

ARTICLE

DOI: 10.1038/s41467-018-04830-4

OPEN

# Structure determination from single molecule X-ray scattering with three photons per image

Benjamin von Ardenne<sup>1</sup>, Martin Mechelke<sup>1</sup> & Helmut Grubmüller<sup>1</sup>

Scattering experiments with femtosecond high-intensity free-electron laser pulses provide a new route to macromolecular structure determination. While currently limited to nanocrystals or virus particles, the ultimate goal is scattering on single biomolecules. The main challenges in these experiments are the extremely low signal-to-noise ratio due to the very low expected photon count per scattering image, often well below 100, as well as the random orientation of the molecule in each shot. Here we present a de novo correlation-based approach and show that three coherently scattered photons per image suffice for structure determination. Using synthetic scattering data of a small protein, we demonstrate near-atomic resolution of 3.3 Å using  $3.3 \times 10^{10}$  coherently scattered photons from  $3.3 \times 10^9$  images, which is within experimental reach. Further, our three-photon correlation approach is robust to additional noise from incoherent scattering; the number of disordered solvent molecules attached to the macromolecular surface should be kept small.

<sup>1</sup>Department of Theoretical and Computational Biophysics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. Correspondence and requests for materials should be addressed to H.Gül. (email: [hgrubmu@mpibpc.mpg.de](mailto:hgrubmu@mpibpc.mpg.de))

First proposed by Neutze et al.<sup>1</sup>, single-particle scattering experiments with high-intensity X-ray free-electron lasers (XFELs) hold the promise to solve the three-dimensional atomic structure of biological macromolecules such as proteins without the need for crystallization<sup>2–5</sup>. High-repetition femtosecond X-ray pulses are used to outrun the severe radiation damage due to Auger decay and Coulomb explosion and thus allow for extremely high peak brilliance pulses to the point where single molecules can be imaged. Indeed, the first proof of principle experiments<sup>6,7</sup> determined the 3D structure of single mimivirus particles to a resolution of 125 nm and Hosseinizadeh et al. recently demonstrated the structure determination of a coliphage virus with 9 nm resolution<sup>8</sup>. In these experiments, more than  $10^7$  photons per X-ray pulse were scattered by the virus and recorded on a pixel detector (Fig. 1a). In contrast, for a medium-sized molecule and an expected XFEL fluence of  $1.3 \times 10^6$  photons  $\text{nm}^{-2}$  ( $10^{12}$  photons) at a 1  $\mu\text{m}$  focus diameter<sup>9</sup>, only about 10–50 coherently scattered photons per scattering image are expected at a beam energy of 5 keV (2.5 Å wavelength)<sup>9–11</sup>.

The high statistical noise in this extreme Poisson regime poses considerable methodological challenges, and hence XFEL structure determination attempts almost exclusively focus on nanocrystals<sup>12–18</sup>. A particular challenge is to determine the orientation of the molecule for each image to assemble all recorded images in 3D Fourier space for subsequent electron density determination. For macroscopic 2D objects and 3D objects rotated around a single axis, Philipp et al. showed structure recovery from only 2.5 photons per image on average<sup>19–21</sup>, but the method was not extended or applied to three-dimensional objects or molecules with unknown orientation. For single-molecule scattering experiments, several orientation determination methods were developed<sup>22–29</sup>, which, however, require at least 100 photons per image. Alternatively, manifold reconstruction algorithms (manifold embedding)<sup>30–33</sup> forego the explicit assembly in Fourier space and instead use the similarity between scattering images to determine the manifold of orientations. However, also for these methods, successful structure determination was only reported for much more than 100 photons per image.

In fluorescence microscopy or cryo-electron microscopy, time integrated and time-correlated single-photon counting is used at extremely low signal-to-noise ratios<sup>34</sup>. In the context of single-molecule X-ray scattering, two-photon correlations were successfully used to determine the molecular shape of symmetric particles<sup>35,36</sup> and the structure of particles randomly oriented around one axis<sup>37,38</sup>. However, two photons are not sufficient to retrieve the structure de novo.

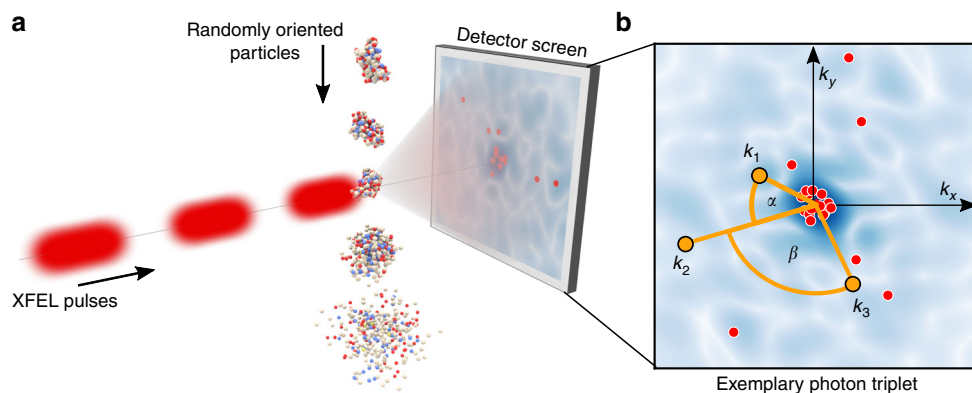
Based on early analytic work on degenerate three-photon correlations<sup>39</sup>, structure determination of mesoscopic cylindrical particles<sup>40</sup> and of a highly symmetric icosahedral virus<sup>41,42</sup> was demonstrated. This approach is limited to only a small fraction of the recorded correlations; however, also this method has so far not been applied to de novo single-molecule structure determination.

Here, we use the full three-photon correlation as an orientation-independent representation of the scattering images. We demonstrate that only three coherently scattered photons per image are required for de novo structure determination, such that near-atomic resolution for single biomolecules should in principle be possible even at extremely low photon counts.

## Results

**Structure determination.** Like in X-ray crystallography, the photon distribution of each scattering image follows the intersection between the Ewald sphere and the 3D intensity,  $I(\mathbf{k}) \propto |\mathcal{FT}[\rho(\mathbf{x})]|^2$ , which is proportional to the absolute square of the Fourier-transformed electron density  $\rho(\mathbf{x})$ . The orientation of the Ewald sphere depends on the molecular orientation and so does the scattering image. In contrast to X-ray crystallography,  $I(\mathbf{k})$  is continuous for single-molecule scattering, rendering the phase problem accessible to established methods<sup>43–46</sup>. Because the orientation of the molecule is unknown, here  $I(\mathbf{k})$  is determined via the three-photon correlation function  $t(k_1, k_2, k_3, \alpha, \beta)$  which is accumulated from all photon triplets in the recorded scattering images as illustrated in Fig. 1b.

