

Direct Phasing of Finite Crystals Illuminated with a Free-Electron Laser

Richard A. Kirian,^{1,2,*} Richard J. Bean,^{1,3} Kenneth R. Beyerlein,¹ Miriam Barthelmess,¹ Chun Hong Yoon,^{1,3} Fenglin Wang,¹ Flavio Capotondi,⁴ Emanuele Pedersoli,⁴ Anton Barty,¹ and Henry N. Chapman^{1,5}

¹Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany

²Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

³European XFEL GmbH, Albert Einstein Ring 19, 22761 Hamburg, Germany

⁴Fermi, Elettra Sincrotrone Trieste, SS 14km 163.5, 34149 Basovizza, Trieste, Italy

⁵University of Hamburg, Luruper Chaussee 149, Hamburg, 22761, Germany

(Received 31 July 2014; revised manuscript received 5 November 2014; published 12 February 2015)

It has been suggested that the extended intensity profiles surrounding Bragg reflections that arise when a series of finite crystals of varying size and shape are illuminated by the intense, coherent illumination of an x-ray free-electron laser may enable the crystal's unit-cell electron density to be obtained *ab initio* via well-established iterative phasing algorithms. Such a technique could have a significant impact on the field of biological structure determination since it avoids the need for *a priori* information from similar known structures, multiple measurements near resonant atomic absorption energies, isomorphous derivative crystals, or atomic-resolution data. Here, we demonstrate this phasing technique on diffraction patterns recorded from artificial two-dimensional microcrystals using the seeded soft x-ray free-electron laser FERMI. We show that the technique is effective when the illuminating wavefront has nonuniform phase and amplitude, and when the diffraction intensities cannot be measured uniformly throughout reciprocal space because of a limited signal-to-noise ratio.

DOI: 10.1103/PhysRevX.5.011015

Subject Areas: Optics

I. INTRODUCTION

Serial femtosecond crystallography (SFX) [1–3] exploits coherent, femtosecond-duration hard x-ray pulses produced by an x-ray free-electron laser (XFEL) to record diffraction patterns from protein crystals on a time scale that is faster than atomic motion [4]. This method of “diffraction-before-destruction” allows for a delivered radiation dose (energy/mass) that is orders of magnitude larger than the tolerable dose using longer duration exposures because the pulse terminates before the onset of significant atomic motion [5], thereby bypassing some of the known limitations of conventional synchrotron-based macromolecular crystallography. Such high intensity snapshot diffraction patterns allow for room-temperature structure determination of radiation-sensitive biological macromolecules from crystals of just a few hundred nanometers in size. In this way, SFX has enabled the room-temperature determination of novel structures from natively inhibited *in-vivo* grown protein microcrystals [6] and G protein-coupled receptor micro-crystals grown in lipidic cubic phase [7], and it permits atomic-resolution structure determination in static [8] and time-resolved studies.

SFX requires the averaging of many snapshot diffraction patterns from different crystals in order to measure reliable diffraction intensity to high resolution. These averaged data suffer from the same phase problem as encountered in conventional crystallography. Existing methods for determining diffraction phases require either a known structure of sufficient similarity to produce initial phase estimates [9] or high-resolution data in conjunction with either (1) small molecular structures (about 1000 atoms or less) [10], (2) isomorphous derivative crystals labeled with heavy atoms [11], (3) diffraction data recorded near resonant conditions [12], or (4) other physical modifications of the structure (e.g., by inducing radiation damage [13]). Most of the aforementioned phasing methods are likely applicable to SFX data [14], though some modifications may be necessary in cases of extremely high x-ray fluence [15]. The method we demonstrate here differs from all of them and may be applied to SFX data from submicron crystals without restrictions on resolution.

The genesis of SFX is, in part, due to the advances in x-ray coherent diffractive imaging (CDI) [16,17], which has previously shown that lenseless imaging could be performed on single-shot XFEL images taken before the object was destroyed [18]. CDI works on the basis that a fine sampling of the continuous coherent diffraction pattern contains enough information to determine the phases of the recorded far-field diffraction amplitudes [19]. This information may be used to produce an image of the illuminated object without the use of a lens, at a resolution limited, in

*richard.kirian@asu.edu

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

principle, only by the wavelength of the radiation and the maximum recorded diffraction angle.

As Sayre pointed out [20], traditional crystallographic diffraction data only measure the diffraction intensity at reciprocal lattice points that satisfy the Bragg condition, which is a factor of 2 less than the critical sampling rate needed to recover an unaliased autocorrelation function of the crystal unit cell. For that reason, the common iterative phasing strategies associated with CDI have only rarely been applied directly to crystallographic data in situations of especially high solvent fraction [21,22]. Work by Perutz carried out decades ago [23] aimed to determine oversampled molecular transforms by physically modifying the unit-cell size of haemoglobin crystals, but this approach has not developed into a practical phasing strategy so far. However, early SFX experiments demonstrated strong diffraction from submicron protein crystals that exhibited measurable intensities sampled between Bragg spots without any need to modify the crystals [1]. More recent work has shown that the average over tens of thousands of diffraction patterns, each from a different crystal, into a finely sampled three-dimensional average intensity map contains a recoverable signal between Bragg reflections at medium resolutions [24]. This is remarkable evidence that SFX data allow one to bridge the divide between the crystallographic phase problem and the continuous Fraunhofer diffraction phase problem. A variety of iterative phasing strategies that have been developed in the context of CDI [25–27] can be applied to this three-dimensional average intensity, as suggested previously and demonstrated in simulations [28]. Additional simulation studies have been performed that extend the original concept to include the effects of disorder [29–31] and have explored the sensitivity of the method to noise [32,33]. Furthermore, the effects of different crystal-edge truncations have been investigated to find that they can greatly influence the diffracted intensity between Bragg spots. This turns out to be an important consideration when attempting iterative phasing, and methods to account for it have been suggested [31,34].

The possibility of directly imaging finite crystals via oversampled diffraction was considered prior to the work of Spence *et al.* [35]. As shown by the simulations of Miao *et al.*, iterative phase-retrieval methods may be applied directly to individual (appropriately sampled) finite-crystal diffraction patterns in the same way as they are applied to noncrystalline targets. In practice, the recovery of a three-dimensional structure by such means would require (1) a signal-to-noise ratio in single diffraction patterns that is sufficient to support image recovery, (2) a reciprocal-space sampling frequency corresponding to the size of the whole crystal, (3) a resolution limited by the projection approximation (for given wavelength), and (4) tomographic assembly of many real-space images after iterative phase retrieval is applied to an equal number of diffraction

patterns, each of which introduces errors into the final real-space structure. The method we (and colleagues) proposed previously [28] avoids all of the aforementioned challenges. As we discuss in the following section, the key difference in our proposed technique is that we decouple the underlying unit-cell transform from the average finite-lattice transform. By decoupling these terms, we open the possibility to average many patterns from size- and shape-varying crystals directly in the translationally insensitive three-dimensional reciprocal space. The phase problem may then be solved for the entire data set in a single phase-retrieval step performed directly on a reciprocal-space volume sampled according to the unit-cell size (not the whole crystal size), with signal-to-noise ratio limited, in principle, only by the number of available diffraction patterns. Recent experimental work has indicated the need to merge tens or hundreds of thousands of XFEL diffraction patterns in order to observe intensities between protein-crystal Bragg reflections at moderate resolutions [24]; the low signal detected in individual patterns from protein crystals continues to be the primary challenge to directly applying iterative phasing on a pattern-by-pattern basis. Radiation damage makes this approach particularly challenging at synchrotron sources [36].

In this paper, we present the first proof-of-principle experimental demonstration of the direct determination of a crystal unit-cell structure via the oversampled average over many coherently illuminated nonidentical crystal diffraction patterns, using soft x-ray FEL diffraction data collected from artificial two-dimensional crystal targets. A theoretical description of the method will be given, followed by a description of the samples, experimental apparatus, and a detailed description of the data reduction procedure and phasing algorithm employed. The paper concludes with an analysis of the recovered structures and a discussion of the results.

