Diffraction data from aerosolized Coliphage PR772 virus particles imaged with the Linac Coherent Light Source

Haoyuan Li et al.*

Single Particle Imaging (SPI) with intense coherent X-ray pulses from X-ray free-electron lasers (XFELs) has the potential to produce molecular structures without the need for crystallization or freezing. Here we present a dataset of 285,944 diffraction patterns from aerosolized Coliphage PR772 virus particles injected into the femtosecond X-ray pulses of the Linac Coherent Light Source (LCLS). Additional exposures with background information are also deposited. The diffraction data were collected at the Atomic, Molecular and Optical Science Instrument (AMO) of the LCLS in 4 experimental beam times during a period of four years. The photon energy was either 1.2 or 1.7 keV and the pulse energy was between 2 and 4 mJ in a focal spot of about 1.3 μm x 1.7 μm full width at half maximum (FWHM). The X-ray laser pulses captured the particles in random orientations. The data offer insight into aerosolised virus particles in the gas phase, contain information relevant to improving experimental parameters, and provide a basis for developing algorithms for image analysis and reconstruction.

Background & Summary

Since the establishment of the single particle initiative1, several experiments have been conducted at the Linac Coherent Light Source (LCLS) to identify and resolve experimental challenges in high-resolution Single Particle Imaging (SPI) experiments2,3. Coliphage PR772 viruses were utilized extensively in these experiments as the standard control sample due to its high structural homogeneity, uniformity, stability, suitable particle concentration in solution, and the ability to be aerosolized for injection into the LCLS beam using aerosol injector technology2,4,5.

An initial dataset from experiments using Coliphage PR772 performed at the LCLS in 2015 was published in 20176 to assist in the development of analysis methods. Since that experiment, several additional experiments have been performed to push the method to higher resolutions and carry out testing of different aerosolization and sample delivery methods. Coliphage PR772 was also used as a standard reference sample in those subsequent experiments. This provides an opportunity to investigate the influence of experiment conditions on data quality and to check the reproducibility of SPI experiments in addition to obtaining higher resolution data. The purpose of this paper is to describe data from these additional experiments7.

Four experiment runs with PR772 have been performed in the years from 2015 to 2018 (amo87215, amo06516, amo11416, amox34117). This paper summarizes the data collected in those experiments, the experimental conditions, and classification results for single-hit diffraction patterns. We provide appropriate metadata for interpreting the images including: photon energy, X-ray pulse energy and length, position of each pixel relative to the interaction region, bad-pixel mask, the run number and index for all classified hits and the run number and index for all single hits. Analysis of diffraction patterns from real experiments with a variety of experimental configurations can potentially facilitate the development of a robust data processing pipeline for the processing of experimental single particle diffraction data.

*A full list of authors and their affiliations appears at the end of the paper.
Wavefront sensor measurements taken in 2017 show the focused X-ray beam to be nearly Gaussian in shape with coated Kirkpatrick-Baez (KB) mirrors capable of focusing the FEL beam to a nominal 1.5 μm diameter focal spot. A schematic of these experiments is shown in Fig. 3. The instrument uses a pair of Boron Carbide GDVN nozzles with an asymmetric “syringe tip” design to produce a much smaller inner diameter of order 15 to 20 μm. The Glass GDVN Nozzles were melted to create a much smaller inner diameter of order 15 to 20 μm. The virus particles were then passed through a differentially pumped skimmer that was used to reduce the background scattering from the carrier gas and to protect the detector from thermal drift and high voltage arcing. The samples were then focused into the sample chamber’s interaction point of the X-ray instrument using an aerodynamic lens stack injector. All four experiments were conducted at the LAMP endstation of the AMO instrument at the LCLS. A schematic of these experiments is shown in Fig. 3. The instrument uses a pair of Boron Carbide coated Kirkpatrick-Baez (KB) mirrors capable of focusing the FEL beam to a nominal 1.5 μm diameter focal spot, Wavefront sensor measurements taken in 2017 show the focused X-ray beam to be nearly Gaussian in shape with a FWHM of 1.3 μm × 1.7 μm (vertical x horizontal). Shot by shot X-ray pulse energies were measured with gas monitors located upstream of the AMO optics. Measured pulse energies varied between 2 and 4 mJ per pulse and are included in the metadata for each diffraction image. It is noted that the X-ray optical transport system of the AMO instrument is not perfect and has been measured to be ~40% efficient. Background scatter, from the upstream optics and residual gas in the chamber, was reduced using a beveled silicon nitride 4-jaw slit followed by a FWHM of 1.3 μm.
by two motorized 1 mm × 1 mm opening silicon nitride apertures used to reduce scatter from the 4-jaw slit. The 4-jaw slit was located ~20 cm upstream of the focus and the two apertures were located ~15 cm and ~7 cm respectively upstream of the focus. Additionally, adjustable rolled B₄C slits were used 2.0 m upstream of the KB mirrors to define the entrance aperture of the focusing system (not shown in Fig. 3).

