Single Mimivirus particles intercepted and imaged with an X-ray laser

M. Marvin Seibert[1*], Tomas Ekeberg[1*], Filipe R. N. C. Maia[1*], Martin Svenda[1], Jakob Andreasson[1], Olof Jönsson[1], Duško Odić[1], Bianca Iwan[1], Andrea Rocker[1], Daniel Westphal[1], Daniel P. DePonte[2], Anton Barty[2], Joachim Schulz[2], Lars Gumprecht[2], Nicola Coppola[2], Andrew Aquila[2], Mengning Liang[2], Thomas A. White[2], Andrew Martin[2], Carl Caleman[1,2], Stephan Stern[2,3], Chantal Abergel[4], Virginie Seltzer[4], Jean-Michel Claverie[4], Christoph Bostedt[5], John D. Bozek[5], Sébastien Boutet[5], A. Alan Miahnahri[5], Marc Messerschmidt[5], Jacek Krzywinski[5], Garth Williams[5], Keith O. Hodgson[5], Michael J. Bogan[6], Christina Y. Hampton[6], Raymond G. Sierra[6], Dmitri Starodub[6], Inger Andersson[7], Saša Bajt[8], Miriam Barthelmess[8], John C. H. Spence[9], Petra Fromme[10], Uwe Weierstall[9], Richard Kirian[9], Mark Hunter[10], R. Bruce Doak[9], Stefano Marchesini[11], Stefan P. Hau-Riege[12], Matthias Frank[12], Robert L. Shoeman[13], Lukas Lomb[13], Sascha W. Epp[14,16], Robert Hartmann[15,17], Daniel Rolles[13,14], Artem Rudenko[14,16], Carlo Schmidt[14,16], Lutz Foucar[13,14], Nils Kimmel[15,18], Peter Holl[15,17], Benedikt Rudek[14,16], Benjamin Erk[14,16], André Hömke[14,16], Christian Reich[15,17], Daniel Pietschner[15,18], Georg Weidenspointner[15,18], Lothar Strüder[14,15,18,19], Günter Hauser[15,18], Hubert Gorke[20], Joachim Ullrich[14,16], Ilme Schlichting[13,14], Sven Herrmann[15,18], Gerhard Schaller[15,18], Florian Schopper[15,18], Heike Soltau[15,17], Kai-Uwe Kühnel[16], Robert Andritschke[15,18], Claus Dieter Schröter[16], Faton Krasniqi[13,14], Mario Bott[13], Sebastian Schorb[21], Daniela Rupp[21], Marcus Adolph[21], Tais Gorkhover[21], Helmut Hirsemann[8], Guillaume Potdevin[8], Heinz Graafsma[8], Björn Nilsson[8], Henry N. Chapman[2,3], Janos Hajdu[1]

(1) Laboratory of Molecular Biophysics, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3 (Box 596), SE-751 24 Uppsala, Sweden

(2) Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany

(3) University of Hamburg, Notkestrasse 85, 22607 Hamburg, Germany

(4) Information Génomique et Structurale, CNRS-UPR2589, Aix-Marseille Université, Institut de Microbiologie de la Méditerranée, Parc Scientifique de Luminy, Case 934, 13288 Marseille Cedex 9, France

(5) LCLS, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

(6) Stanford PULSE Institute, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

(7) Department of Molecular Biology, Swedish University of Agricultural Sciences, Uppsala Biomedical Centre, Box 590, S-751 24 Uppsala, Sweden

(8) Photon Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany

(9) Department of Physics, PSF470, Arizona State University, Tempe, AZ 85287-1504, USA

(10) Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287- 1604 USA

(11) Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

(12) Lawrence Livermore National Laboratory, 7000 East Avenue, Mail Stop L-211, Livermore, CA 94551, USA

(13) Max-Planck-Institut für medizinische Forschung, Jahnstr. 29, 69120 Heidelberg, Germany

(14) Max Planck Advanced Study Group, Center for Free Electron Laser Science, Notkestrasse 85, 22607 Hamburg, Germany