II. METHOD FOR PHASING SIZE- AND SHAPE-VARYING FINITE CRYSTALS

Consider a finite crystal with real-space unit-cell density, $\rho(\mathbf{r})$, that repeats at the periodically positioned real-space lattice points \mathbf{R}_j , so that the whole-crystal density of the n th crystal may be written as

$$\rho_n(\mathbf{r}) = \sum_j \rho(\mathbf{r} - \mathbf{R}_{nj}). \quad (1)$$

Under the far-field kinematic diffraction approximation, the complex-valued diffraction amplitudes $\tilde{\rho}_n(\mathbf{q})$ corresponding to the n th crystal are proportional to the Fourier transform of $\rho_n(\mathbf{r})$:

$$\tilde{\rho}_n(\mathbf{q}) \propto \sum_j \tilde{\rho}(\mathbf{q}) e^{iq \cdot \mathbf{R}_{nj}}, \quad (2)$$

where the scattering vectors are defined as $\mathbf{q} = 2\pi(\mathbf{s} - \mathbf{s}_0)/\lambda$, with \mathbf{s} and \mathbf{s}_0 the outgoing and incoming wave vectors and λ the monochromatic illumination wavelength, and the tilde symbol denotes the n -dimensional Fourier transform of a function:

$$\tilde{f}(\mathbf{u}) = \int_{-\infty}^{\infty} f(\mathbf{x}) e^{-i\mathbf{u}\cdot\mathbf{x}} d^n x. \quad (3)$$

Note that we have neglected multiplicative terms such as the beam polarization factor, detector solid angle, and incident fluence, since these parameters and functions are known and can be corrected for at the data reduction stage if the need arises.

Under the assumption that the crystal is constructed entirely by repeating a common unit-cell density, the unit-cell transform in Eq. (2) can be moved out of the summation so that the diffraction amplitude can be written as the product of the unit-cell transform and the finite-lattice transform:

$$\tilde{\rho}_n(\mathbf{q}) \propto \tilde{\rho}(\mathbf{q}) \sum_j e^{i\mathbf{q}\cdot\mathbf{R}_{nj}}. \quad (4)$$

The measured diffracted intensity is proportional to the modulus square of the diffracted amplitude, which we may write as

$$I_n(\mathbf{q}) \propto |\tilde{\rho}(\mathbf{q})|^2 S_n(\mathbf{q}) \quad (5)$$

where

$$S_n(\mathbf{q}) = \left| \sum_j e^{i\mathbf{q}\cdot\mathbf{R}_{nj}} \right|^2. \quad (6)$$

If we choose two points \mathbf{q} , \mathbf{q}' in reciprocal space that are shifted with respect to each other by an integer combination of reciprocal-lattice vectors $\mathbf{q}' = \mathbf{q} + \mathbf{g}_{hkl}$, where $\mathbf{g}_{hkl} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$ and \mathbf{a}^* , \mathbf{b}^* , \mathbf{c}^* are reciprocal-lattice vectors, it is clear that the $S(\mathbf{q}) = S(\mathbf{q} + \mathbf{g}_{hkl})$ since $\mathbf{R}_j \cdot \mathbf{g}_{hkl} = m2\pi$, where m is an integer. In other words, the squared modulus of the finite-lattice transform is periodic about the reciprocal-space lattice vectors. Moreover, since there is inversion symmetry about the origin, $S(\mathbf{q}) = S(-\mathbf{q})$, it follows from the periodicity that there is inversion symmetry about every lattice point: $S(\mathbf{g}_{hkl} + \mathbf{q}) = S(\mathbf{g}_{hkl} - \mathbf{q})$.

We now consider the average over many diffraction patterns, which produces the intensity

$$\langle I_n(\mathbf{q}) \rangle_n \propto |\tilde{\rho}(\mathbf{q})|^2 \langle S_n(\mathbf{q}) \rangle_n. \quad (7)$$

If a sufficient number of patterns contribute to this average, $\langle S_n(\mathbf{q}) \rangle_n$ will possess the same symmetry properties as $S_n(\mathbf{q})$. Spence *et al.* [28] suggested that the average squared finite-lattice transform can be determined from the average

diffracted intensity directly by taking the average of $\langle I_n(\mathbf{q}) \rangle_n$ over all translations $\mathbf{q} \rightarrow \mathbf{q} - \mathbf{g}_{hkl}$:

$$\langle \langle I_n(\mathbf{q} - \mathbf{g}_{hkl}) \rangle_n \rangle_{hkl} \propto \langle |\tilde{\rho}(\mathbf{q} - \mathbf{g}_{hkl})|^2 \rangle_{hkl} \langle S_n(\mathbf{q} - \mathbf{g}_{hkl}) \rangle_n. \quad (8)$$

If the terms $|\tilde{\rho}(\mathbf{q})|^2$ and $\langle S_n(\mathbf{q}) \rangle_n$ are uncorrelated with each other, then the average product is equal to the product of averages, and hence

$$\langle \langle I_n(\mathbf{q} - \mathbf{g}_{hkl}) \rangle_n \rangle_{hkl} \propto \langle |\tilde{\rho}(\mathbf{q} - \mathbf{g}_{hkl})|^2 \rangle_{hkl} \langle S_n(\mathbf{q} - \mathbf{g}_{hkl}) \rangle_n. \quad (9)$$

Spence *et al.* further postulated that the periodic average over the isolated unit-cell transform would likely produce a flat function, so one might assume the approximation

$$\langle \langle I_n(\mathbf{q} - \mathbf{g}_{hkl}) \rangle_n \rangle_{hkl} \propto \langle S_n(\mathbf{q} - \mathbf{g}_{hkl}) \rangle_n. \quad (10)$$

The operation described by Eq. (10) is the periodic average over reciprocal-space intensities, which can be accomplished by computing the average peak profile over all reciprocal-space Wigner-Seitz cells (i.e., the smallest primitive unit cell that can be constructed) and then replicating the average profile throughout reciprocal space. With the squared finite-lattice determined directly from the diffraction data in this way, the finely sampled squared unit-cell transform may be expressed as

$$|\tilde{\rho}(\mathbf{q})|^2 \propto \langle I_n(\mathbf{q}) \rangle_n / \langle \langle I_n(\mathbf{q} - \mathbf{g}_{hkl}) \rangle_n \rangle_{hkl}, \quad (11)$$

for which the phases may be determined via well-established iterative phasing algorithms [27]. Spence *et al.* [28] demonstrated this method through simulated diffraction data—here, we extend the proof of principle for this technique by using experimental data from artificial two-dimensional crystals under realistic conditions.

It is important to note that Eq. (11) takes a different form if the crystal density, defined in Eq. (4), does not have identical unit-cell content throughout. As discussed in previous work [30,31], for space groups other than $P1$, it is possible that incomplete unit cells exist at the boundaries of the crystal. If such defects are present, a reformulation of the reconstruction algorithm is required because such crystals do not have a well-defined physical unit cell. In this situation, a reconstruction algorithm might instead seek to determine the molecular asymmetric unit of the crystal, but conventional phasing algorithms developed in the context of CDI are not effective. While a general solution to this problem has not yet been proposed, an approximate solution that works in limited cases has been demonstrated through simulations of crystals bounded by randomized partial unit cells [31]. In this paper, crystals bounded by partial unit cells are not considered, although

we demonstrate that it is the physical crystal boundary that determines the recovered unit cell, as previously postulated.

III. SAMPLE DESIGN AND FABRICATION

In order to perform a controlled experiment on well-characterized targets, we designed finite periodic patterns of Pt islands consisting of simple unit cells containing four symmetry-related elliptical objects, as shown in Fig. 1. The ellipses were positioned with sufficient space between them so as to avoid any complications during fabrication. We designed four different unit-cell configurations, in order to demonstrate the importance of the edge truncation of the crystals. Note that internal densities within each crystal are essentially identical and that the differing unit-cell definitions are distinguished only by the way the edges of the

crystals terminate. As we will show, the edge termination determines which unit cell is reconstructed by the iterative phasing procedure described later. For each of the four unit cell configurations, we designed 20 different crystals of rectangular shape with randomly chosen dimensions ranging from 5 to 12 unit cells along an edge.

We fabricated our two-dimensional crystals by deposition with a focused ion beam (FIB). Patterns were deposited in Pt onto 30-nm-thick Si_3N_4 membranes supported by a Si frame with $100 \times 100 \mu\text{m}$ windows. Pt was deposited via a gas injection system ($(\text{CH}_3)_3\text{Pt}(\text{CpCH}_3)$) with an electron beam current of 43 pA at an energy of 5 keV. The crystal unit-cell size was approximately $1.25 \times 1.25 \mu\text{m}^2$ and the deposition thickness (Pt content of approximately 16%) was about 20 nm. Eighty targets were fabricated in total, an example of which is shown in Fig. 1.