Initial alignment of the aerodynamic lens injector to the focal spot position was performed using the beam-line alignment laser and a retractable alignment pin coated in a powdered phosphor to directly align the center of the injector with the X-ray focus. The injector was positioned 3 mm above the X-ray focus. Lateral scans of the injector were conducted for each experiment to optimize hit rates. The focus of the particle stream was found to be approximately 100 μm (full width at half maximum) with variation in focal spot size depending on inlet and chamber pressures.

The samples exiting the aerodynamic lens injector and entering the X-ray interaction region of the instrument are in random orientations and also enter the interaction point at random time intervals, as the aerodynamic lens does not align the particles in any particular orientation. As the sample delivery focus was far greater than that of the X-ray pulses in width (as illustrated in the inset of Fig. 3) the majority of X-ray pulses miss the sample and do not interact with any particles. The LCLS provides 120 equally spaced X-ray pulses per second and typically 1% of these will intersect with a sample, depending on the sample concentration, GDVN and skimmer operating conditions.

Diffracted X-rays are collected, downstream of the interaction point on two 512 × 1024 pixel pnCCD panels. The detector consists of two panels which are movable jointly along the X-ray beam axis, Z, and the two panels can also be moved independently vertically, Y, with respect to the horizontal gap between the two detector panels.
panels. When no particle is present in the X-ray focus the measured intensity corresponds to instrument background due to scatter from residual gas, slits, and so on; however, when a sample particle interacts with the XFEL beam a coherent diffraction pattern is additionally measured on the detector. The position of both panels and the camera length of the detector from the interaction region was determined using the known diffraction of Silver Behenate prior to each experiment. An example of such a calibration is shown in Fig. 4.

An X-ray photon energy of 1.7 keV (0.73 nm) was used for most of the experiments reported here, except for during runs 38–58 of the amo87215 experiment where an X-ray photon energy of 1.2 keV (1.03 nm) was used (other runs in amo87215 were at 1.7 keV).

Both the detector distance and the detector gap size have been optimized for the measurement of high resolution data throughout the experiments. The detector distance and the detector edge resolution for each experiment can be found in Table 1. Notice that, in amo11416, for runs 55 and 56, the gap size is different from the previous runs to reach a higher edge resolution of 2.8 nm.

Data processing. The pnCCD detector is an integrating detector that reads out the deposited charge incident on each pixel in analog-to-digital units (ADUs). Photon counting detectors cannot be used for this type of experiment due to the arrival of multiple photons in an individual pixel within the space of a few femtoseconds. However, integrating detectors (such as the pnCCD) can still achieve single photon sensitivity under certain conditions. A series of corrections and calibrations are required in order to convert the data from ADUs to photon counts per pixel. In this report, we use psana, an LCLS software framework, to retrieve the data, obtain the detector pixel positions, mask for bad pixels and apply various corrections to convert the ADUs into photon counts. Corrections applied to the pnCCD data include (in order) pedestal subtraction, common-mode correction and gain correction followed by conversion to photon counts. As each photon strikes a given pixel, an electron cloud is generated in the substrate of the detector panel, with the number of electrons being proportional to the number of incident photons, the photon energy and the degree of charge sharing between neighbouring pixels. This current is then integrated to form the ADU count for that pixel. In Fig. 5 we show a histogram of the measured ADU counts from silicon fluorescence ($K\alpha = 1.74$ keV) after pedestal and common-mode correction (i.e. subtraction of average CCD dark current and voltage offsets). The modal ADU values corresponding to zero, one and two incident photons are situated at the peaks of the three Gaussian profiles (black dashed lines) with values of 0, 134 and 268 ADUs for a gain setting of 4, respectively. The spread in the ADU values about these peaks are due to the stochastic nature of the pedestal, gain and charge sharing processes. Thus, simple division of ADUs by the mean ADUs-per-photon yields poor photon conversion. We used a psana built-in function (detector.photons) to convert the ADUs into photon counts for each pixel which accounts for charge sharing and incident photon energy.