To recover  $I(\mathbf{k})$ , an analytic expression of the full three-photon correlation as a function of the 3D intensity  $I(\mathbf{k})$  was derived using shell-wise spherical harmonics (SH) expansions<sup>47</sup> for  $I(\mathbf{k}) = \sum_{lm} A_{lm}(|\mathbf{k}|) Y_{lm}(\theta, \varphi)$  (Methods and Supplementary Notes 1–3). This choice allows for adapting the number  $K(L^2 + 3L + 2)/2$  of SH basis functions to the target resolution via the largest considered wave number  $k_{\text{cut}}$ , the number  $K$  of used shells between  $0 \dots k_{\text{cut}}$ , and the expansion order  $L$ . We were unable to invert the analytic expression of the three-photon correlation, and the number of unknowns (e.g., 4940 for  $K = 26$ ,  $L = 18$ ) is too large for a straightforward numeric solution. To circumvent this problem, we used a probabilistic approach and solved for those SH coefficients  $\{A_{lm}(k)\}$  that maximize the probability,  $p(\{k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i\} | \{A_{lm}(k)\}) = \prod_{i=1 \dots T} \tilde{t}(k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i)_{\{A_{lm}(k)\}}$  (Bayesian with uniform prior), of observing all  $T$  recorded triplets (Methods and Supplementary Notes 4 and 5). Due to their statistical independence,  $p$  is the product of the probabilities of



**Fig. 1** Single-molecule scattering and photon correlations. **a** A stream of randomly oriented particles is injected into the XFEL beam, hit sequentially by femtosecond X-ray pulses, and the coherently scattered photons (red dots) are recorded on the pixel detector. **b** In the detector plane  $k_x, k_y$ , the recorded photons are grouped into triplets, each of which is characterized by distances  $k_1, k_2, k_3$  to the detector center (orange lines) and the angles  $\alpha$  and  $\beta$  between the respective photons (orange circular arcs)

observing each recorded photon triplet which is given by the normalized three-photon correlation  $\tilde{t}(k_1, k_2, k_3, \alpha, \beta)_{\{A_m(k)\}}$ . The search space was further reduced by utilizing the analytic inversion of the two-photon correlation<sup>39</sup> (Methods and Supplementary Note 6), rendering the problem accessible to Monte Carlo simulated annealing<sup>48</sup>. We found that independent Monte Carlo runs converged to similar intensities (Pearson correlation of 0.99), suggesting that the solution of the inversion of the three photon correlations is unique.

Contrary to intuition, smaller molecules are more demanding than larger ones<sup>24</sup>. We therefore challenged our approach by using the 46 residue comprising Crambin protein, which is known to 0.8 Å resolution<sup>49</sup> (Fig. 2e). We estimated an average of 14 coherently scattered photons per Crambin shot, a number which is achieved, e.g., at the XFEL at DESY using an X-ray intensity of  $10^{12}$  photons per pulse at 5 keV and a 1 μm beam diameter. The estimates were calculated with the Condor package by Hantke et al.<sup>10</sup> using a flat-top beam profile. An independent calculation using the SimEx simulation framework for imaging single particles at the European XFEL by Fortmann-Grote et al.<sup>9,11</sup> using a realistic beam profiles yielded similar numbers.

As a conservative test case, and to challenge our method, we generated up to  $3.3 \times 10^9$  synthetic scattering images with only 10 photons on average, totaling up to  $3.3 \times 10^{10}$  recorded photons (Methods). With an expected XFEL repetition rate of up to 27 kHz<sup>50</sup>, and assuming a hit rate of 10%, we expect this data to be collected within a few days (Fig. 3d). As discussed in Supplementary Note 8, the data acquisition time substantially decreases to, e.g., approx. 30 min when on average 100 photons per image are recorded (e.g., by shrinking the beam diameter by a factor of 3 to approx. 300 nm), reducing the total number of required photons by a factor 100 to  $3.3 \times 10^8$  (and reducing the number of images by a factor 1000 to  $3.3 \times 10^6$ ). Even for a lower hit rate such as 1%, 300 min would suffice in this case.

From the synthetic scattering images, we performed 20 independent structure determination runs (Methods and Supplementary Fig. 7). For all runs we used an expansion order  $L = 18$ ,  $K = 26$  shells and a cutoff  $k_{\text{cut}} = 2.15 \text{ \AA}^{-1}$  (Supplementary Note 7 discusses the optimal parameters), thus setting the maximum achievable resolution to 2.9 Å. Fig. 2a–c compares the average intensity obtained from these 20 runs (green) with the reference intensity derived from the known X-ray structure (blue). Overall, the shape of the intensity is recovered very well and only minor deviations in the outer shells, where fewer photons are recorded, are present.

To assess the achievable resolution of the determined Fourier intensities, we calculated 20 real space electron density maps using an iterative phase retrieval algorithm<sup>45</sup>. Figure 2d and e compares the average of the 20 retrieved densities (d, green shaded structure) with the the reference electron density (e, blue shaded structure) which has been calculated from the Fourier density (including phases) with same cutoff  $k_{\text{cut}}$  as (d). The cross-correlation between the two densities is 0.9. The Fourier shell correlation (FSC) between the known reference electron density of Crambin and the retrieved averaged electron density was calculated as a function of the wave number  $k$ <sup>51</sup> (note that we use  $k_{\text{in}} = 2\pi/\lambda$  for all wave number calculations). Similar to single-particle electron microscopy<sup>51</sup>, the wave number  $k_{\text{res}}$  at which  $\text{FSC}(k_{\text{res}}) = 0.5$  was used to estimate the achieved radial resolution  $\Delta r = 2\pi/k_{\text{res}}$ . Here, a near-atomic resolution of 3.3 Å was achieved.

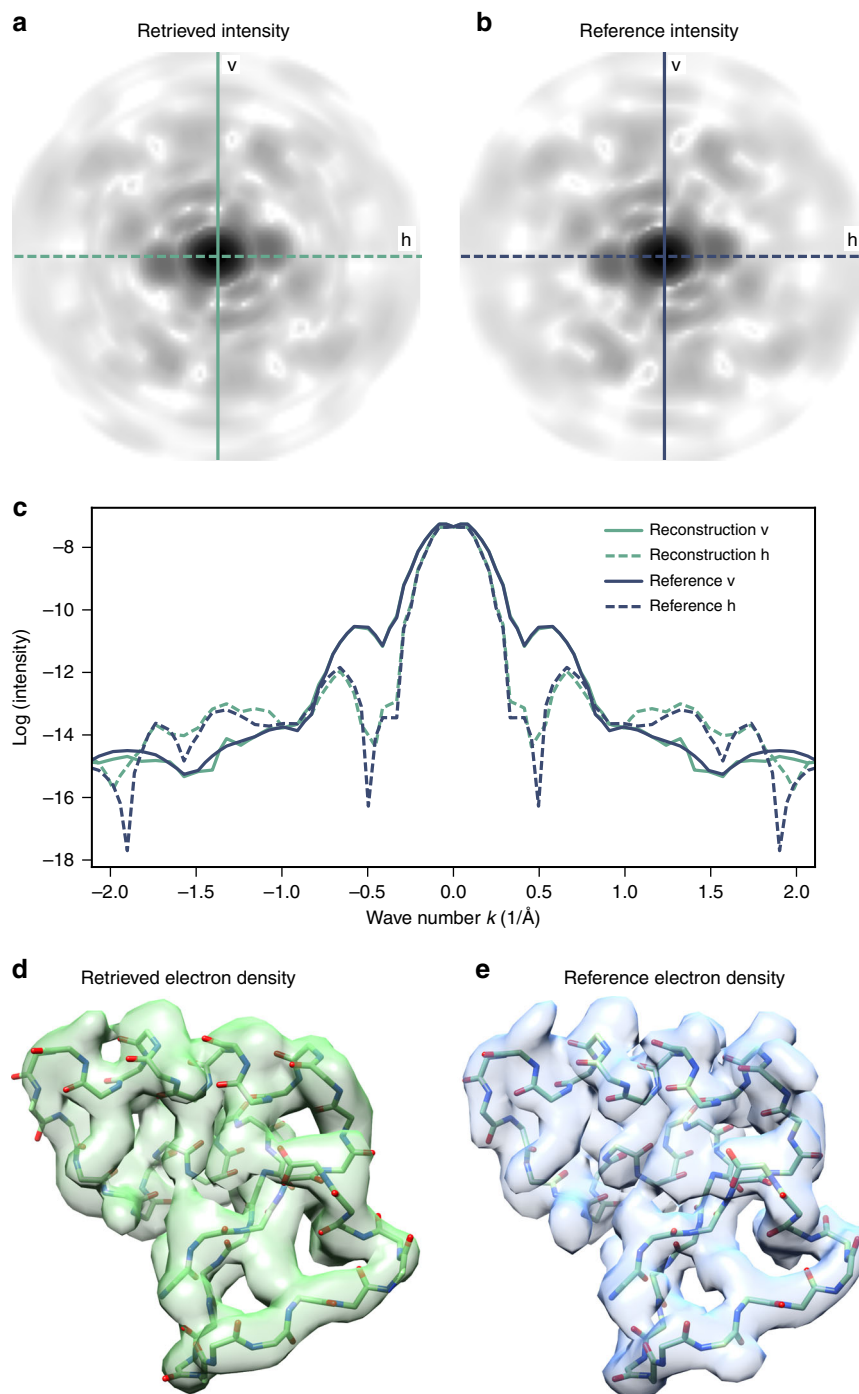
**Resolution as function of number of recorded images.** Next we explored how the achieved resolution depends on the number of observed photons (and triplets, respectively), and hence the