IV. INSTRUMENTATION AND DATA COLLECTION

Measurements were carried out at the coherent diffraction imaging (CDI) experimental station at the DiProI beamline [37] of the FERMI@Elettra FEL [38–40]. We chose to perform our investigations at an FEL source so that the relevant effects of shot-by-shot wavefront phase and intensity variations would be present in our data. Compared to FELs that operate through self-amplified spontaneous emission, the FERMI FEL x-ray pulses are seeded and thus nearly monochromatic. The narrow spectral width ($\Delta\lambda/\lambda \approx 5 \times 10^{-4}$) [39] is ideal for our experiments since, strictly speaking, the lattice transform recovered by our method is periodic only for a monochromatic beam.

The experimental station hosts all necessary components and diagnostics for beam cleaning and sample positioning. The 32.5-nm wavelength FEL beam was focused with a Kirkpatrick-Baez (KB) mirror system [41] and entered the chamber through a circular aperture of 5 mm diameter in order to reduce the stray radiation coming from the beamline. The focus of the KB optical system was optimized to a spot size of about $25 \mu\text{m}$ full width at half maximum using scintillator phosphorus screen and indentations on a Poly(methyl methacrylate) coated silicon wafer placed at the sample plane. The average incoming FEL pulse energy was approximately $27 \pm 5 \mu\text{J}$, with root-mean-square shot-to-shot fluctuations of about 20% during our measurements. The number of pulses incident on the sample was controlled by a fast shutter with pulse intensities controlled with a gas cell and Al filters that allow beam attenuation of up to 4 orders of magnitude. For each exposure (both in single-shot mode and in multishot accumulative mode), the beam intensity and spectrum were acquired shot by shot.

Diffraction patterns were collected using a detection system in an indirect configuration where the scattered x rays were reflected onto a CCD (Princeton Instrument MTE2048B; 2048×2048 pixels, each $13.5 \mu\text{m}$ in size) by

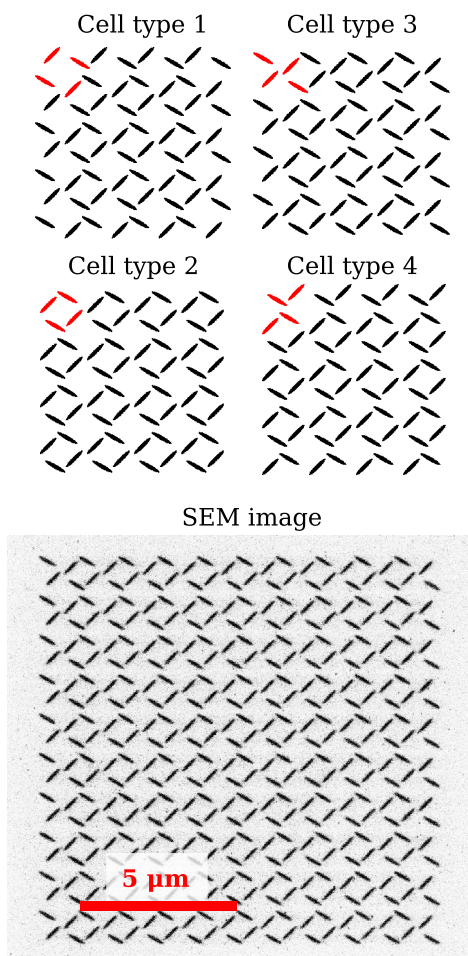


FIG. 1. Top figures: Illustration of the four different crystal types that were designed for the diffraction experiment. The unit-cell contents for each type are indicated in red. The full-size crystals used in the experiment had randomly chosen edge lengths ranging from 5 to 12 unit cells, with the two dimensions varying independently of each other (allowing for rectangular crystals). Bottom figure: SEM image of an actual sample, with distance scale indicated in red.

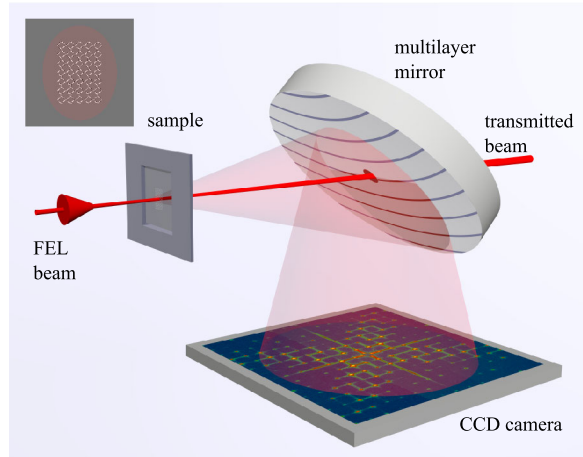


FIG. 2. Schematic illustration of the experiment. After passing through the optics and beam-cleaning aperture, the $\sim 25\text{-}\mu\text{m}$ focused FEL beam illuminates the SiN sample window. The direct FEL beam passes through a hole in the 45° multilayer mirror that reflects the diffracted x rays to the CCD camera. The inset at the top left shows the approximate illumination profile of the FEL beam on a typical microcrystal target.

a 45° multilayer mirror, mounted on a motorized optical gimbal, with a central hole to allow the passage of the direct beam [42]. A schematic of this geometry is shown in Fig. 2. At our operational conditions, the 16-bit detector is saturated at approximately 20,000 photons per pixel. In order to improve the signal-to-noise ratio of detected diffraction, a 200-nm Al filter (transmission of 30%, taking into account a 7-nm Al_2O_3 oxide layer on both sides) was placed upstream of the focusing optics in order to remove the contamination of the 260-nm seeding laser.

The samples were optically aligned onto the FEL beam path using a four-axis manipulator motor stage and a long-range optical microscope. Diffraction patterns were first obtained in accumulative mode, using an attenuated x-ray beam in order to center each target using low-intensity exposures of less than $5\text{ mJ}/\text{cm}^2$. We verified that there were no signs of significant radiation damage at this intensity by comparing a series of successive diffraction patterns. Best centering was determined by inspecting the Bragg reflections in each exposure and seeking the position that produced the greatest degree of inversion symmetry. The breakdown of inversion symmetry about Bragg reflections indicates that the incident wavefront had a nonuniform phase, and we observed that nearly all patterns contained some degree of deviation from ideal inversion symmetry, even when centering was optimized. This is likely because the maximum size of the crystals was $15\text{ }\mu\text{m} \times 15\text{ }\mu\text{m}$, so even the well-centered $25\text{-}\mu\text{m}$ beam did not provide uniform illumination over the crystal. Once the targets were centered, we collected a diffraction pattern at full fluence, which destroyed the target. The final 1/3 of the diffraction patterns were captured at full fluence without careful centering because of time constraints

(we captured diffraction from all 80 targets over the course of about one day).

Figure 3 shows a typical raw diffraction pattern at full fluence, along with three additional shots at lower fluence. The typical number of pixels between Bragg reflections was approximately 105. This detector sampling was sufficient to reveal the interference fringes from our largest crystal targets, although, as we describe in the following sections, we did not require these fine features in our analysis. In addition to the effects of the nonuniform

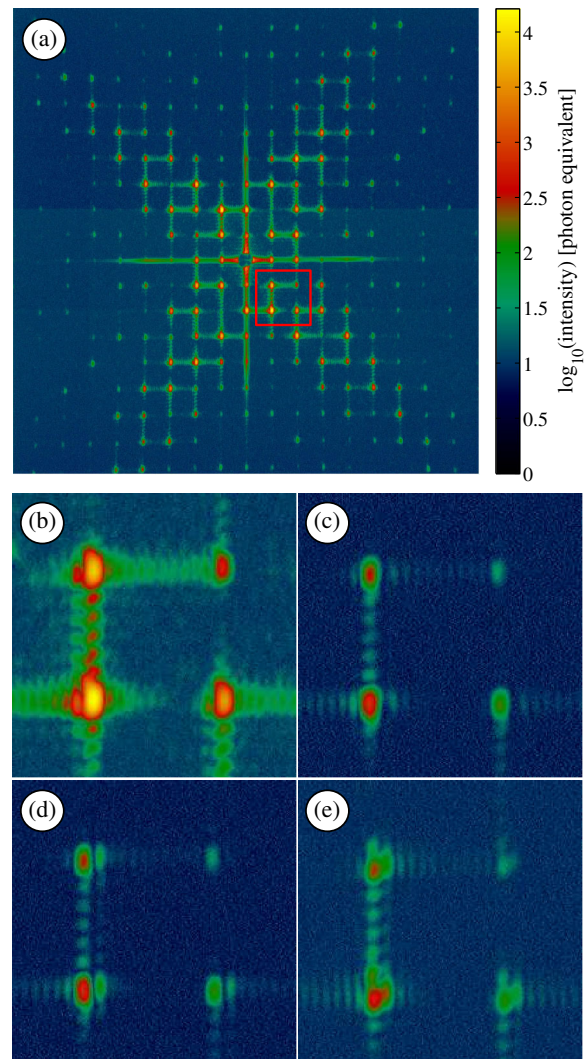


FIG. 3. (a) Single full-fluence shot of a 2D crystal target on a logarithmic intensity scale. The red rectangular box within the full pattern indicates the zoomed regions shown in panels (b-e). Panel (b) corresponds to that in the full pattern (a), while the additional three zoomed regions (c-e) were recorded from the same target but at lower fluence and with the target shifted by a few microns between shots. The breakdown of inversion symmetry, and its sensitivity to beam position, is evident. The prominent central streaks of intensity, which are due to edge scatter from the Si window, were ignored throughout our data analysis.

illuminating wavefront, apparent Bragg peak asymmetries are partly caused by the underlying unit-cell transform, which can be thought of as a modulation of the finite lattice transform according to Eq. (5).