(15) Max-Planck-Institut Halbleiterlabor, Otto-Hahn-Ring 6, 81739 München, Germany

(16) Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany

(17) PNSensor GmbH, Römerstr. 28, 80803 München, Germany

(18) Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, 85741 Garching, Germany

(19) Universität Siegen, Emmy-Noether Campus, Walter Flex Str. 3, 57068 Siegen, Germany

(20) Forschungszentrum Jülich, Institut ZEL, 52425 Jülich, Germany

(21) Institut für Optik und Atomare Physik, Technische Universität Berlin, Hardenbergstrasse 36, 10623 Berlin, Germany

* These authors contributed equally to this work

X-ray lasers are creating new capabilities in understanding the structure of biological systems, complex materials, and matter under extreme conditions[1-4]. Very short and extremely bright coherent X-ray pulses can be used to outrun key damage processes and obtain a single diffraction pattern from a large macromolecule, a virus, or a cell before the sample explodes and turns into

plasma[1]. The continuous diffraction pattern of non-crystalline objects permits over-sampling and direct phase retrieval[2]. Here we show that high quality diffraction data can be obtained with a single X-ray pulse from a non-crystalline biological sample, a single Mimivirus particle, which was injected into the pulsed beam of the first hard X-ray free-electron laser, the Linac Coherent Light Source (LCLS)[5]. Calculations indicate the energy deposited into the virus by the pulse heated the particle to over 100,000 K after the pulse had left the sample. The reconstructed exit wave front (image) yielded 32 nm full period resolution in a single exposure, and shows no measurable damage. The reconstruction reveals a compartmentalized interior inside the virion with a distinct arrangement of dense material (presumably the DNA genome). We expect significantly higher resolutions in such experiments with shorter and brighter photon pulses focused to a smaller area. Resolution in such experiments can be further extended for samples available in multiple identical copies.

Diffraction studies on crystalline samples have led to spectacular breakthroughs in physics, chemistry, and biology over the past hundred years. Many important targets are difficult or impossible to crystallise, and this creates systematic blank areas in structural sciences. X-ray lasers offer the possibility to step beyond crystallography, and extend structural studies to single, non-crystalline particles or molecules[1]. This paper presents first results on biological imaging with an X-ray free-electron laser, bringing together all elements required for structural studies on single, non-crystalline objects.

Mimivirus (*Acanthamoeba polyphaga* Mimivirus) is the largest known virus[6]. Its size is comparable to the size of the smallest living cells (in fact, the name Mimivirus stands for microbe mimicking virus). The viral capsid (0.45 μm in diameter) has a pseudo-icosahedral appearance and is covered by an outer layer of dense fibrils[7,8], giving a total diameter of about 0.75 μm for the particle. Mimivirus is too big for a full

three-dimensional reconstruction by cryo-electron microscopy (cryo-EM, ref. 7) and its fibrils prevent crystallisation. The genome[9] has 1.2 million base pairs (comparable to a small bacterium) and contains some genes previously thought only to be present in cellular organisms, including components of the protein translation apparatus. Mimivirus can be infected by a smaller virus, named a "virophage"[10,11], which itself appears to be the first example of a virus, behaving as a parasite of another virus. Studies on Mimivirus are creating a paradigm shift in virology and have led to renewed debates about the origin and the definition of viral and cellular life[8].

Figure 1 shows the experimental arrangement for imaging single virus particles. The sample injector, which uses aerodynamic focusing, was mounted into the CFEL-ASG Multi-Purpose (CAMP) instrument[12] at the Atomic, Molecular and Optical Science beam line[13] (AMO) of the LCLS[5]. Far-field diffraction patterns were recorded at a reduced pressure (10^{-6} mbar) to minimize background scattering. Mimivirus was aerosolised from a volatile buffer (250 mM ammonium acetate, pH 7.5) using a gas dynamic nebuliser[14] in a helium atmosphere. The beam of adiabatically cooled virus particles was guided through an aerodynamic lens stack (similar to the one described in ref. 15) and entered the interaction zone with an estimated velocity of 60-100 m/s. The particles were intercepted randomly by the LCLS pulses. The X-ray energy was 1.80 keV (6.9 Å wavelength) and the pulse length was 70 fs (full width at half maximum, FWHM). The X-ray beam diameter at the interaction point was about 10 μm (FWHM), with a maximum of 1.6×10^{10} photons/μm^2 in the centre of this beam. This translates to a peak power density of 6.5×10^{15} W/cm^2. Forward-scattered diffraction patterns were recorded on a pair of pn-junction charge-coupled device detectors[12] (pnCCD). The direct beam exited through an opening between the two detector halves, and was absorbed in a beam dump at a distance behind the detectors (Figure 1). The detector pair was placed 564 mm away from the interaction point, giving a maximum full period