Hit rates in these experiments were typically ~1% as previously mentioned. Hits are defined as frames containing discernible diffraction from the sample, which are identified as frames with significantly elevated diffraction intensity. This process is accomplished using the program psocake. First, one designs a mask for each run defining bad regions, usually blocking the zeroth order diffraction fringe, pixels too far away from the diffraction center and other “bad” regions in the detector where there is significant instrument scattering or there are readout issues with specific pixels. The total photon numbers in the remaining region is calculated, and then patterns are sorted according to the total photon counts per frame as shown in Figs. 6, 7, 8, and 9. The threshold at which to stop accepting frames is then determined by inspection of individual data frames from high intensity to lower intensity. Below a certain number of photons in the region of interest, the diffraction fringes are no longer visible.
When diffraction fringes are no longer visible by eye, the image is considered to contain not enough data to be classified as a hit and is classified as empty or blank for preliminary processing. Frames with higher total photon counts than that value are considered hits and retained for subsequent analysis.

Not all the patterns retained above are valid diffraction patterns from a single PR772 virus particle. These patterns are further classified manually to select the single-hit patterns, from those consisting of clusters of PR772 virus particles. This clustering occurs when two or more PR772 virus particles are contained in a single aerosolization droplet causing the viruses to stick together. A trade off between higher isolated particle hit rates and a higher number of clusters is observed as increasing hit rates to higher levels usually requires changing sample concentration or GDVN conditions in the same direction that also increases the probability of multiple particles existing in an aerosolization droplet. It is acknowledged that this analysis process is influenced by human bias, however it is relatively straightforward to distinguish good single hit patterns from the others for PR772 particles when the intensity is high enough, because the PR772’s shell possess pseudo-icosahedral symmetry this lends itself to a distinct diffraction pattern at low diffraction angles.
Data Records

We provide access to the experiment data, both in the native file format used by the LCLS and in the CXI file format. The LCLS stores beamtime data in the XTC format, which is optimised for sequential reading and writing. The XTC files contain the unprocessed “raw” detector data and metadata for every event in the selected experiment runs. Instructions for extracting data from XTC formatted files can be found at the LCLS data analysis website: https://confluence.slac.stanford.edu/display/PSDM/LCLS+Data+Analysis. The CXI format is based on the popular HDF5 format, which is a self-describing container for multidimensional data structures. The CXI format can be understood as simply a set of conventions for storing scientific data relating to coherent x-ray imaging in a HDF5 file. The CXI files contain the processed and selected diffraction patterns following version 1.6 of the standard, as shown in Fig. 10. There is one cxi file per experiment. The data corresponding to the nth experiment run is stored in a separate “entry” /entry_n, for example, the data for run 90 of the AMO06516 experiment is stored in /entry_1 of the file amo06516.cxi, since this is the first run that has been selected from that experiment.
Fig. 9  Histograms and typical single hits for experiment AMOX34117. (a) The histogram of the total photon counts of the single hit patterns in this experiment. (b–e) Each is a random pattern selected from the 1st, 3rd, 5th and 7th column in the histogram. The boundary is colored with the same color as that of the corresponding column. Single hit patterns are rendered with matplotlib.pyplot.imshow functions with color map “jet” and $v_{\text{max}} = 4$. Before rendering, the photon count patterns are first down-sampled 4-by-4 times.