V. DATA ANALYSIS AND REDUCTION

Our data-processing pipeline began with a visual inspection of all patterns. We rejected frames that were clearly damaged or contaminated by foreign debris. We saw no obvious need to mask any malfunctioning pixels, but we ignored the central cross-shaped region of the detector shown in Fig. 3 throughout our analysis since it was contaminated by scatter from the sample Si frame. We also ignored the central region of the diffraction patterns since no intensities were observed because of the hole in the 45° mirror that the direct beam passed through. We assumed the beam center to be the point about which an image had maximum inversion symmetry, which we determined by computing the cross correlation between the image and a copy of itself inverted about its center. (The vector pointing to the peak of this cross-correlation function is twice the length of the vector pointing to the inversion center.) The detector distance of 56.7 mm was determined from the known unit-cell size of the targets.

Since the orientations of each target varied slightly from one to the next, we determined the crystal reciprocal-lattice vectors (relative to the laboratory reference frame) for each pattern via an “auto-indexing” algorithm similar to that described by Steller *et al.* [43] (details can be found in the Appendixes A and B). Our auto-indexing routine was not vulnerable to indexing ambiguities because we knew *a priori* the approximate orientation of each target—we would have otherwise needed to include the peak intensity information in the reciprocal-lattice determination [44].

Using the crystal reciprocal-lattice vectors, we remapped the raw intensity data onto a symmetric orthogonal grid by averaging pixel intensities according to their nearest fractional Miller indices h, k . By remapping intensities according to only two Miller indices, we effectively projected the intensities onto a plane in reciprocal space, which is justified for our targets because they are known *a priori* to be thin; we expect no significant reciprocal-space intensity variation along the normal to the crystal plane. We found that it was also necessary to correct for slight distortions in the 45° mirror that reflected diffraction patterns to the CCD. Details of the remapping procedure are discussed in the Appendix C, and a typical resulting pattern along with predicted and found peaks is displayed in Fig. 4.

The intensity division step in Eq. (11) was found to be highly sensitive to the background signal that was present in our data. Moreover, the subtraction of a background estimated from blank Si_3N_4 targets was not sufficiently accurate for our purposes because the targets had varying amounts of debris that scattered onto the detector, and

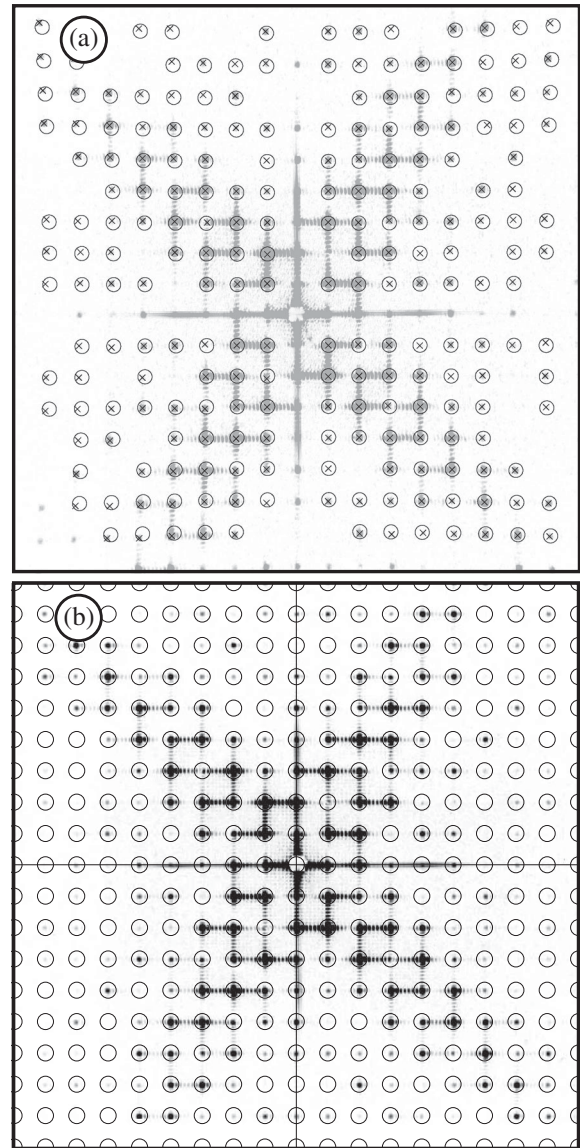


FIG. 4. Comparison between found and predicted Bragg peaks (a) and remapped diffraction intensities projected onto a plane in reciprocal space, after correcting minor distortions caused by the 45° mirror (b). The majority of remapped patterns appeared qualitatively similar to this one, with peaks well centered within the predicted locations.

because the background arising from both optical and x-ray photons was slightly dependent on the position of the sample stage. In order to estimate backgrounds on a frame-by-frame basis, we fit a polynomial function to the intensities averaged within small regions centered at the diagonal midpoints between neighboring Bragg reflections, where we expect little diffraction signal from the crystals. The polynomial functions were used to generate an estimate for the background throughout each entire pattern, which was subsequently subtracted.

For each of the crystal types 1–4 shown in Fig. 1, we averaged a total of 12, 11, 6, and 16 full-fluence

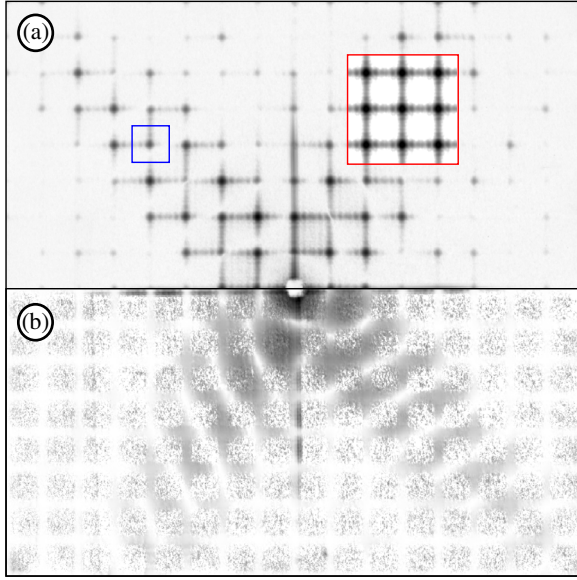


FIG. 5. Extraction of the unit-cell transform. (a) The average over eight diffraction patterns with unit-cell type 4. The intensities in the red box show the extracted average reciprocal-lattice transform, generated by averaging over all Wigner-Seitz cells of the average pattern. A single Wigner-Seitz cell that contributed to this average is indicated by the smaller blue box. (b) The unit-cell transform that results from dividing the average pattern by the average reciprocal-lattice transform. Note that the gridlike patches of noisy pixels are due to the relatively low diffraction signal in those regions.

background-subtracted diffraction patterns, respectively. These counts were lower than the total of 20 patterns that we recorded from each type due to the fact that many full-fluence patterns had a large number of saturated peaks, many of them contained diffraction from debris on the Si_3N_4 membranes, and a few of them were damaged. Each of the four unit-cell transforms were extracted from their averaged crystal diffraction patterns according to Eq. (11), as shown in Figs. 5 and 6. Average lattice transforms were constructed by first identifying the reciprocal-space coordinates of each pixel relative to the center of the nearest Wigner-Seitz cell (specifically, these coordinates are $h - h'$, where h' is rounded to the nearest integer), averaging the peak profile contained within all Wigner-Seitz cells and then distributing the resulting intensity profile periodically about the full diffraction field (the result of this procedure is shown in Fig. 5). When calculating the average reciprocal-lattice transform, we ignored pixels falling nearby the direct beam and the streaks caused by the edges of the Si sample support frame (the central cross shown in Fig. 3). We found that the inclusion of patterns with saturated peaks degraded the resulting unit-cell transforms because the resulting truncated Bragg peaks biased the extracted reciprocal-space lattice transforms.