resolution of 10.16 nm at the edges and 7.20 nm at the corners of the compound detector at 1.8 keV energy.

Figure 2a and 2b show single shot X-ray diffraction patterns of individual Mimivirus particles, and Figure 2c shows a transmission electron micrograph of a Mimivirus particle. Each of the diffraction patterns contains about 1.7 million scattered photons. The lowest-resolution data are missing between the two detector halves and so the total number of scattered photons exceeds this number. Figures 2d and 2e show autocorrelation functions calculated from the diffraction patterns. Missing low-resolution data act as a high-pass filter. For an object of extent D, the extent of its autocorrelation is $2D$, and the diffraction intensities are band-limited with a Nyquist rate of $1/(2D)$. The size and shape of the autocorrelation functions in Figures 2d and 2e are indicative of hits on single virus particles. Figures 2f and 2g show the reconstructed exit wave fronts for these Mimivirus particles. The shape and size of the reconstructed objects agree with data from prior cryo-EM studies in which 30,000 images were averaged[7]. In contrast, the reconstructed structures in Figure 2f and 2g come from single shots from single particles, and demonstrate the power of this imaging concept[1].

Image reconstruction was performed by iterative phase retrieval implemented in the Hawk software package[16], using the RAAR algorithm[17] enhanced with both reality and positivity constraints. The support was handled by a Shrinkwrap algorithm[17] with the constraint to have a specific area that was estimated from the autocorrelation function. Weakly constrained modes[19] in the reconstructions were identified and removed, using the formalism of Thibault et al.[19]. This is a linear algebra method to compensate for noise, or the lack of constraints in the missing central region of the pattern. The uncertainty in the overall density was less than 10% after the identification and removal of the unconstrained modes. We then fitted these modes to match the total density of a spherical or a suitably rotated icosahedral profile. The missing modes were

adjusted to give a total density that best matched the target. Residual phase fluctuations were suppressed by averaging many hundred reconstructions following this procedure, using different random seeds. The results gave improved image reliability. For details see "Full Methods".

We estimate the image resolution in the reconstruction by computing the phase-retrieval transfer function[2,20] (PRTF, Figures 2h and 2i), which represents the confidence in the retrieved phases as a function of resolution. No consensus has emerged so far on what single PRTF value should be used as the measure of resolution (values between 0.5 and 0.1 can be found in the literature, see ref. 32 in Full Methods). We characterise resolution by the point where the PRTF drops to 1/e (ref. 20) and this corresponds to a full period resolution of 32 nm in both cases. We expect significantly higher resolutions in such experiments with shorter and brighter photon pulses focused to a smaller area.

In principle, resolution could reach less than one nanometer in a single exposure with a biological object of similar size as the Mimivirus particle[3]. This resolution would require a FEL pulse shorter than about 5 fs at 1.8 keV energy and a photon flux of more than 3×10^{11} photons/μm^2 on the sample[3]. This pulse length and photon flux are beyond the initial capabilities of the LCLS, although there are indications already now for near transform-limited LCLS pulses, lasting only a few femtoseconds and containing about 5×10^{11} photons per pulse in the unfocused beam[21].