Fig. 10  The structure of the CXI file containing the photon converted and selected diffraction data.
### Table 1. Summary of experiment conditions and dataset statistics.

| Exp Name | AMO87215 | AMO06516 | AMO11416 | AMOX34117 |
|----------|----------|----------|----------|-----------|
| Run Range | 49–58 59–78 90–143 | 38–50 55,56 | 130–236 |
| Photon Energy (eV) | 1210.6 1536.0 1656.4 | 1653.1 1701.6 |
| Detector Distance (cm) | 360 283 | 219 130 |
| Edge Resolution (nm) | 9.7 7.6 | 5.5 4.2 2.8 1.8 |
| Single Hit Number | 24 2450 9033 | 211 2450 1393 |
| Total Hit Number | 216 11230 84596 | 4546 11230 197667 |
| Single/Total Ratio | 11.1% 21.8% 10.7% | 4.6% 21.8% 0.7% |
| Approx. Run Time (hr) | 1.25 6.45 10.34 | 4.05 0.87 22.83 |

### Table 2. Summary of experiment conditions and dataset statistics.

| Exp Name | AMO87215 | AMO06516 | AMO11416 | AMOX34117 |
|----------|----------|----------|----------|-----------|
| Run | 49 54 55 56 57 58 59 60 61 62 63 64 65 |
| Single | 0 8 4 5 7 14 139 126 320 378 324 |
| Total | 5 1 36 25 112 37 74 74 76 77 78 |
| Single | 160 33 96 5 6 3 1 171 163 172 58 203 |
| Total | 487 115 935 208 564 206 78 365 239 1182 700 1439 1678 1186 |
| Run | 66 67 68 69 71 72 73 74 75 76 77 78 |
| Single | 159 33 96 5 6 3 1 171 163 172 58 203 |
| Total | 487 115 935 208 564 206 78 365 239 1182 700 1439 1678 1186 |
| Run | 90 91 93 94 95 96 97 99 100 101 102 104 105 |
| Single | 106 101 12 60 22 475 128 70 189 200 29 67 300 |
| Total | 1122 984 217 902 379 6850 1938 1009 1396 2723 289 900 3238 |
| Run | 106 107 108 109 111 113 114 116 117 118 119 121 122 |
| Single | 74 481 484 409 461 3 376 487 438 406 375 432 410 |
| Total | 708 4681 4711 4155 4088 26 3028 3759 3592 3404 3022 2945 3364 |
| Run | 123 124 126 127 128 129 132 133 137 138 143 |
| Single | 355 385 350 369 13 395 201 0 6 9 |
| Total | 3373 2705 2511 4009 3786 287 376 438 406 375 432 410 |
| Run | 130 131 132 133 134 135 136 141 147 148 149 150 151 |
| Single | 18 19 24 4 6 3 3 1 0 0 0 0 0 |
| Total | 379 507 521 108 280 1598 126 111 460 494 165 1570 1044 |
| Run | 152 153 154 155 156 157 158 159 160 163 164 165 168 |
| Single | 0 0 0 0 0 0 0 0 0 0 0 0 0 |
| Total | 194 351 376 750 1437 61 231 94 114 1.6e4 2.5e4 2.5e4 1052 |
| Run | 169 170 172 173 174 175 176 177 178 179 180 181 182 |
| Single | 0 0 0 0 0 0 10 6 3 3 0 1 16 |
| Total | 343 104 131 86 223 698 4749 3131 1032 2338 1196 1850 6191 |
| Run | 183 184 185 186 187 188 189 190 191 192 193 194 200 |
| Single | 70 5 61 65 119 3 4 13 4 6 1 2 0 |
| Total | 3532 980 2070 976 7466 1.5e4 1.0e4 8007 4350 4252 2601 1523 643 |
| Run | 201 202 203 204 205 206 209 210 211 212 213 214 215 |
| Single | 0 0 0 0 0 0 2 2 5 17 15 0 0 |
| Total | 1209 1484 714 6300 5841 26 354 79 402 322 19 423 989 |
| Run | 216 217 218 219 220 221 222 225 226 227 228 229 230 |
| Single | 0 0 0 0 0 0 0 7 86 59 79 96 164 |
| Total | 273 266 152 374 128 134 10 315 1078 1170 1043 2831 2909 |
| Run | 231 232 233 234 235 236 |
| Single | 63 33 42 17 139 86 |
| Total | 1267 284 574 396 2388 2141 |
The pnCCD detector used to collect these data is composed of 2 panels, as stated above, with two readout electronic back-ends per panel (each containing 4 analogue to digital converters). Each readout is composed of a 2D pixel array of shape $512 \times 512$. In the stack format, the recorded image data are presented in an array with a shape of $(4, 512, 512)$. In this 3D array, the first index is the index of the electronic readout, and the last two are the indexes