number of recorded images. To this end, electron densities were calculated and averaged as above from  $1.3 \times 10^7$  up to  $3.3 \times 10^{10}$  photons gathered from images with 10 photons on average ( $4.7 \times 10^8$  up to  $1.2 \times 10^{12}$  triplets). Figure 3a depicts the respective FSC curves for different photon counts along with the 0.5 cutoff (vertical dashed line) and the corresponding resolutions (inset).

As mentioned before, for  $3.3 \times 10^{10}$  photons a near-atomic resolution of 3.3 Å was achieved. Decreasing the number of photons by a factor of 10 decreased the resolution only slightly by 0.4 Å to 3.7 Å, which indicates that very likely fewer than  $3.3 \times 10^{10}$  photons suffice to achieve near-atomic resolution. If much fewer photons are recorded, e.g.  $1.3 \times 10^7$  ( $4.4 \times 10^8$  triplets), the resolution decreased markedly to 14 Å. To address the question how much further the resolution can be increased, we mimicked an experiment with infinite number of photons by determining the intensity from the analytically calculated three-photon correlation using Eq. (3) from the Methods section. As can be seen in Fig. 3a (purple line), the resolution only slightly improved by 0.1 Å to about 3.2 Å indicating that at this point either the expansion order  $L$  or insufficient convergence of the Monte Carlo-based structure search became resolution limiting. To distinguish between these two possible causes, we phased the electron density directly from the reference intensity, using the same expansion order  $L = 18$  as in the other experiments. The reference intensity is free from convergence issues of the Monte Carlo structure determination and the resulting electron density only includes the phasing errors introduced by the limited angular resolution of the SH expansion in Fourier space. The FSC curve of the optimal phasing (gray dashed) shows only a minor increase in resolution to 3.1 Å indicating that the Monte Carlo search decreases the resolution by 0.1 Å. The remaining 0.2 Å difference to the optimal resolution of 2.9 Å at the given  $k_{\text{cut}}$  (not shown) is attributed to the finite expansion order  $L$  and the corresponding phasing errors. We have also independently assessed the overall phasing error by calculating the intensity shell correlation (ISC) between the intensities of the phased electron densities  $I_{\text{phased}} = |\mathcal{FT}[\rho_{\text{retrieved}}]|^2$  and the intensities before phasing  $I_{\text{retrieved}}$  (Methods and Supplementary Fig. 8). As discussed in the Methods section, the phasing method does not markedly deteriorate our structures.

Because a large expansion order  $L$  requires a larger number of shells  $K$ , and, therefore, much larger numbers of unknowns (Supplementary Note 7), the question remains at which point overfitting occurs. To quantify this effect for our sets of images, we calculated the achieved resolution as a function of expansion order  $L$  for four different total photon counts  $5.1 \times 10^7$ ,  $2.0 \times 10^8$ ,  $8.2 \times 10^8$ , and  $3.3 \times 10^{10}$  ( $1.8 \times 10^9$ ,  $7.1 \times 10^9$ ,  $2.8 \times 10^{10}$ , and  $1.2 \times 10^{12}$  triplets, respectively) at a fixed number of shells  $K = 26$ . Indeed, as shown in Fig. 3b, for up to  $2.0 \times 10^8$  photons, the obtained three-photon correlation is too noisy to yield an improved resolution when increasing the model detail and for larger  $L$ , the probability  $p$  of the intensity model still increases whereas the resolution decreases again, indicating overfitting. In contrast, for larger photon counts ( $>8.2 \times 10^8$ ), the resolution improves even up to the expansion order  $L = 18$  and no overfitting is expected here. However, due to the large parameter space, convergence of the simulated annealing becomes computationally demanding (Supplementary Notes 4, 5, and 7).