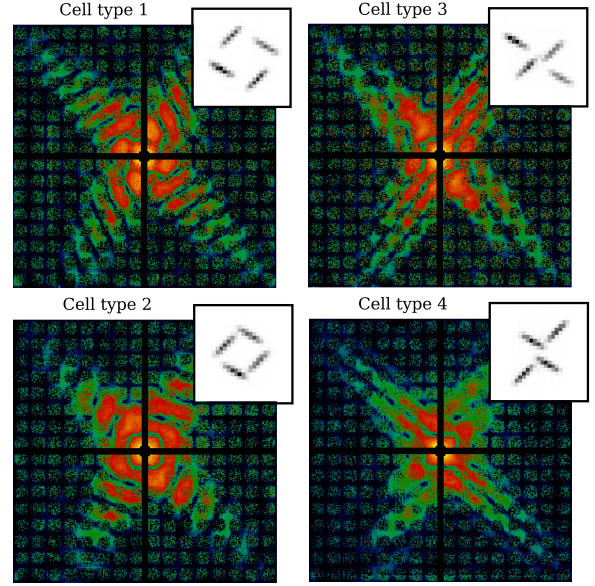


FIG. 6. Extracted unit-cell transforms (on a logarithmic scale) and corresponding unit-cell density reconstructions (linear scale, without an applied support). Iterative phasing was performed on re-binned intensity maps, rather than the finely sampled maps shown here. The gridlike appearance of regions with large errors is caused by the lack of a strong diffraction signal in those regions. Only the intensities where either of the Miller indices are nearly integer-valued were used as Fourier-amplitude constraints, while the remaining intensities with larger errors were allowed to float. Each of the reconstructions shown are the averages over 250 individual images, which were selected from the total of 500 phasing trials via pairwise correlations.

VI. RECONSTRUCTION OF THE UNIT CELLS

Defining $U(\mathbf{q})$ as the squared unit-cell transform $|\tilde{\rho}(\mathbf{q})|^2$ that we extracted from the data as described in the previous section (we have one for each of the four crystal types), we employed two projection operators in the iterative phase-retrieval process. The intensity projection \tilde{P}_I has the action of bringing the magnitudes of the i th estimate of the unit-cell transform, $\tilde{\rho}_i(\mathbf{q})$, into correspondence with the measured intensities:

$$\tilde{P}_I \tilde{\rho}_i(\mathbf{q}) = \begin{cases} \tilde{\rho}_i(\mathbf{q}) \sqrt{\frac{U(\mathbf{q})}{|\tilde{\rho}_i(\mathbf{q})|^2}} & \text{if } \mathbf{q} \in M \\ \tilde{\rho}_i(\mathbf{q}) & \text{otherwise,} \end{cases} \quad (12)$$

where M is the set of constrained intensities (those assumed to be reasonably accurate). The support projection \tilde{P}_S sets the real-space densities to zero in the regions outside of the support S and enforces real and positive values in real space:

$$\tilde{P}_S \tilde{\rho}_i(\mathbf{q}) = \mathcal{F} P_S \mathcal{F}^{-1} \tilde{\rho}_i(\mathbf{q}), \quad (13)$$

where \mathcal{F} is the Fourier transform operator and \mathcal{F}^{-1} its inverse. The real-space component of the projection operator is

$$P_S \rho_i(\mathbf{r}) = \begin{cases} \max(\text{Re}\{\rho_i(\mathbf{r})\}, 0) & \text{if } \mathbf{r} \in S \\ 0 & \text{otherwise,} \end{cases} \quad (14)$$

where $\text{Re}\{\}$ is the function that returns the real part of a complex number, and the max function returns the maximum of its two input values. With the two projection operations \tilde{P}_I and \tilde{P}_S , we utilized the update rule of the hybrid input-output (HIO) algorithm [26],

$$\tilde{\rho}_{i+1}(\mathbf{q}) = (1 + \tilde{P}_S[(1 + \beta)\tilde{P}_I - 1] - \beta\tilde{P}_I)\tilde{\rho}_i(\mathbf{q}), \quad (15)$$

along with the update rule of the error-reduction (ER) algorithm [25],

$$\tilde{\rho}_{i+1}(\mathbf{q}) = \tilde{P}_S \tilde{P}_I \tilde{\rho}_i(\mathbf{q}). \quad (16)$$

Prior to phasing, we rebinned our unit-cell-transform intensities with a coarse sampling corresponding to four points between each pair of Bragg reflections, producing 5×5 -pixel Wigner-Seitz cells. This down-sampling step, which reduced noise levels as well as computing time, is justified because the unit-cell transform varies over longer distance scales than the finite-lattice transform. We ignored noisy intensities by forming a meshlike mask M that rejected pixels with noninteger values for either the h or k Miller indices. (We found that a thicker mesh, as well as circular regions surrounding Bragg reflections, reduced success rates in our phasing trials, presumably because the thin mesh provided the best compromise between number and accuracy of intensities.) We used a value $\beta = 0.9$ and alternated between HIO and ER, beginning with HIO. Each cycle ran for 20 iterations, with 15 HIO steps followed by 5 ER steps, with the support S updated using the Shrinkwrap algorithm [45] at the end of each cycle. The updated support was generated by applying a threshold to the real-space image estimate $\rho_i(\mathbf{q})$ after convolution with a Gaussian kernel, where the width of the Gaussian smoothing kernel was gradually reduced after each cycle. The initial support estimate was generated from a threshold applied to the autocorrelation function $\mathcal{F}^{-1}I(\mathbf{q})$, and the initial estimate $\tilde{\rho}_1(\mathbf{q})$ was taken to be the square root of the full set of measured intensities $U(\mathbf{q})$ with uniformly random phases. Each phasing trial ran for 1000 iterations.

Typical reconstructions are shown in Fig. 6. An example of the mesh of constrained intensities along with the final retrieved floating intensities is shown in Fig. 7. Approximately 50%–100% of all phasing trials produced reconstructions that appeared to be accurate. Cell types 1 and 4 produced the best results (nearly 100% for both), perhaps because these two cells were determined from the greatest number of averaged patterns. For cell types 1, 2,

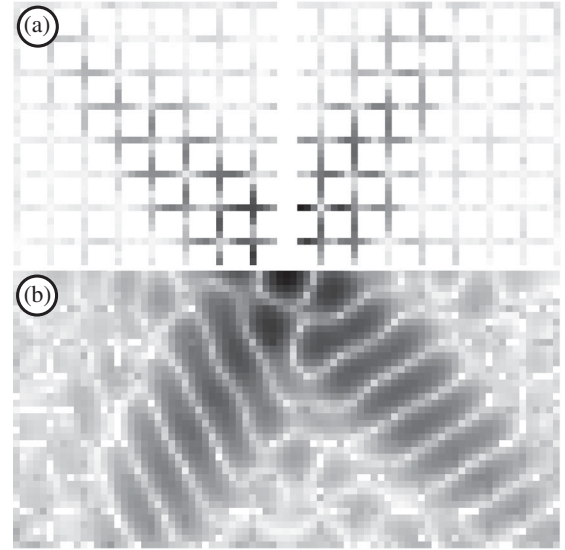


FIG. 7. The down-sampled mesh of intensities used as intensity constraints during phase iterative retrieval (a), and the combined intensity constraints and retrieved floating values (b). Intensities are displayed on a logarithmic scale.

and 4, we also found reasonable convergence without the assertion of a real, positive-valued object, though the apparent quality and fraction of accurate reconstructions was reduced in the absence of these assertions.

We first quantified the errors in our phasing trials using an R factor defined as

$$R = \frac{\sum_i \|\tilde{\rho}(\mathbf{q}_i)\|^2 - U(\mathbf{q}_i)}{\sum_i U(\mathbf{q}_i)}, \quad (17)$$

where i are the indices that correspond to the unmasked intensities. The resulting histograms are shown in Fig. 8 and were formed from 500 independent phasing trials. We note that the highest and lowest mean R factors correspond to the cell types with the fewest and greatest number of contributing patterns, respectively, as one might expect.