---

**Fig. 11** Pseudo SAXS patterns for six different configurations; (first and third rows) pseudo 1D SAXS profile, with the $x$-axis scaled to resolution in nm, and the $y$-axis in arbitrary units. (second and fourth rows) 2D summed SAXS patterns from single-hits after mapping the detector panels to $x$-$y$ coordinates in the laboratory frame. Note: the red circles are to show the center of the pattern and the tile locations and not resolution. As all of the images are of the same size PR772 virus capsid the resolution of the diffraction speckle fringes is an indication of the camera length and hence resolution.
of a specific pixel in that panel. When one would like to represent the actual spatial arrangement of the pixels with a 2D array, one can use psana functions to assemble arrays in the stack format and obtain the corresponding array in the 2D format. Alternatively, one can use the the `corner_positions` and `basis_vectors` datasets to determine the x and y coordinates of each pixel, as documented in the CXIDB file description. In the CXI file, this diffraction data (after conversion to photon counts) is stored in the data set `entry_n/data_1/data`, which is an N × 4 × 512 × 512 unsigned 16bit integer dataset, where N is the number of frames in the experiment run.

In addition to the diffraction data, the datasets `energy`, `pulse_energy` and `pulse_length` contain the X-ray pulse properties, `basis_vectors` and `corner_positions` the detector geometry, mask the detector mask and tags the image classification labels (1 if the diffraction was deemed to have originated from an isolated PR772 molecule and 0 otherwise). For a detailed explanation of these datasets, see the version 1.6 format description at [23].

**Data access.** All datasets described above are deposited in the Coherent X-ray Imaging Data Bank (CXIDB)[23] in the CXIDB data format.

**Data statistics.** The run number range, total hit number, single hit number and the single hit to total hit number ratio are summarized in Table 1. The hit threshold, the number of measured photons required to be classified as a "hit", for amo34117 has been set to a lower value, compared to the other experimental runs, which causes the drastic drop in the single to total hit number ratio.

The detailed distribution of total hits and single hits during each run are summarized in Table 2.

**Technical Validation**

As a measure of the reliability of the datasets, all single-hits from each experiment were summed to form pseudo small angle X-ray scattering (SAXS) patterns (see the first and third rows of Fig. 11). These SAXS patterns are calculated as a function of resolution, accounting for the missing diffraction data and changing detector distance in each dataset, thus one can compare the SAXS profiles across the 6 groups of data.

The second and fourth rows of Fig. 11 show the 2D summed images corresponding to each of the 1D pseudo SAXS profiles. In these summed patterns background and detector artifacts are observable. It is noted that for amo87215 one of the panels had an issue with the readout electronics so that two of the analogue to digital converters read out at a different gain levels. For amo06516 there was a gap in the scatter shield of the second aperture, resulting in an increased level of beamline background signal in the unshielded area, located on the side of the detector (upper part of the image). For amo11416 an analogue to digital converter readout gain issue, similar to amo067215, is also observed. Additionally after run 55 one can observe the increase in the gap of the detector to allow one of the panels to obtain higher resolution. For amo34117 the center four of the analogue to digital converters readouts on one of the panels were not operational.

**Usage Notes**

The dataset contains the recorded data during the experiment in both XTC and CXIDB formats. The dataset also contains a set of pre-selected hits and metadata as described in this paper. XTC files are the native format of LCLS and can be read using analysis frameworks provided by the LCLS (see https://confluence.slac.stanford.edu/display/PSDM/LCLS+Data+Analysis).

**Code availability**

Instructions for downloading and installing `psana` can be found: https://confluence.slac.stanford.edu/display/PSDM/Offsite+Installation.