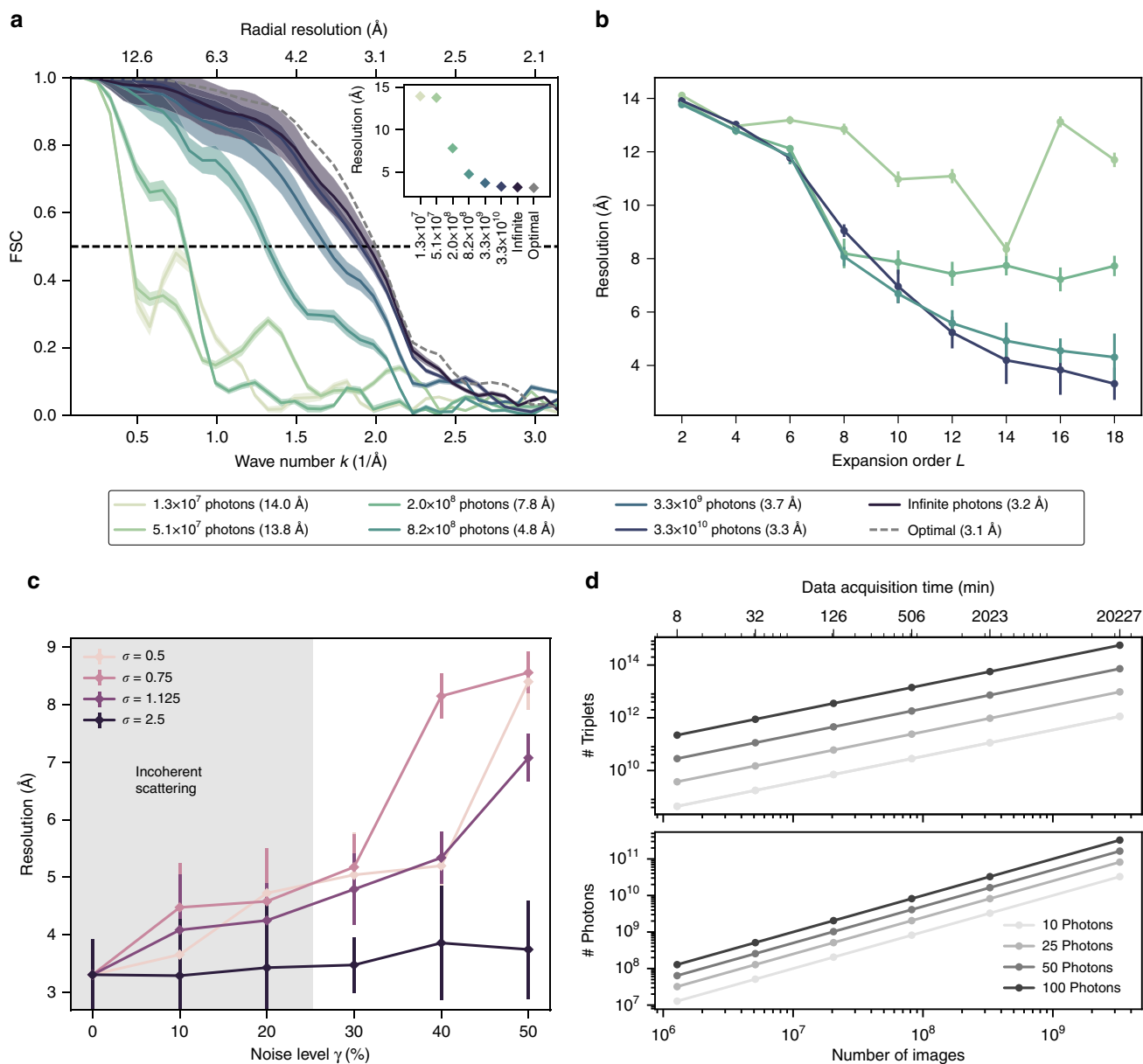
**Robustness to noise.** We finally assessed how robust our approach is in the presence of additional experimental noise due to, e.g., incoherent scattering, background radiation, detector noise, or scattering at the unstructured fraction of water molecules that may adhere to the surface of the macromolecules<sup>1</sup>. Since only very few single-molecule scattering experiments have



**Fig. 2** Structure determination at 3.3 Å resolution. Comparison of the retrieved density (green lines and structure) and the reference density of Crambin (blue lines and structure) in Fourier space and real space. Shown are averages over 20 structure determination runs, each using the same  $3.3 \times 10^9$  images with an average of 10 photons per image yielding  $3.3 \times 10^{10}$  photons. A cutoff in reciprocal space of  $k_{\text{cut}} = 2.15 \text{ \AA}^{-1}$  was used and the intensity was expanded with  $K = 26$  shells using an expansion order of  $L = 18$ . **a, b** Comparison of the the retrieved intensity (**a**) and the reference intensity (**b**) in the  $k_x k_y$ -plane (logarithmic shading). **c** Comparison of two orthogonal linear cuts (vertical, v, and horizontal, h) through the  $k_x k_y$ -planes shown in **a** and **b**. **d, e** Comparison of the retrieved electron density (**d**) and the reference electron density (**e**). The latter was calculated from the known Fourier density using the same cutoff  $k_{\text{cut}} = 2.15 \text{ \AA}^{-1}$  in reciprocal space as in **d**. The resolution of the retrieved density is 3.3 Å, the resolution of the reference density is 2.9 Å, and the cross-correlation between the two densities is 0.9

been carried out so far, quantitative noise models are available only for incoherent scattering, for which a noise level of ca.  $\gamma = 25\%$ <sup>52</sup> is expected. Here we modeled the noise as a Gaussian distribution,  $G(k, \sigma) = \gamma(2\pi\sigma^2)^{-1/2} \exp(-k^2/2\sigma^2)$ . Depending on the width  $\sigma$ , different signal-to-noise ratios are expected in the

low-resolution and high-resolution regions of the image, respectively. For incoherent scattering (indicated as gray background) a width of  $\sigma = 2.5 \text{ \AA}^{-1}$  was assumed<sup>53</sup> (Supplementary Note 9), which corresponds to a relatively uniform noise distribution. Figure 3c (black line) shows a moderate decrease in resolution to



**Fig. 3** Resolution as a function of collected photons and noise. **a** Average Fourier shell correlations (FSC) between the reference electron density of Crambin and the densities retrieved from  $1.3 \times 10^7$  to  $3.3 \times 10^{10}$  photons ( $4.7 \times 10^8$  to  $1.2 \times 10^{12}$  triplets) and infinite photon number. As a reference, the optimal FSC is shown (dashed gray), which was calculated directly from the known intensity using the same expansion parameters. The inset shows the corresponding resolutions estimated from  $\text{FSC}(k_{\text{res}}) = 0.5$ . The opaque error band was calculated from the standard deviation of the 20 individual FSCs before averaging. **b** Achieved average resolution as a function of the expansion order  $L$  using  $5.1 \times 10^7$ ,  $2.0 \times 10^8$ ,  $8.2 \times 10^8$ , and  $3.3 \times 10^{10}$  photons ( $1.8 \times 10^9$ ,  $7.1 \times 10^9$ ,  $2.8 \times 10^{10}$ , and  $1.2 \times 10^{12}$  triplets, respectively). The error bars show the standard error of mean of 20 independent runs. **c** Achieved average resolution for synthetic experiments with  $3.3 \times 10^9$  images including an additional fraction  $\gamma$  of random photons ( $\gamma = 0$ –50%) following a Gaussian distribution with varying width  $\sigma$ . For noise from incoherent scattering (indicated as gray background), we assumed a width  $\sigma = 2.5 \text{\AA}^{-1}$ , but also included narrower widths  $\sigma = [0.5, 0.75, 1.125] \text{\AA}^{-1}$  as a model for other sources of noise. The error bars show the standard error of mean of 20 independent runs. **d** Expected number of photons and triplets as a function of image numbers for 10, 20, 50, and 100 photons per image. Also shown (top axis) is the estimated data acquisition time given a 27 kHz XFEL repetition rate and 10% hit rate

approx.  $3.5 \text{\AA}$  when this noise is included within our synthetic experiments (as described in Supplementary Note 9). Additional noise with a uniform distribution from, e.g. background radiation or detector noise, slightly decreased the resolution to  $3.8 \text{\AA}$  at 50% noise level.

For scattering from disordered water molecules that are attached to the macromolecular surface, a narrower intensity distribution is expected (Supplementary Fig. 4). To also

investigate this effect and the effect of other potential noise sources with non-uniform distribution, in Fig. 3c, we considered noise with widths of  $\sigma = [0.5, 0.75, 1.125] \text{\AA}^{-1}$  and noise levels  $\gamma$  between 10 and 50%, the latter corresponding e.g. to up to 100 disordered water molecules per Crambin molecule. The resolution remained better than  $5 \text{\AA}$  within the 25% noise level but decreases markedly to  $9 \text{\AA}$  with  $\gamma = 50\%$ , in particular for narrow noise widths of  $\sigma = [0.5, 0.75] \text{\AA}^{-1}$ .