We further quantified the quality of our reconstructions with the phase-retrieval transfer function defined as

$$\text{PRTF}(\mathbf{q}_i) = |\langle \exp(i\phi(\mathbf{q}_i)) \rangle_i| = \left| \frac{\langle \tilde{\rho}(\mathbf{q}_i) \rangle}{|\tilde{\rho}(\mathbf{q}_i)|} \right|, \quad (18)$$

which effectively measures the consistency of the recovered phases. The phase-retrieval transfer function (PRTF) is sensitive to phase ramps that result from the relative shift of each reconstruction in the image plane. Therefore, prior to computing the PRTF, we shifted each reconstruction to best match a template (the first reconstruction) via the up-sampled cross-correlation procedure described in Ref. [46]. Since as many as 50% of the reconstructions were visibly inconsistent with the others, we calculated the PRTFs only

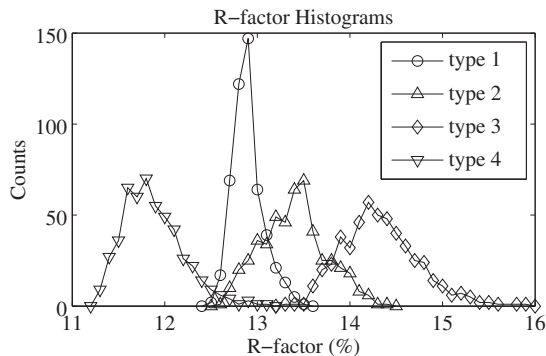


FIG. 8. R -factor histograms formed from 500 independent phasing trials.

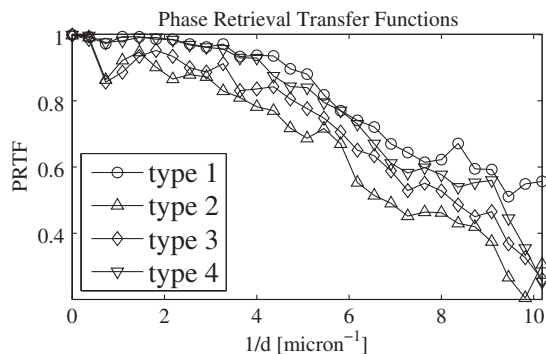


FIG. 9. Phase-retrieval transfer functions as a function of inverse resolution d^{-1} , formed from the 250 patterns that correlated best, on average, with the other 500 phasing trials.

for the 50% of patterns that correlated best on average with all of the other patterns, as determined by the Pearson correlations computed from all pairs of reconstructions for a given unit-cell type. The PRTFs are shown in Fig. 9, which maintain values of greater than 0.5 to resolutions better than 150 nm.

VII. DISCUSSION AND CONCLUSIONS

The results presented here provide the first experimental proof-of-principle demonstration of the method for phasing coherently illuminated crystals originally proposed by Spence *et al.* [28]. We have demonstrated that this method for decoupling the average crystal-lattice transform from the underlying unit-cell transform is effective when the phase and intensity of the illuminating wavefront is noticeably nonuniform and varies from one crystal to the next. We have also demonstrated that the method works in the presence of noise, to the extent that many of the intensity data are completely unreliable and must be ignored. The subset of data at high signal-to-noise ratio nonetheless provided sufficient information for accurate object reconstructions. A data-processing strategy that copes with nonuniform background signal was described.

Despite the fact that all four crystal types that we fabricated were identical in their internal structures, the significance of the crystal boundaries was clearly demonstrated by the four different unit-cell configurations that we reconstructed from each type. As noted in previous work [31,47], the situation becomes more complicated if the finite crystals are not composed of identical real-space unit cells. While we did not consider crystals bounded by partial unit cells here, the present results suggest that methods for coping with partial unit cells such as the special case presented by Kirian *et al.* [31] will also work with experimental data of quality similar to those considered here.

It is important to note that our crystal targets have a high “solvent” fraction (regions of uniform density) that generously exceeds 50%, in which case there is sufficient information to solve the phase problem using only the integrated Bragg peak intensities. We indeed found this to be true for simulations of our data, as we were able to reconstruct a unit cell from the Bragg intensities alone using a combination of the HIO [26] and Shrinkwrap [45] algorithms. Likely, we could also obtain reasonable results with real data, though our attempts thus far have not produced accurate reconstructions. However, as discussed elsewhere [31], a high solvent fraction is not a *requirement* of the presented phasing method; the principle advantage of having access to a continuous, oversampled intensity function is that it overdetermines the phase problem in general, without restrictions on resolution or solvent fraction. The approach presented also allows for a good initial estimate of the object support through the autocorrelation function, whereas previous work has required prior knowledge of a molecular envelope when applying the HIO algorithm to Bragg intensity data from protein crystals with a high solvent fraction [21].

Improvements beyond our present results can be expected if we include knowledge of noise levels in our phasing, as discussed in Refs. [48–50]. Experimental errors are an important challenge in SFX data analysis in general, and it will likely be necessary to address this in future work aimed at phasing protein-crystal targets that have more complex electron densities and where the signal-to-noise ratios are far lower than in the data considered here. An alternative approach to phasing has been suggested by Elser [51], where only the strong intensities located at the Bragg conditions and the associated intensity gradients are utilized for phase determination. In simulations, this algorithm was shown to be effective at high resolution and at signal-to-noise ratios of 2 in the intensity gradients.

It should also be noted that the interactive target positioning that we carried out at low doses for the majority of our data collection is not possible using existing technology for delivering protein nanocrystals to the XFEL beam [52,53]. However, in a protein SFX experiment, a small-angle detector may be used to inspect finely sampled shape transforms, as demonstrated by

Chapman *et al.* [1]. If necessary, poorly positioned targets that are intercepted at severely distorted regions of the illuminating wavefront can be rejected prior to merging the diffraction data. As we have demonstrated here, minor (but clearly observable) phase and amplitude distortions can likely be neglected in most situations where the focal spot size spans several unit cells. This might be understood by the fact that the intensities recovered by averaging many patterns with a randomly positioned, nonuniform wavefront are equal to the convolution between the modulus-squared wavefront transform and the modulus-squared object transform. The procedure we have applied here would then produce a convolved unit-cell transform. Real-space reconstructions are possible without the need to incorporate knowledge of the wavefront, provided that this convolution is not severe (i.e., that the beam size is not severely distorted and that it is large enough to span many unit cells).

ACKNOWLEDGEMENTS

We thank Joe Chen and Rick Millane for many insightful discussions. We acknowledge support from the Helmholtz Association through program-oriented funds, the Bundesministerium für Bildung und Forschung through Projects No. 05K2012, No. 05K12CH1, and No. 05E13GU1, the Deutsche Forschungsgemeinschaft through the Centre for Ultrafast Imaging, and the National Science Foundation through the Science and Technology Center Grant No. 1231306. FERMI is a project of Elettra-Sincrotrone Trieste, partially supported by the Italian Ministry of University and Research under Grants No. FIRB- RBAP045JF2 and No. FIRB-RBAP06AWK3 and by a grant from the Friuli Venezia Giulia Region: Nanotex 0060-2009.

APPENDIX A: CRYSTAL LATTICE DETERMINATION

We first determined Bragg peak positions by convolving each image with a function consisting of a positive-valued top hat and negative-valued concentric annulus. The values of the top hat and annulus were chosen such that the resulting map was equivalent to locally averaged intensities with local background subtracted. Peaks were identified as the centroids of regions in which several connected pixels had local average intensities above a threshold (all tunable parameters were determined by visual inspection of the results). The reciprocal-space vectors corresponding to Bragg peaks, \mathbf{p}_m where m is the peak number, were determined according to

$$\mathbf{p}_m = \frac{2\pi}{\lambda} (\hat{\mathbf{v}}_{i_m j_m} - \hat{\mathbf{b}}), \quad (\text{A1})$$

where $\hat{\mathbf{v}}_{i_m j_m}$ is the vector pointing from the sample to the peak centroid (which has noninteger pixel coordinates

i_m, j_m), and $\hat{\mathbf{b}}$ points in the incident beam direction. The auto-indexing algorithm identified unit vectors $\hat{\mathbf{d}}_\omega$ that maximized the periodicity of the histograms $H_\omega(x)$ formed from the projected values $x_m = |\hat{\mathbf{d}}_\omega \cdot \mathbf{p}_m|$. This criterion is met when the first peak (excluding the origin) of the Fourier transform of $H_\omega(x)$ is maximized within the limited angular search space ω (within 10 degrees of the guessed value). A second vector $\hat{\mathbf{d}}'_\omega$ was found in an angular search space assumed to be nearly orthogonal to the first. The reciprocal lattice vectors were taken to be $\mathbf{a}^* = \hat{\mathbf{d}}_\omega / \eta \Delta x$ and $\mathbf{b}^* = \hat{\mathbf{d}}'_\omega / \eta' \Delta x$, where Δx is the bin size of $H_\omega(x)$ and η is the noninteger bin number corresponding to the center of the peak position.

