the maximal incident angle $\theta = 3.6$ mrad, one achieves an overall high-reflectivity close to 100% over the photon energy range between 3 keV and 5 keV. It can be shown that in our case of interest the HOMs are expected to preserve the radiation wavefront. Once the radiation pulse enters the experiment area, it is focused by a KB mirror system to about 100 nm size [2]. A baseline layout for KB system is shown in Fig. 4. Two elliptical mirrors with a 950 mm clear aperture along the mirror surface and a fixed incidence angle of 3.5 mrad are assumed in the vertical and horizontal direction in order to achieve high efficiency at high photon energies. Considering a 950 mm clear aperture and a 3.5 mrad reflection angle, one obtains a lateral aperture of 3.3 mm. However, the about 900 m-long propagation distance from source to sample leads to a large lateral beam size at the focusing optics. In fact, accepting 4$\sigma$ of the beam, for a photon energy of 4 keV the desired lateral aperture for the ultra-short pulse case is about 8 mm. As a result, due to the large divergence of a nominal X-ray pulse shorter than 10 fs, one suffers major diffraction effects from the KB mirror aperture, leading to about a hundred-fold decrease in fluence at photon energies around 4 keV. However, it is possible to obtain an X-ray source capable of producing X-ray pulses with smaller angular divergence of about 2 µrad and, simultaneously, a few fs duration, by taking advantage of a minimal modification in the accelerator complex, amounting to the introduction of a slotted foil in the last electron bunch compressor.

**RADIATION FROM SASE1**

We consider current profile, normalized emittance, energy spread profile, electron beam energy spread and wakefields.
from start-to-end simulations for the European XFEL at 14 GeV [53]. The electron beam charge is 1 nC, and the peak current is 10 kA. Detailed computer simulations with $2 \cdot 10^5$ macroparticles have been carried out to evaluate the performance of the slotted spoiler using the tracking code ELEGANT [54]. They include multiple Coulomb scattering in a 2 µm thin aluminum foil for two case studies referring to pulse length of 12 fs and 4 fs. A slit full-width of 0.7 mm selects a small fraction of electrons, about 20%, and produces an unspoiled electron bunch slice after BC2. Self-seeding and post-saturation tapering simulations were performed with the help of the FEL code Genesis [55]. The output power and spectrum of the entire setup, that is after the second part of the output undulator for the 4 fs case, are shown in Fig. 5.

**NANO-SCALE FOCAL SPOT**

We carried out wavefront propagation simulations to investigate the evolution of the radiation beam profile through the SPB optics. Our wave optics analysis takes into account aberrations and errors from each optical element. In our case of interest, a reflection from the mirror becomes similar to the propagation through a transparency at the mirror position, which just changes the phase of the reflected beam without changing its amplitude. Applying the Marechal criterion, i.e. requiring a Strehl ratio larger than 0.8, and treating the errors from the different optics independently, we conclude that an height error $h_{\text{rms}} < 1.5$ nm should be sufficiently small for diffraction-limited propagation through the SBP beamline at a photon energy of 4 keV. In fact, the SPB instrument designers are planning to use mirrors capable to preserve the geometrical focus properties at much shorter wavelength range. The effects of the horizontal offset mirrors in the X-ray beam transport are modeled using the code SRW [56] as a combination of two apertures with sizes determined by the mirror length (800 mm in our case) and two phase shifters describing the mirror surface errors. The SRW code has further the capability of modeling the KB optics by elliptical mirrors (with length 950 mm in our case) and to account for all aberrations. The KB mirror surface errors are simulated by two phase shifters, similar to the case of offset mirrors. The plot in Fig. 6 shows the intensity profile at the focus, integrated over the radiation pulse. This is thus a simulation of the energy profile per unit surface that can be measured by a detector that integrates over a single radiation pulse, placed in the plane of interest. The maximal fluence in the focus may now be obtained from $F = N_p/S$, where $N_p$ is the number of photons into the radiation pulse and $S$ is the effective focal spot squared. We found that $1/S \sim 1.5 \cdot 10^9$ cm$^{-2}$. For $10^{13}$ photons per pulse, which can be achieved as discussed previously, this amounts to a fluence of about $1.5 \cdot 10^{22}$ photons/cm$^2$. This result can be achieved without additional cost for the baseline optical layout of the SPB beamline, and with very moderate costs for the installation of the slotted foil setup into the beam formation system. The X-ray optical layout of the SPB instrument provides an option of operation with an intermediate source point in the horizontal plane allowing for a fluence of about $0.5 \cdot 10^{23}$ photons/cm$^2$. In order to achieve additional tightening of the focusing, pending feasibility study, one might also additionally install two vertical refocusing mirrors. In this way, with a moderate additional cost, a higher fluence of about $10^{23}$ photons/cm$^2$ is within reach.