**Sample application to experimental data.** To test if our method is also robust against noise in real experimental data, we have determined the structure of the coliphage PR772 virus from the Reddy et al. data set<sup>54</sup> (Supplementary Note 10), albeit at much higher photon counts than our method is targeted for. As described in ref. 54, this image-set has been obtained by filtering the raw images for single molecule hits with diffusion map embedding. Therefore, to mimic low photon counts, we down-sampled the images, which contain over 400,000 photons per image, and generated  $3 \times 10^{12}$  triplets using the same rejection sampling method that we used to generate the Crambin images, and subsequently applied the same reconstruction procedure (Supplementary Fig. 5). A resolution of 11.7 nm was achieved, as evidenced from the FSC between two independently determined structures (Supplementary Fig. 6). This resolution is slightly lower than the 9 nm obtained by Hosseinizadeh et al.<sup>55</sup>, which may be due to the fact that we used fewer photons, implying additional Poisson noise. Also, in contrast to Hosseinizadeh et al., we have not implicitly imposed any icosahedral symmetry in our reconstructions.

## Discussion

The presented method demonstrates de novo structure determination from as few as three photons per XFEL scattering image at near-atomic resolution. Our synthetic scattering experiments with subsequent structure determination have shown that, for the most challenging case of small biomolecules, a resolution better than 3.3 Å should be achievable with available technology at realistic beam times; specifically, as our conservative estimate rests on a beam fluence of  $5.0 \times 10^{11}$  photons per pulse. Assuming a 10% hit rate, our method requires only ca.  $10^{10}$  molecules, which is, compared to nano-crystallography, smaller by a factor of 10 ( $10^5$  nano-crystals with  $10^6 \text{ nm}^3$  volume)<sup>13</sup>.

Even higher resolutions are conceivable for larger molecules due to the larger scattering signal<sup>24</sup>, albeit computational resources may become a limiting factor when determining larger structures at the same resolution of around 3 Å. However, as shown for the structure determination of the much larger coliphage virus in Supplementary Note 10, the computational complexity only depends on the ratio between the size of the molecule and the desired resolution. For a given resolution, the computational complexity scales slightly faster than the molecular weight cubed.

Given that currently available de novo refinement methods require at least 100 photons per image, we consider our finding that only three photons per image suffice quite unexpected. Further, in this extreme Poisson regime, our three-photon correlation approach—in contrast to previous structure determination methods—allows to compensate for fewer photons per image  $P$  by acquiring more images  $I$ . In particular, because two photons per image do not uniquely determine the structure<sup>39</sup>, here we have reached the fundamental limit.

Our analysis also suggests that the method is robust against noise from incoherent scattering, and that removing as much as possible disordered water (or other contaminants) from the molecule in the experiment is crucial. Further, fluctuations of the beam intensity—both in time and due to beam-particle impact parameter fluctuations, which are a limiting factor for image-wise orientation-based methods, should not deteriorate the resolution in our approach, as the correlations are insensitive to such fluctuations. Clearly, further experimental data and improved noise models are required to study the effect of these and other potential noise sources such as background radiation from the evaporated water and detector noise. Structural fluctuations and inhomogeneities of the sample turn more and more into a

limiting factor for all current structure determination methods—particularly for high resolutions. Notably, for mixtures of several structures, single-particle scattering implies that the three-photon correlation on which our method rests is a linear superposition of the three-photon correlations of the individual structures. Hence, our approach should be generalizable in a straight-forward way to refine such mixtures, albeit at the cost of more required images, larger computational effort, and more severe convergence issues. Further, due to the averaging properties of the three-photon correlations, our method should be more robust than methods that rely on an accurate orientation of individual scattering images.

We have tested our approach for a conservative estimate of 10 coherently scattered photons. Should the number of coherently scattered photons per shot be larger, e.g., by reducing the size of the beam focus, our method might even bring single-molecule structure determination within reach of less bright free electron lasers or even table top setups<sup>56</sup>.

Overall, our results suggest that near-atomic structure determination by single-molecule X-ray scattering is within experimental reach. We would like to point out that our correlation-based method can also determine structures from images containing more than one particle which may further reduce the data acquisition time and facilitate sample delivery (Supplementary Note 11 discusses how the two-photon and three-photon correlation of single molecules is calculated from multi-particle correlations). The method is potentially also useful to extract as much as possible information from other types of scattering experiments, in particular when 3D structures are inferred from noisy two-dimensional projections, such as cryo-EM<sup>57,58</sup>, X-ray microscopy, sub-diffractive optical microscopy<sup>59,60</sup>, and from fluctuations in correlated X-ray scattering.

## Methods

**Three-photon correlations expressed in SH.** The three-photon correlation  $t(k_1, k_2, k_3, \alpha, \beta)$  is the orientational average  $\langle \cdot \rangle_\omega$  of the product between three intensities  $I(\mathbf{k})$  that lie on the intersection between the Ewald sphere and the 3D Fourier density (see Supplementary Note 12),

$$t(k_1, k_2, k_3, \alpha, \beta)_{I(\mathbf{k})} = \langle I_\omega(\mathbf{k}_1^*(k_1, 0)) \cdot I_\omega(\mathbf{k}_2^*(k_2, \alpha)) \cdot I_\omega^*(\mathbf{k}_3^*(k_3, \beta)) \rangle_\omega \quad (1)$$

Here, without loss of generality, the three vectors  $\mathbf{k}_1^*$ ,  $\mathbf{k}_2^*$ , and  $\mathbf{k}_3^*$ , are the projection onto the Ewald sphere of the three photons  $\mathbf{k}_1 = (k_1, 0, 0)$ ,  $\mathbf{k}_2 = k_2(\cos \alpha, \sin \alpha, 0)$ , and  $\mathbf{k}_3 = k_3(\cos \beta, \sin \beta, 0)$  in the detector plane. Using a shell-wise SH decomposition of the intensity<sup>47</sup>,

$$I(\mathbf{k}) = \sum_{lm} A_{lm}(k) Y_{lm}(\theta, \varphi), \quad (2)$$

with the coefficients  $A_{lm}(k)$  describing the intensity function on the respective shells, the three-photon correlation is expressed in sums of products of SH coefficients together with known Wigner-3j symbols and SH basis functions  $Y_{lm}(\theta, \varphi)$ ,

$$t(k_1, k_2, k_3, \alpha, \beta)_{\{A_{lm}(k)\}} = \sum_{l_1 l_2 l_3} \sum_{m_1 m_2 m_3} A_{l_1 m_1}(k_1) A_{l_2 m_2}(k_2) A_{l_3 m_3}^*(k_3) \times \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \sum_{m_1' m_2' m_3'} (-1)^{m_3 - m_3'} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1' & m_2' & -m_3' \end{pmatrix} \times Y_{l_1 m_1'}(\theta_1(k_1), 0) Y_{l_2 m_2'}(\theta_2(k_2), \alpha) Y_{l_3 m_3'}^*(\theta_3(k_3), \beta). \quad (3)$$

See Supplementary Note 1 for the full derivation of Eq. (3).