It is useful to place the result of the auto-indexing routine into a 3×3 reciprocal-lattice matrix \mathbf{U} with column vectors equal to $[\mathbf{a}^* \ \mathbf{b}^* \ \mathbf{c}^*]$. The third lattice vector is chosen to be $\mathbf{c}^* = \mathbf{a}^* \times \mathbf{b}^* / |\mathbf{a}^* \times \mathbf{b}^*|$ for convenience, though its magnitude is arbitrary because of the lack of periodicity along the third direction.

APPENDIX B: PEAK PREDICTION

Three-dimensional reciprocal-space coordinates are mapped to detector pixels with indices i, j according to the formula

$$\mathbf{q}_{ij} = \frac{2\pi}{\lambda} (\hat{\mathbf{v}}_{ij} - \hat{\mathbf{b}}), \quad (\text{B1})$$

where λ is the wavelength, $\hat{\mathbf{b}}$ is the incident beam direction, and $\hat{\mathbf{v}}_{ij}$ is the unit vector pointing from the x-ray-target interaction point to the center of pixel i, j :

$$\mathbf{v}_{ij} = \mathbf{t} + i\mathbf{f} + j\mathbf{s}. \quad (\text{B2})$$

The vector \mathbf{t} points from the interaction point to the center of the corner detector pixel, \mathbf{f} points along rows of pixels, and \mathbf{s} points along columns of pixels. We found that the assumption of mutually perpendicular \mathbf{f} , \mathbf{s} , and $\hat{\mathbf{b}}$ was satisfactory in our data reduction scheme.

The reciprocal-lattice vectors allow reciprocal-space peak positions to be predicted by finding the intersection between the points $\boldsymbol{\epsilon}$ on the Ewald sphere that satisfy

$$|\boldsymbol{\epsilon} + \mathbf{s}_0|^2 = s_0^2 \quad (\text{B3})$$

and a line corresponding to Miller indices h, k defined as

$$\boldsymbol{\epsilon} = \mathbf{o}_{hk} + \mathbf{l}_{hk}, \quad (\text{B4})$$

where the vector $\mathbf{o}_{hk} = \mathbf{U} \cdot \mathbf{h}$, \mathbf{U} is the 3×3 reciprocal-lattice matrix containing the column vectors $[\mathbf{a}^* \ \mathbf{b}^* \ \mathbf{c}^*]$, \mathbf{h} is the column vector $[h \ k \ 0]$, and \mathbf{l}_{hk} is the vector pointing in the direction of $\mathbf{c}^* = \mathbf{a}^* \times \mathbf{b}^*$. The two possible solutions to this system of equations are

$$l_{hk}^{\pm} = -\hat{l}_{hk} \cdot (\mathbf{o}_{hk} + \mathbf{s}_0) \pm \sqrt{(\hat{l}_{hk} \cdot (\mathbf{o}_{hk} + \mathbf{s}_0))^2 - (\mathbf{o}_{hk} + \mathbf{s}_0)^2 + s_0^2}, \quad (\text{B5})$$

and thus the predicted peak location is

$$\mathbf{p}'_{hk} = \min(\mathbf{U} \cdot \mathbf{h} + l_{hk}^{\pm}), \quad (\text{B6})$$

where the min function selects the vector with minimum length (this is appropriate for our data, where the vector normal to the sample plane was nearly parallel to the incident beam direction). The coordinates i_{hk}, j_{hk} of the predicted peaks in the detector plane are equal to

$$i_{hk} = \left(\frac{\lambda}{2\pi} \mathbf{p}'_{hk} + \hat{\mathbf{b}} - \mathbf{t} \right) \cdot \hat{\mathbf{f}}, \quad (\text{B7})$$

$$j_{hk} = \left(\frac{\lambda}{2\pi} \mathbf{p}'_{hk} + \hat{\mathbf{b}} - \mathbf{t} \right) \cdot \hat{\mathbf{s}}. \quad (\text{B8})$$

APPENDIX C: INTENSITY REMAPPING

We remapped the raw intensity data I_{ij} onto a symmetric orthogonal grid I'_{hk} by averaging values in I_{ij} according to their nearest fractional Miller indices h, k . The mapping from the native detector indices to the Miller indices is given by the expression $\mathbf{h} = \mathbf{U}^{-1} \mathbf{q}_{ij}$, where \mathbf{U} is the 3×3 reciprocal-lattice matrix determined by the auto-indexing routine and \mathbf{h} is the column vector $[h \ k \ l]$.

As shown in Fig. 4, there are systematic errors between the predicted and found Bragg peaks in our patterns, which were nearly identical for all diffraction patterns. We concluded that the observed systematic errors were likely caused by small slope errors in the 45° mirror that reflects the diffracted x rays to the CCD because we found no experimental geometry that could reproduce the patterns with reasonable accuracy throughout the detector plane. In order to rectify the discrepancies between predicted and found peak locations, we noted that the offsets between peaks could be reasonably well described with smoothly varying functions. We assumed that the correct indices i', j' (those corresponding to measurements made in the absence of mirror distortions) could be related to the native detector coordinates with the formula $(i', j') = [i + f_1(i, j), j + f_2(i, j)]$, where $f_n(i, j)$ are two-dimensional fifth-order polynomial functions. The polynomial coefficients were determined via linear least-squares minimization using the observed residuals $\Delta i_m = i'_m - i_m$, where i'_m was taken to be the prediction corresponding to the found peak m . The rectified reciprocal-space vectors are equal to $\mathbf{q}'_{ij} = \mathbf{q}_{i'j'}$. This method of determining mirror slope errors, based on diffraction from a periodic calibration object, would also be useful for other coherent diffractive imaging experiments that utilize a similar experimental geometry.

- [1] H. N. Chapman *et al.*, *Femtosecond X-Ray Protein Nanocrystallography*, *Nature (London)* **470**, 73 (2011).
- [2] J. C. H. Spence, U. Weierstall, and H. N. Chapman, *X-Ray Lasers for Structural and Dynamic Biology*, *Rep. Prog. Phys.* **75**, 102601 (2012).
- [3] I. Schlichting and J. Miao, *Emerging Opportunities in Structural Biology with X-Ray Free-Electron Lasers*, *Curr. Opin. Struct. Biol.* **22**, 613 (2012).
- [4] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, *Potential for Biomolecular Imaging with Femtosecond X-Ray Pulses*, *Nature (London)* **406**, 752 (2000).
- [5] A. Barty *et al.*, *Self-Terminating Diffraction Gates Femtosecond X-Ray Nanocrystallography Measurements*, *Nat. Photonics* **6**, 35 (2012).
- [6] L. Redecke *et al.*, *Natively Inhibited Trypanosoma Brucei Cathepsin B Structure Determined by Using an X-Ray Laser*, *Science* **339**, 227 (2013).
- [7] W. Liu *et al.*, *Serial Femtosecond Crystallography of G Protein-Coupled Receptors*, *Science* **342**, 1521 (2013).
- [8] S. Boutet *et al.*, *High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography*, *Science* **337**, 362 (2012).
- [9] M. G. Rossmann, *The Molecular Replacement Method*, *Acta Crystallogr. Sect. A* **46**, 73 (1990).
- [10] H. Hauptman, *Phasing Methods for Protein Crystallography*, *Curr. Opin. Struct. Biol.* **7**, 672 (1997).
- [11] M. F. Perutz, *Isomorphous Replacement and Phase Determination in Non-centrosymmetric Space Groups*, *Acta Crystallogr.* **9**, 867 (1956).
- [12] D. W. Green, V. M. Ingram, and M. F. Perutz, *The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method*, *Proc. R. Soc. A* **225**, 287 (1954).
- [13] R. B. G. Ravelli, H. K. S. Leiros, B. C. Pan, M. Caffrey, and S. McSweeney, *Specific Radiation Damage Can Be Used to Solve Macromolecular Crystal Structures*, *Structure* **11**, 217 (2003).
- [14] T. R. M. Barends, L. Foucar, S. Botha, R. Bruce Doak, R. L. Shoeman, K. N., J. E. Koglin, G. J. Williams, S. Boutet, M. Messerschmidt, and I. Schlichting, *De Novo Protein Crystal Structure Determination from X-Ray Free-Electron Laser Data*, *Nature (London)*, **505**, 244 (2013).
- [15] S.-K. Son, H. N. Chapman, and R. Santra, *Multiwavelength Anomalous Diffraction at High X-Ray Intensity*, *Phys. Rev. Lett.* **107**, 218102 (2011).
- [16] J. W. Miao, P. Charalambous, J. Kirz, and D. Sayre, *Extending the Methodology of X-Ray Crystallography to Allow Imaging of Micrometre-Sized Non-crystalline Specimens*, *Nature (London)* **400**, 342 (1999).
- [17] *Science of Microscopy*, edited by P. W. Hawkes and J. C. H. Spence (Springer, New York, NY, 2005).
- [18] H. N. Chapman *et al.*, *Femtosecond Diffractive Imaging with a Soft-X-Ray Free-Electron Laser*, *Nat. Phys.* **2**, 839 (2006).
- [19] H. N. Chapman and K. A. Nugent, *Coherent Lensless X-Ray Imaging*, *Nat. Photonics* **4**, 833 (2010).
- [20] D. Sayre, *Some Implications of a Theorem due to Shannon*, *Acta Crystallogr. Sect. A* **5**, 843 (1952).
- [21] Z.-C. Liu, R. Xu, and Y.-H. Dong, *Phase Retrieval in Protein Crystallography*, *Acta Crystallogr. Sect. A* **68**, 256 (2012).