**NOISY X-RAY DIFFRACTION PATTERNS**

The calculations in this article were carried out for a photon energy of 4.1 keV, corresponding to a wavelength $\lambda = 0.3$ nm. For reference, note that in order to reach the resolution $d = 2\pi/(0.4$ nm) with a 200 mm by 200 mm detector, the sample to detector distance needs to be as short as 10 cm. For this resolution, and for a molecule size of $w = 10$ nm, we estimate the requirement for the number of Shannon pixels as $N_x = 4w/d = 100$. The Adaptive Integrating Detector (AGIPD) [57], which will be installed at the SPB instrument features a pixel size of 0.2 mm and frame of 1 megapixel, amounting to a total lateral size $D = 200$ mm. In this case, the average size of a Shannon pixel can be estimated as $D/N_x = 2$ mm. However, our calculations assume that the detector can be placed at the necessary propagation distance (10 cm) to realize the desirable resolution. The numerical simulations carried out here are based on
some simplifying assumptions. In particular, noise is only considered in terms of photon noise, i.e. we assume Poisson shot noise. No additional sources of noise such as detector readout noise are considered. A quantum efficiency of about 85% is assumed for the AGIPD detector, for standard window and at the photon energy of 4 keV. Concerning issues related with detector electronics it can be shown that the expected number of false hits per detector pixel (0.2 mm by 0.2 mm) can be neglected. We simulated 30000 randomly oriented diffraction patterns for the RNA Pol II structure. The plot in Fig. 7 shows the radial average of the photon count. The simulated array size was 200 by 200 pixels, with sampling ratio per dimension of \( s = 2 \), and binning \( b = 5 \). Note that larger molecules do not necessary give larger signals; there are a fixed number of photons per pulse and larger molecules require a proportionally larger focal spot size, hence giving a lower fluence. The plot in Fig. 7 demonstrates that a signal of the order of 0.1 photons per pixel, corresponding to 0.4 photons per Shannon angle can be achieved. A typical diffraction pattern from a single FEL pulse as seen by AGIPD detector is shown in Fig. 8. A difficulty to be considered is that the 6 mm gap width between two detector modules corresponds, in our case of single biomolecule imaging, to the size of three Shannon pixels.

**DISCUSSIONS AND CONCLUSIONS**

The imaging method “diffraction before destruction” promises to be a revolutionary technique for protein determination, capable of resolving the structure of molecules that cannot crystallize. Here we propose a cost-effective proof-of-principle experiment, aiming to demonstrate the actual feasibility of a single molecule diffraction experiment using the baseline European XFEL accelerator complex and the SPB beamline hardware, with minimal modifications only. More specifically, we want to determine the structure of a relatively small (about 30000 non-hydrogen atoms), well-known protein molecule and compare it with results in the protein data bank. We developed a complete package of computational tools for start-to-end simulations predicting the performance of this experiment. Its composition is sketched in Fig. 9. In this paper we reported about detailed simulations from the photocathode of the European XFEL injector up to the collection of noisy 2D diffraction data set. Several issues are discussed. As a follow-up of this work, we will perform image reconstruction up to the determination of the electron density distribution, after image orientation and assembly of the 3D diffraction data set.

**REFERENCES**


SCRYM, Inhouse code written by V. Kocharyan and G. Geloni, partly based on xframework routines by I. Agapov, available on request only.

MOLTRANS, Inhouse code written by E. Weckert, available on request only.

CORREL, Inhouse code written by E. Weckert, available on request only.

ORIENT, Inhouse code written by O. Yefanov, available on request only.

RECON, Inhouse code written by O. Yefanov, available on request only.