**Synthetic data generation.** We validated our structure determination approach using synthetic scattering experiments on the structure of the 46 residue protein Crambin (PDB descriptor: 3U7T)<sup>49</sup> which has been determined to 0.8 Å resolution. To this end, we approximated the 3D electron density  $\rho(\mathbf{x})$  by a sum of Gaussian functions centered at the atomic positions with height  $\gamma$  and variance  $\sigma$  depending on the atom type. The absolute square of the electron densities' Fourier transformation  $I(\mathbf{k}) = |\mathcal{FT}[\rho(\mathbf{x})]|^2$  was used to generate synthetic scattering images. In each synthetic scattering experiment, the molecule, and thus also  $I(\mathbf{k})$ , was randomly oriented. On average  $P$  photons per image were generated each shot,

according to the distribution given by the randomly oriented Ewald slice of the intensity  $I_w(\mathbf{K})$ .

To generate the distributions numerically, first, a random set of  $N_{\text{pos}}$  positions  $\{\mathbf{K}_i\}$  in the  $k_x k_y$ -plane was generated according to a 2D Gaussian distribution  $G(\mathbf{K})$  with width  $\sigma = 1.05 \text{ \AA}^{-1}$ . Given a random 3D rotation  $\mathbf{U}$  (see Supplementary Note 4 for uniform sampling of  $\text{SO}(3)$ ), rejection sampling method was used to accept or reject each position according to  $\xi < I_w(\mathbf{U} \cdot \mathbf{K}_i) / (M \cdot G(\mathbf{K}_i))$  using uniformly distributed random numbers  $\xi \in [0, 1]$  each. Here, the constant  $M$  was chosen as  $I_{\text{max}} \cdot \max(G(\mathbf{K}))$  such that the ratio  $I_w(\mathbf{U} \cdot \mathbf{K}_i) / (M \cdot G(\mathbf{K}_i))$  is below 1 for all  $\mathbf{K}$ . In accordance with our most conservative estimate discussed in the main text, the number of positions  $N_{\text{pos}}$  was chosen such that on average 10 scattered photons were generated. For assessing the dependency of the resolution on the number of scattered photons, additional image sets with 25, 50, or 100 scattered photons were also generated (Supplementary Note 8).

For technical reasons, we used a SH expansion of the intensity with a high expansion order  $L = 35$  as a sufficiently accurate approximation for  $I(\mathbf{k})$  to generate the images. The accuracy of the intensity model was cross-checked with the intensity calculated on a cubic grid (150 grid size) using the Fast Fourier Transform, resulting in a 0.9999 correlation, thus establishing sufficient accuracy. Altogether, up to  $3.3 \times 10^{10}$  images were generated using a high degree of parallelism.

**Probability of observing a set of triplets.** Because we were not able to derive an analytic inversion for Eq. (3), we chose a probabilistic approach and asked which intensity  $I(\mathbf{k})$  is most likely to have generated the complete set of measured scattering images and triplets, respectively. To this end, we considered the probability  $p$  that a given intensity  $I(\mathbf{k})$ , expressed in SH by  $\{A_{lm}(k)\}$ , generated the set of triplets,  $\{k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i\}_{i=1 \dots T}$ ,

$$p\left(\{k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i\}_{i=1 \dots T} | \{A_{lm}(k)\}\right) = \prod_{i=1}^T \bar{t}(k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i)_{\{A_{lm}(k)\}} \cdot \quad (4)$$

Due to the statistical independence of the triplets, this probability  $p$  is a product over the probabilities  $\bar{t}(k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i)$  of observing the individual triplets  $i$  which is given by the normalized three-photon correlation  $\bar{t}(k_1, k_2, k_3, \alpha, \beta)$ . Here,  $\bar{t}(k_1, k_2, k_3, \alpha, \beta)$  was calculated using Eq. (3) for varying intensity coefficients  $\{A_{lm}(k)\}$  and the coefficients that maximized  $p(\{k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i\})$  were determined using a Monte Carlo scheme.

In contrast to the direct inversion, the probabilistic approach has the benefit of fully accounting for the Poissonian shot noise implied by the limited number of photon triplets that are extracted from the given scattering images. We note that this approach also circumvents the limitation faced by Kam<sup>39</sup>, where only triplets with two photons recorded at the same position could be considered. Because all other triplets had to be discarded, Kam's approach is limited to very high beam intensities, and cannot be applied in the present extreme Poisson regime.

**Reduction of the search space using two-photon correlations.** In our approach, we used the structural information contained within the two-photon correlation to reduce the high-dimensional search space. In analogy to the three-photon correlation, the two-photon correlation is expressed as a sum over products of SH coefficients  $A_{lm}(k)$  weighted with Legendre polynomials  $P_l^{35,39}$ ,

$$c_{k_1, k_2, \alpha} = \sum_l P_l(\cos(\alpha^*)) \sum_m A_{lm}(k_1)(\omega) A_{lm}^*(k_2). \quad (5)$$

Please note that the  $\alpha$  which is seen on the detector is different from the angle  $\alpha^* = \cos^{-1}(\sin(\theta_1)\sin(\theta_2)\cos(\alpha) + \cos(\theta_1)\cos(\theta_2))$  between the two points in 3D intensity space due to the Ewald curvature ( $\theta = \cos^{-1}(kl/4\pi)$ ).

The inversion of Eq. (5) yields coefficient vectors  $\mathbf{A}_l^0(k) = (A_{l-m}^0, \dots, A_{lm}^0)$  for all  $l \leq L \leq K_{\text{max}}/2$  and  $-l < m < l$ , as first demonstrated by Kam<sup>39</sup>. However, all rotations in the  $2l + 1$ -dimensional coefficient eigenspaces of  $\mathbf{A}_l^0(k)$  by  $\mathbf{U}_l$  are also solutions,

$$\mathbf{A}_l(k) = \mathbf{U}_l \mathbf{A}_l^0(k). \quad (6)$$

The result implies that the inversion only gives a degenerate solution for the coefficients and the intensity cannot be determined solely from two photons. Here, we used Eq. (6) to search for the optimal rotations  $\mathbf{U}_l$  instead of optimal coefficients  $A_{lm}^{\text{all}}(k)$ , which reduced the size of the search space from  $(\frac{1}{2}L^2 + \frac{3}{2}L + 1) \cdot K$  to  $\frac{1}{3}(L^3 + \frac{15}{4}L^2 + \frac{7}{2}L)$  unknowns (e.g., reducing the number of unknowns from 4940 coefficients to 2370 rotation angles for  $L = 18$  and  $K = 26$ ). See Supplementary Note 6 for more details.

**Monte Carlo simulated annealing.** The probability  $p$  from Eq. (4) was maximized by a Monte Carlo/simulated annealing approach on the energy function:

$$E(\{k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i\} | \{A_{lm}(k)\}) = -\log p(\{k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i\} | \{A_{lm}(k)\}) \\ = -\sum_i \log \bar{t}(k_1^i, k_2^i, k_3^i, \alpha^i, \beta^i)_{\{A_{lm}(k)\}}, \quad (7)$$

in the space of all rotations  $\mathbf{U}_l$  given by the inversion of the two-photon correlation. Each Monte Carlo run was initialized with a random set of rotations  $\{\mathbf{U}_l\}$  and the set of unaligned coefficients  $\{A_l^0\}$ . In each Monte Carlo step  $j$ , all rotations  $\mathbf{U}_l^j$  were varied by small random rotations  $\Delta_l(\beta_l)$  such that the updated rotations for each  $l$  ( $l \leq L$ ) read  $\mathbf{U}_l^{j+1} = \Delta_l(\beta_l) \cdot \mathbf{U}_l^j$  using stepsizes  $\beta_l$ . In order to escape local minima, a simulated annealing was performed using an exponentially decaying temperature protocol,  $T(j) = T_{\text{init}} \exp(j/\tau)$ . Steps with an increased energy were also accepted according to the Boltzmann factor  $\exp(-\Delta E/T)$ . We further used adaptive stepsizes such that all  $\beta_l$  were increased or decreased by a factor  $\mu$  when accepting or rejecting the proposed steps, respectively. Convergence was improved by using a hierarchical approach in which the intensity was first determined with low angular resolution and further increased to high resolution. To this end, the variations of low-resolution features were frozen out faster than the variations of high-resolution features. See Supplementary Note 4 on how to generate random rotations in  $\text{SO}(n)$  and how the parameters of the Monte Carlo search were determined.