Received: 11 August 2020; Accepted: 29 October 2020;
Published online: 19 November 2020

**References**

1. Aquila, A. et al. The linac coherent light source single particle imaging road map. *Structural Dynamics* 2, 041701 (2015).
2. Selbert, M. M. et al. Single mimivirus particles intercepted and imaged with an x-ray laser. *Nature* 470, 78–81, https://doi.org/10.1038/nature09748 (2011).
3. Ekberg, T. et al. Single-shot diffraction data from the mimivirus particle using an x-ray free-electron laser. *Scientific Data* 3, 160060, https://doi.org/10.1038/sdata.2016.60 (2016).
4. Benner, W. H. et al. Non-destructive characterization and alignment of aerodynamically focused particle beams using single particle charge detection. *Journal of Aerosol Science* 39, 917–928, https://doi.org/10.1016/j.jaerosci.2008.05.008 (2008).
5. Hanke, M. F. et al. High-throughput imaging of heterogeneous cell organelles with an x-ray laser. *Nature Photonics* 8, 943–949, https://doi.org/10.1038/nphoton.2014.270 (2014).
6. Reddy, H. K. et al. Coherent soft x-ray diffraction imaging of coliphage pr772 at the linac coherent light source. *Scientific data* 4, 170079 (2017).
7. Morgan, A. J. Diffraction data from aerosolized coliphage pr772 virus particles imaged with the linac coherent light source. *Coherent X-ray Imaging Data Bank*, https://doi.org/10.11577/1645124 (2020).
8. 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).
9. DePonte, D. P. et al. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics* 41, 195505, https://doi.org/10.1088/0022-3777/41/19/195505 (2008).
10. Weierstall, U., Spence, J. C. H. & Doak, R. B. Injector for scattering measurements on fully solvated biospecies. *Review of Scientific Instruments* 83, 035108, https://doi.org/10.1063/1.3693040 (2012).
11. Nazari, R. et al. 3d printing of gas-dynamic virtual nozzles and optical characterization of high-speed microjets. *Optics Express* 28, 21749, https://doi.org/10.1364/OE.390131 (2020).
12. Ferguson, K. R. et al. The atomic, molecular and optical science instrument at the linac coherent light source. *Journal of Synchrotron Radiation* 22, 492–497, https://doi.org/10.1107/S1600577515004646 (2015).
Author contributions

A.A., M.B., G.C., H.N.C., M.F., A.M., Z.S., P.W., G.W. participated in instrument development & alignment. B.A., R.A., P.B., J.B., L.F., M.S.H., H.O.J., R.A.K., R.N., M.M.S., D.W., S.Z. participated in sample delivery development. B.A., A.A., K.A., A.Ba., J.B., M.B., L.F., M.S.H., H.O.J., R.A.K., R.N., M.M.S., R.G.S., D.W., S.Z. participated in sample delivery. A.A., M.B., G.C., H.N.C., M.F., A.M., Z.S., P.W., G.W. participated in data collection. A.A., K.A., A.Ba., B.J.D., A.H., R.P.K., H.L., D.L., F.R.N.C.M., A.J.M., J.A.S., M.S., P.S., P.L.Z. participated in sample preparation & characterization. The project “Structural dynamics of biomolecular systems (ELIBIO)” (NO. CZ.02.1.01/0.0/0.0/15_003/0000447) from the European Regional Development Fund (to J.H.). The project “Structural dynamics of biomolecular systems (ELIBIO)” (NO. CZ.02.1.01/0.0/0.0/15_003/0000447) from the European Regional Development Fund (to J.H.).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.A. or A.J.M.

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 Creative Commons Public Domain Dedication waiver http://creativecommons.org/publicdomain/zero/1.0/ applies to the metadata files associated with this article.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2020.