- [22] R. P. Millane and V. L. Lo, *Iterative Projection Algorithms in Protein Crystallography. I. Theory*, *Acta Crystallogr. Sect. A* **69**, 517 (2013).
- [23] M. F. Perutz, *The Structure of Haemoglobin. 3. Direct Determination of the Molecular Transform*, *Proc. R. Soc. A* **225**, 264 (1954).
- [24] O. M. Yefanov, C. Gati, G. Bourenkov, R. A. Kirian, T. A. White, J. C. H. Spence, H. N. Chapman, and A. Barty, *Mapping the Continuous Reciprocal Space Intensity Distribution of X-Ray Serial Crystallography*, *Phil. Trans. R. Soc. B* **369**, 20130333 (2014).
- [25] R. W. Gerchberg and W. O. Saxton, *Practical Algorithm for Determination of Phase from Image and Diffraction Plane Pictures*, *Optik (Stuttgart)* **35**, 237 (1972).
- [26] J. R. Fienup, *Reconstruction of an Object from the Modulus of Its Fourier Transform*, *Opt. Lett.* **3**, 27 (1978).
- [27] S. Marchesini, *A Unified Evaluation of Iterative Projection Algorithms for Phase Retrieval*, *Review Sci. Instrum.* **78**, 011301 (2007).
- [28] J. C. H. Spence, R. A. Kirian, X. Wang, U. Weierstall, K. E. Schmidt, T. White, A. Barty, H. N. Chapman, S. Marchesini, and J. Holton, *Phasing of Coherent Femtosecond X-Ray Diffraction from Size-Varying Nanocrystals*, *Opt. Express* **19**, 2866 (2011).
- [29] R. A. Dilanian, V. A. Streltsov, H. M. Quiney, and K. A. Nugent, *Continuous X-Ray Diffractive Field in Protein Nanocrystallography*, *Acta Crystallogr. Sect. A* **69**, 108 (2013).
- [30] J. P. J. Chen and R. P. Millane, *Diffraction by Nanocrystals*, *J. Opt. Soc. Am. A* **30**, 2627 (2013).
- [31] R. A. Kirian, R. J. Bean, K. R. Beyerlein, O. M. Yefanov, T. A. White, A. Barty, and H. N. Chapman, *Phasing Coherently Illuminated Nanocrystals Bounded by Partial Unit Cells*, *Phil. Trans. R. Soc. B* **369**, 20130331 (2014).
- [32] J. P. J. Chen, J. C. H. Spence, and R. P. Millane, *Direct Phasing in Femtosecond Nanocrystallography. I. Diffraction Characteristics*, *Acta Crystallogr. Sect. A* **70**, 143 (2014).
- [33] J. P. J. Chen, J. C. H. Spence, and R. P. Millane, *Direct Phasing in Femtosecond Nanocrystallography. II. Phase Retrieval*, *Acta Crystallogr. Sect. A* **70**, 154 (2014).
- [34] H. Liu, N. A. Zatsepin, and J. C. H. Spence, *Ab-Initio Phasing Using Nanocrystal Shape Transforms with Incomplete Unit Cells*, *IUCrJ* **1**, 19 (2013).
- [35] J. Miao and D. Sayre, *On Possible Extensions of X-Ray Crystallography through Diffraction-Pattern Oversampling*, *Acta Crystallogr. Sect. A* **56**, 596 (2000).
- [36] S. Boutet and I. K. Robinson, *Coherent X-Ray Diffractive Imaging of Protein Crystals*, *J. Synchrotron Radiat.* **15**, 576 (2008).
- [37] F. Capotondi *et al.*, *Invited Article: Coherent Imaging Using Seeded Free-Electron Laser Pulses with Variable Polarization: First Results and Research Opportunities*, *Rev. Sci. Instrum.* **84**, 051301 (2013).
- [38] E. Allaria, C. Callegari, D. Cocco, W. M. Fawley, M. Kiskinova, C. Masciovecchio, and F. Parmigiani, *The FERMI@Elettra Free-Electron-Laser Source for Coherent X-Ray Physics: Photon Properties, Beam Transport System and Applications*, *New J. Phys.* **12**, 075002 (2010).
- [39] E. Allaria *et al.*, *Highly Coherent and Stable Pulses from the FERMI Seeded Free-Electron Laser in the Extreme Ultraviolet*, *Nat. Photonics* **6**, 699 (2012).
- [40] E. Allaria *et al.*, *Two-Stage Seeded Soft-X-Ray Free-Electron Laser*, *Nat. Photonics* **7**, 913 (2013).
- [41] L. Raimondi *et al.*, *Microfocusing of the FERMI@Elettra FEL Beam with a K-B Active Optics System: Spot Size Predictions by Application of the WISE Code*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **710**, 131 (2013).
- [42] S. Bajt, H. N. Chapman, E. A. Spiller, J. B. Alameda, B. W. Woods, M. Frank, M. J. Bogan, A. Barty, S. Boutet, S. Marchesini, S. P. Hau-Riege, J. Hajdu, and D. Shapiro, *Camera for Coherent Diffractive Imaging and Holography with a Soft-X-Ray Free-Electron Laser*, *Appl. Opt.* **47**, 1673 (2008).
- [43] I. Steller, R. Bolotovskiy, and M. G. Rossmann, *An Algorithm for Automatic Indexing of Oscillation Images Using Fourier Analysis*, *J. Appl. Crystallogr.* **30**, 1036 (1997).
- [44] W. Brehm and K. Diederichs, *Breaking the Indexing Ambiguity in Serial Crystallography*, *Acta Crystallogr.* **D70**, 101 (2014).
- [45] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence, *X-Ray Image Reconstruction from a Diffraction Pattern Alone*, *Phys. Rev. B* **68**, 140101 (2003).
- [46] J. R. Fienup, *Invariant Error Metrics for Image Reconstruction*, *Appl. Opt.* **36**, 8352 (1997).
- [47] R. P. Millane and J. P. J. Chen, *Aspects of Direct Phasing in Femtosecond Nanocrystallography*, *Phil. Trans. R. Soc. B* **369**, 20130498 (2014).
- [48] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. H. Spence, U. Weierstall, T. Beetz, C. Jacobsen, and D. Shapiro, *High-Resolution Ab Initio Three-Dimensional X-Ray Diffraction Microscopy*, *J. Opt. Soc. Am. A* **23**, 1179 (2006).
- [49] R. A. Dilanian, G. J. Williams, L. W. Whitehead, D. J. Vine, A. G. Peele, E. Balaur, I. McNulty, H. M. Quiney, and K. A. Nugent, *Coherent Diffractive Imaging: A New Statistically Regularized Amplitude Constraint*, *New J. Phys.* **12**, 093042 (2010).
- [50] A. V. Martin, F. Wang, N. D. Loh, T. Ekeberg, F. R. N. C. Maia, M. Hantke, G. van der Schot, C. Y. Hampton, R. G. Sierra, and A. Aquila, *Noise-Robust Coherent Diffractive Imaging with a Single Diffraction Pattern*, *Opt. Express* **20**, 16650 (2012).
- [51] V. Elser, *Direct Phasing of Nanocrystal Diffraction*, *Acta Crystallogr. Sect. A* **69**, 559 (2013).
- [52] U. Weierstall, J. C. H. Spence, and R. B. Doak, *Injector for Scattering Measurements on Fully Solvated Biospecies*, *Rev. Sci. Instrum.* **83**, 035108 (2012).
- [53] U. Weierstall, *Lipidic Cubic Phase Injector Facilitates Membrane Protein Serial Femtosecond Crystallography*, *Nat. Commun.* **5**, 1 (2014).