**Calculation of real space electron densities and resolutions.** Supplementary Fig. 7 summarizes the calculation of the electron densities as carried out in this work. All intensities were obtained up to an arbitrary Euler rotation ( $\theta, \phi, \psi$ ) and were therefore rotationally fit to the known reference intensity for subsequent comparison. The phases of the aligned intensities were calculated using the relaxed averaged alternating reflections (RAAR) method by Luke<sup>45</sup>. The resolution of the electron densities was characterized by the FSC,

$$\text{FSC}(k) = \frac{\sum_{k_i \in k} F_1(k_i) \cdot F_2(k_i)^*}{\sqrt{\sum_{k_i \in k} |F_1(k_i)|^2 \cdot \sum_{k_i \in k} |F_2(k_i)|^2}}. \quad (8)$$

In analogy to cryo-EM<sup>51</sup>, the resolution is defined as the wave number  $k_{\text{res}}$  at which  $\text{FSC}(k) = 0.5$ , yielding a radial resolution  $\Delta r = 2\pi/k_{\text{res}}$ .

Starting from an individual set of doublet and triplet histograms (Supplementary Fig. 1), 20 independent intensity determination runs were carried out to assess and improve convergence of the Monte Carlo simulated annealing runs. To reduce the phasing error, the phase retrieval of one intensity was carried out eight times and the resulting eight electron densities were averaged. The final electron density, for which the resolution is given, is the average of those 20 individual densities and the resolution error was estimated from the standard deviation of the resolution of the 20 individual electron densities. We chose to average in real space instead of Fourier space before phasing because we found that this sequence yielded more accurate electron densities.

**Evaluation of phasing errors.** To assess the phasing error, we compared the intensities of the phased electron densities  $I_{\text{phased}} = |\mathcal{FT}[\rho_{\text{retrieved}}]|^2$  with the intensities  $I_{\text{retrieved}}$  before phasing. To this end, the ISC was calculated as:

$$\text{ISC}(k) = \frac{\sum_{k_i \in k} (I_{\text{res}}(k_i) - \bar{I}_{\text{res}}(k_i)) (I_{\text{ref}}(k_i) - \bar{I}_{\text{ref}}(k_i))}{\sqrt{\sum_{k_i \in k} (I_{\text{res}}(k_i) - \bar{I}_{\text{res}}(k_i))^2} \sqrt{\sum_{k_i \in k} (I_{\text{ref}}(k_i) - \bar{I}_{\text{ref}}(k_i))^2}}. \quad (9)$$

In analogy to the FSC, we considered  $\text{ISC}(k) = 0.5$  as a resolution measure. As can be seen in Supplementary Fig. 8, the phasing shifted this crossover from approx. 2.8 to 3.1 Å, but does not distort the shapes and relative heights of the ISC curves. Assuming that the phasing error can be estimated from the shift of this crossover, for our high-resolution density result with 3.3 Å resolution (retrieved from  $3.3 \times 10^{10}$  photons), a decrease in resolution of ca. 0.3 Å is expected to be due to phasing.

**Data availability.** All relevant data are available from the authors.

**Code availability.** The code is available at <https://github.com/h4rm/ThreePhotons> and the data analysis was done using IJulia notebooks which are available at <https://github.com/h4rm/ThreePhotonsNotebook>. For more information, please visit <http://www.mpiibpc.mpg.de/grubmueller/threephotons>.