Acknowledgements

Use of the Linac Coherent Light Source, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The research conducted at UWM was supported by the US Department of Energy, Office of Science, Basic Energy Sciences under award DE-SC0002164 (algorithm design and development), and by the US National Science Foundation under awards STC 1231306 (numerical trial models and data analysis) and 1551489 (underlying analytical models). This work was supported by the Cluster of Excellence “CUI: Advanced Imaging of Matter” of the Deutsche Forschungsgemeinschaft (DFG) - EXC 2056 - project ID 390715994. The US National Science Foundation Award 1231306. The NSF Science and Technology Center grant NSF-1231306 (Biological with X-ray Lasers, BioXFEL). NIH grant 5R01GM117342. Helmholtz Associations Initiative and Networking Fund and Russian Science Foundation (Grant No. HRSF-0002/18-41-00001). This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 701647. The Swedish Research Council (to J.H.), The Knut and Alice Wallenberg Foundation (to J.H.), The European Research Council (to J.H.), and the project “Structural dynamics of biomolecular systems (ELIBIO)” (NO. CZ.02.1.01/0.0/0.0/15_003/0000447) from the European Regional Development Fund (to J.H.). The project “Structural dynamics of biomolecular systems (ELIBIO)” (NO. CZ.02.1.01/0.0/0.0/15_003/0000447) from the European Regional Development Fund. The European Research Council—Frontiers in Attosecond X-ray Science: Imaging and Spectroscopy (AXSIS). The Australian Research Council Centre of Excellence in Advanced Molecular Imaging (AMI).

Author contributions

A.A., M.B., G.C., H.N.C., M.F., A.M., Z.S., P.W., G.W. participated in instrument development & alignment. B.A., R.A., P.B., J.B., L.F., M.S.H., H.O.J., R.A.K., R.N., M.M.S., R.G.S., D.W., S.Z. participated in sample delivery & sample delivery development. B.A., A.A., K.A., A.Ba., J.B., M.B., L.F., M.F., R.A.K., A.M., A.J.M., J.A.S., M.S., P.S., P.W., G.W. participated in data collection. A.A., K.A., A.Ba., B.J.D., A.H., R.P.K., H.L., D.L., F.R.N.C.M., A.J.M., C.N., A.O., M.R., P.S., J.A.S., Z.S., I.V. participated in data analysis & software development. J.B., A.P., A.C., H.D., M.F.H., B.G.H., H.O.J., M.M., K.M., A.M., H.K.N.R., M.M.S., M.S., P.L.Z. participated in sample preparation and characterization.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.A. or A.J.M.

Reprints and permissions information is available at www.nature.com/reprints.

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 Creative Commons Public Domain Dedication waiver http://creativecommons.org/publicdomain/zero/1.0/ applies to the metadata files associated with this article.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2020.
Haoyuan Li1,2, Reza Nazari3, Brian Abbey4, Roberto Alvarez5, Andrew Aquila1✉, Kartik Ayer6,7, Anton Barty6,7, Peter Berntsen4, Johan Bielecki8,9, Alberto Pietrini9, Maximilian Bucher10, Gabriella Carini10, Henry N. Chapman6,11, Alice Contreras3, Benedikt J. Daurer8,12, Hasan DeMirci13,14, Leonie Flückiger4, Matthias Frank15, Janos Hajdu9,16, Max F. Hantke9, Brenda G. Hogue3, Ahmad Hosseinizadeh17, Mark S. Hunter2, H. Olof Jönsson18, Richard A. Kirian3, Ruslan P. Kurta9, Duane Loh18, Filipe R. N. C. Maia9, Adrian P. Mancuso6,20, Andrew J. Morgan21✉, Matthew McFadden3, Kerstin Muehlig9, Anna Munke9, Hemanth Kumar Narayana Reddy9, Carl Nettelblad9, Abbas Ourmazd17, Max Rose7, Peter Schwander17, M. Marvin Seibert9, Jonas A. Sellberg18, Raymond G. Sierra1, Zhibin Sun22, Martin Svenda18, Ivan A. Vartanyants7,23, Peter Walter11, Daniel Westphal7, Garth Williams10, P. Lourdu Xavier1,5,6, Chun Hong Yoon1 & Sahba Zaare3

1SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California, 94025, USA. 2Physics Department, Stanford University, 450 Serra Mall, Stanford, California, 94305, USA. 3Arizona State University, 1001 S. McAllister Avenue, Tempe, AZ, 85287, USA. 4ARC Centre of Excellence in Advanced Molecular Imaging, La Trobe University, Bundoora, VIC, 3086, Australia. 5Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761, Hamburg, Germany. 6Center for Free Electron Laser Science, DESY, Notkestrasse 85, 22607, Hamburg, Germany. 7European XFEL, Holzkoppel 4, 22869, Schenefeld, Germany. 8Laboratory of Molecular Biophysics, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3 (Box 596), SE-751 24, Uppsala, Sweden. 9European XFEL, Holzkoppel 4, 22869, Schenefeld, Germany. 10Laboratory of Molecular Biophysics, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3 (Box 596), SE-751 24, Uppsala, Sweden. 11Brookhaven National Laboratory, Bldg 535B, Upton, NY, 11973, USA. 12Centre for Ultrafast Imaging, Luruper Chaussee 149, 22761, Hamburg, Germany. 13Diamond Light Source, Harwell Science & Innovation Campus, Didcot, OX11 0DE, United Kingdom. 14Stanford PULSE Institute, 2575 Sand Hill Road, Menlo Park, California, 94025, USA. 15Koc University, Rumelifeneri, Sariyer Rumeli Feneri Yolu, 34450, Sariyer/Istanbul, Turkey. 16Lawrence Livermore National Laboratory, 7000 East Avenue, L-452, Livermore, California, 94550, USA. 17The European Extreme Light Infrastructure, Institute of Physics, Academy of Sciences of the Czech Republic, Za Radnicí 835, 25241, Dolní Břežany, Czech Republic. 18University of Wisconsin Milwaukee, 3135N. Maryland Ave, Milwaukee, Wisconsin, 53211, USA. 19Department of Applied Physics, KTH Royal Institute of Technology, AlbaNova University Center, KTH Royal Institute of Technology, S-106 91, Stockholm, Sweden. 20Department of Physics, National University of Singapore, 14 Science Drive 4, Blk S1A, Level 2, S1A-02-07, Lee Wee Kheng Building, Singapore, 117557, Singapore. 21Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria, 3086, Australia. 22ARC Centre of Excellence in Advanced Molecular Imaging, School of Physics, University of Melbourne, Parkville, Victoria, 3010, Australia. 23Photon Science Division, Paul Scherrer Institute, CH-5232, Villigen PSI, Switzerland. 24NRNU MEPHI, Kashirskoe shosse 31, 115409, Moscow, Russia. 25✉e-mail: aquila@slac.stanford.edu; morganaj@unimelb.edu.au
Author/s:
Li, H; Nazari, R; Abbey, B; Alvarez, R; Aquila, A; Ayyer, K; Barty, A; Berntsen, P; Bielecki, J; Pietrini, A; Bucher, M; Carini, G; Chapman, HN; Contreras, A; Daurer, BJ; DeMirci, H; Fluckiger, L; Frank, M; Hajdu, J; Hantke, MF; Hogue, BG; Hosseinizadeh, A; Hunter, MS; Joensson, HO; Kirian, RA; Kurta, RP; Loh, D; Maia, FRNC; Mancuso, AP; Morgan, AJ; McFadden, M; Muehlig, K; Munke, A; Reddy, HKN; Nettelblad, C; Ourmazd, A; Rose, M; Schwander, P; Marvin Seibert, M; Sellberg, JA; Sierra, RG; Sun, Z; Svenda, M; Vartanyants, IA; Walter, P; Westphal, D; Williams, G; Xavier, PL; Yoon, CH; Zaare, S

Title:
Diffraction data from aerosolized Coliphage PR772 virus particles imaged with the Linac Coherent Light Source

Date:
2020-11-19

Citation:
Li, H., Nazari, R., Abbey, B., Alvarez, R., Aquila, A., Ayyer, K., Barty, A., Berntsen, P., Bielecki, J., Pietrini, A., Bucher, M., Carini, G., Chapman, H. N., Contreras, A., Daurer, B. J., DeMirci, H., Fluckiger, L., Frank, M., Hajdu, J., ..., Zaare, S. (2020). Diffraction data from aerosolized Coliphage PR772 virus particles imaged with the Linac Coherent Light Source. SCIENTIFIC DATA, 7 (1), https://doi.org/10.1038/s41597-020-00745-2.

Persistent Link:
http://hdl.handle.net/11343/272109

File Description:
Published version

License:
CC BY