Received: 30 June 2017 Accepted: 1 May 2018

Published online: 18 June 2018

## References

- Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406**, 752–757 (2000).
- Hajdu, J. Single-molecule X-ray diffraction. *Curr. Opin. Struct. Biol.* **10**, 569–573 (2000).
- Huldt, G., Szoke, A. & Hajdu, J. Diffraction imaging of single particles and biomolecules. *J. Struct. Biol.* **144**, 219–227 (2003).
- Gaffney, K. J. & Chapman, H. N. Imaging atomic structure and dynamics with ultrafast X-ray scattering. *Science* **316**, 1444–1449 (2007).
- Miao, J., Ishikawa, T., Robinson, I. K. & Murnane, M. M. Beyond crystallography: diffractive imaging using coherent x-ray light sources. *Science* **348**, 530–535 (2015).
- Seibert, M. M. et al. Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature* **470**, 78–81 (2011).
- Ekeberg, T. et al. Three-dimensional reconstruction of the giant mimi-virus particle with an X-ray free-electron laser. *Phys. Rev. Lett.* **114**, 98102 (2015).
- Hosseinzadeh, A. et al. High-resolution structure of viruses from random diffraction snapshots. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130326–20130326 (2014).
- Yoon, C. H. et al. A comprehensive simulation framework for imaging single particles and biomolecules at the European X-ray free-electron laser. *Sci. Rep.* **6**, 24791 (2016).
- Hantke, M. F., Ekeberg, T. & Maia, F. R. Condor: a simulation tool for flash X-ray imaging. *J. Appl. Crystallogr.* **49**, 1356–1362 (2016).
- Fortmann-Grote, C. et al. in *Proceedings of SPIE—The International Society for Optical Engineering* (eds Tschentscher, T. & Patthey, L.) Vol. 10237, 102370S (International Society for Optics and Photonics, Bellingham, 2017).
- Chapman, H. N. et al. Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nat. Phys.* **2**, 839–843 (2006).
- Chapman, H. N. et al. Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011).
- Boutet, S. et al. High-resolution protein structure determination by serial femtosecond crystallography. *Science* **337**, 362–364 (2012).
- Fromme, P. & Spence, J. C. H. Femtosecond nanocrystallography using X-ray lasers for membrane protein structure determination. *Curr. Opin. Struct. Biol.* **21**, 509–516 (2011).
- Kirian, R. A. et al. Femtosecond protein nanocrystallography—data analysis methods. *Opt. Express* **18**, 5713–5723 (2010).
- Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCr* **2**, 246–255 (2015).
- Roedig, P. et al. High-speed fixed-target serial virus crystallography. *Nat. Methods* **14**, 805–810 (2017).
- Philipp, H. T., Ayer, K., Tate, M. W., Elser, V. & Gruner, S. M. Solving structure with sparse, randomly-oriented x-ray data. *Opt. Express* **20**, 13129 (2012).
- Philipp, H. T., Ayer, K., Tate, M. W., Elser, V. & Gruner, S. M. Recovering structure from many low-information 2-D images of randomly-oriented samples. *J. Phys. Conf. Ser.* **425**, 192016 (2013).
- Ayyer, K., Philipp, H. T., Tate, M. W., Elser, V. & Gruner, S. M. Real-space X-ray tomographic reconstruction of randomly oriented objects with sparse data frames. *Opt. Express* **22**, 2403 (2014).
- Shneerson, V. L., Ourmazd, A. & Saldin, D. K. Crystallography without crystals. I. The common-line method for assembling a three-dimensional diffraction volume from single-particle scattering. *Acta Crystallogr. A Found. Crystallogr.* **64**, 303–315 (2008).
- Loh, N. T. D. & Elser, V. Reconstruction algorithm for single-particle diffraction imaging experiments. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **80**, 26705 (2009).
- Walczak, M. & Grubmüller, H. Bayesian orientation estimate and structure information from sparse single-molecule x-ray diffraction images. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **90**, 22714 (2014).
- Flamant, J., Bihan, N. L., Martin, A. V. & Manton, J. H. Expansion-maximization-compression algorithm with spherical harmonics for single particle imaging with X-ray lasers. *Phys. Rev. E* **93**, 053302 (2016).
- Kassemeyer, S. et al. Optimal mapping of x-ray laser diffraction patterns into three dimensions using routing algorithms. *Phys. Rev. E* **88**, 042710 (2013).
- Elser, V. Three-dimensional structure from intensity correlations. *New J. Phys.* **13**, 123014 (2011).
- Donatelli, J. J., Sethian, J. A. & Zwart, P. H. Reconstruction from limited single-particle diffraction data via simultaneous determination of state, orientation, intensity, and phase. *Proc. Natl Acad. Sci. USA* **114**, 7222–7227 (2017).
- Tegze, M. & Bortel, G. Atomic structure of a single large biomolecule from diffraction patterns of random orientations. *J. Struct. Biol.* **179**, 41–45 (2012).
- Fung, R., Shneerson, V., Saldin, D. K. & Ourmazd, A. Structure from fleeting illumination of faint spinning objects in flight. *Nat. Phys.* **5**, 64–67 (2008).
- Moths, B. & Ourmazd, A. Bayesian algorithms for recovering structure from single-particle diffraction snapshots of unknown orientation: a comparison. *Acta Crystallogr. A Found. Crystallogr.* **67**, 481–486 (2011).
- Schwander, P., Giannakis, D., Yoon, C. H. & Ourmazd, A. The symmetries of image formation by scattering. II. *Appl. Opt. Express* **20**, 12827–12849 (2012).
- Giannakis, D., Schwander, P. & Ourmazd, A. The symmetries of image formation by scattering. I. Theoretical framework. *Opt. Express* **20**, 12799–12826 (2012).
- Enderlein, J. Maximum-likelihood criterion and single-molecule detection. *Appl. Opt.* **34**, 514 (1995).
- Saldin, D. K., Shneerson, V. L., Fung, R. & Ourmazd, A. Structure of isolated biomolecules obtained from ultrashort x-ray pulses: exploiting the symmetry of random orientations. *J. Phys. Condens. Matter.* **21**, 134014 (2009).
- Saldin, D. K. et al. Beyond small-angle x-ray scattering: exploiting angular correlations. *Phys. Rev. B* **81**, 1–6 (2010).
- Saldin, D. K. et al. Structure of a single particle from scattering by many particles randomly oriented about an axis: a new route to structure determination? *New J. Phys.* **12**, 35014 (2010).
- Saldin, D. K. et al. New light on disordered ensembles: ab initio structure determination of one particle from scattering fluctuations of many copies. *Phys. Rev. Lett.* **106**, 115501 (2011).
- Kam, Z. The reconstruction of structure from electron micrographs of randomly oriented particles. *J. Theor. Biol.* **82**, 15–39 (1980).
- Starodub, D. et al. Single-particle structure determination by correlations of snapshot X-ray diffraction patterns. *Nat. Commun.* **3**, 1276 (2012).
- Saldin, D., Poon, H.-C., Schwander, P., Uddin, M. & Schmidt, M. Reconstructing an Icosahedral virus from single-particle diffraction experiments. *Opt. Express* **19**, 18 (2011).
- Poon, H. C. & Saldin, D. K. Use of triple correlations for the sign determinations of expansion coefficients of symmetric approximations to the diffraction volumes of regular viruses. *Struct. Dyn.* **2**, 041716 (2015).
- Fienup, J. R. Phase retrieval algorithms: a comparison. *Appl. Opt.* **21**, 2758–2769 (1982).
- Fienup, J. & Wackerman, C. C. Phase retrieval stagnation problems and solutions. *J. Opt. Soc. Am. A* **3**, 1897–1907 (1986).
- Luke, D. R. Relaxed averaged alternating reflections for diffraction imaging. *Inverse Probl.* **37**, 13 (2004).
- Shechtman, Y. et al. Phase retrieval with application to optical imaging: a contemporary overview. *IEEE Signal Process. Mag.* **32**, 87–109 (2015).
- Baddour, N. Operational and convolution properties of three-dimensional Fourier transforms in spherical polar coordinates. *J. Opt. Soc. Am. A* **27**, 2144 (2010).
- Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. Optimization by simulated annealing. *Science* **220**, 671–680 (1983).
- Chen, J. C. H. et al. Room-temperature ultrahigh-resolution time-of-flight neutron and X-ray diffraction studies of H/D-exchanged crambin. *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **68**, 119–123 (2012).
- Barty, A., Küpper, J. & Chapman, H. N. Molecular imaging using X-ray free-electron lasers. *Annu. Rev. Phys. Chem.* **64**, 415–435 (2013).
- Van Heel, M. & Schatz, M. Fourier shell correlation threshold criteria. *J. Struct. Biol.* **151**, 250–262 (2005).
- Hubbell, J. H. et al. Atomic form factors, incoherent scattering functions, and photon scattering cross sections. *J. Phys. Chem. Ref. Data* **4**, 471–538 (1975).
- Klein, O. & Nishina, T. Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac. *Z. für Phys.* **52**, 853–868 (1929).
- Reddy, H. K. N. Data descriptor: coherent soft X-ray diffraction imaging of coliphage PR772 at the Linac coherent light source Background & Summary. *Sci. Data* **4**, 170079 (2017).
- Hosseinzadeh, A. et al. Conformational landscape of a virus by single-particle X-ray scattering. *Nat. Methods* **14**, 877–881 (2017).
- Grüner, F. et al. Design considerations for table-top, laser-based VUV and X-ray free electron lasers. *Appl. Phys. B Lasers Opt.* **86**, 431–435 (2007).
- Ischenko, A. A., Weber, P. M. & Dwayne Miller, R. J. Capturing chemistry in action with electrons: realization of atomically resolved reaction dynamics. *Chem. Rev.* **117**, 11066–11124 (2017).
- Miller, R. J. D. Ultrafast imaging of photochemical dynamics: roadmap to a new conceptual basis for chemistry. *Faraday Discuss.* **194**, 777–828 (2016).
- Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780 (1994).
- Balzarotti, F. et al. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* **355**, 606–612 (2016).

## Acknowledgements

Financial support from the Deutsche Forschungsgemeinschaft (DFG) Grant No. SFB 755.B4 and helpful discussions with Russel Luke are gratefully acknowledged.



**Author contributions**

B.v.A., M.M., H.G. conceived research, B.v.A. carried out research, B.v.A., M.M., H.G. wrote paper.

**Additional information**

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-04830-4>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018