With very short pulses, exposures would be over before there is time for significant Auger emission or for the development of secondary electron cascades in the sample[1]. The conventional handicap of X-rays over electrons in imaging could thus be reversed and made into a net gain over a broad range of sample sizes. First experiments at free-electron lasers show a significant drop in the photoelectric cross section of

hollow atoms[22]. This effect was predicted earlier[1] but it is larger than expected, and can already be measured with LCLS pulses of 20-80 fs duration[22]. In the first of such experiments, photoabsorption decreased 20-fold in hollow neon to equal the cross-section of coherent scattering[22]. The lifetime of double core holes extends to tens of femtoseconds[22]. In C, N and O, more than 90% of the photoelectric cross section comes from 1s electrons with 1.8 keV photons. Ejection of these electrons at the beginning of an intense and short pulse could practically stop photo-ionisation without significantly changing the elastic cross-sections of outer-shell electrons.

We see no measurable sample deterioration. The explosion of micron-sized objects is hydrodynamic[3], and burns the sample from the outside inwards, rarefying and destroying outer contours first. Trapped electrons move inwards to neutralise an increasingly positive core, and leave behind a positively charged outer layer, which then peels off[23]. The reconstructed exit wave front of the Mimivirus particle shows well-defined outer contours, and gives a sample size consistent with the intact virus capsid (we do not expect to see the thin viral fibrils at the length scales accessible here). Other studies on protein nanocrystals[24] at the LCLS at 0.9 nm resolution show no measurable deterioration of Bragg peaks during illumination with pulses similar to those used here. The size of these protein nanocrystals was similar to the size of Mimivirus particles.

At this stage, it is unclear how reproducible is the interior structure of Mimivirus particles (or that of any other viral particles) in terms of atomic positions, and this will need further studies. The viral inner capsid consists of a thin protein shell (about 7 nm thick) followed by phospholipid membranes. The structure of the protein shell appears to be reproducible to at least 6.5 nm resolution[7]. Figures 2d and 2e suggest a compartmentalized structure of the virion core that exhibits an unexpected distribution of dense material (presumably the DNA genome). The internal structure of the

Mimivirus particle does not necessarily follow the pseudo-icosahedral outer shape, which is believed to exhibit a single five-fold symmetry axis[7].

The penetration depth of X-rays permits studies on the interiors of large objects. The methods applied here require no modifications to the sample such as staining, freezing, sectioning, radiolabelling or crystallisation, and can also be used to image cells that are alive at the time of the exposure. The amount of missing data can be reduced by adding an additional detector pair at a distance behind the first pair of detectors. Another necessary improvement is to increase the dynamic range. In these first experiments, there were shots extending to significantly higher resolutions than those reported here but these contained too many saturated pixels, preventing image reconstruction. With reproducible samples, where the experiment can be repeated many times on a new object, a three-dimensional data set can be collected, and the resolution extended (even from weak individual exposures) by merging redundant data[25-29]. Studies on virus particles with higher intensity photon pulses and improved detectors could answer the question of whether the core is reproducible to sub-nanometer resolution or whether the viral genome exhibits the "molecular individualism" that genomic DNA structures explore in vitro[30].

Methods summary

The experiment was carried out with the "CFEL ASG multi purpose" (CAMP) instrument[12] on the Atomic, Molecular and Optical Science beam line[13] at the Linac Coherent Light Source in Stanford, California. The CAMP instrument supports a large variety of photon-imaging and atomic and molecular physics experiments. In the present study, diffraction patterns were recorded on a pair of specifically built pnCCD detectors[12] capable of a maximum readout speed of 250 images/s. The detector pair was placed 564 mm away from the interaction point. The active area for each detector half

was 76.8 mm x 38.4 mm and contained 1024x512 pixels of 75x75 μm^2 dimensions. Each pixel has a full well capacity of 282,000 electrons, corresponding to 571 X-ray photons at 1.8 keV photon energy. The electron bunch length was 70 fs long (FWHM). The corresponding photon bunch is believed to be shorter than the electron bunch[22]. The photon bunch contained 8×10^{11} photons (0.24 mJ at 1.8 keV) and the pulse diameter was about 10 μm (FWHM) at the interaction point, giving 1.6×10^{10} photons/μm^2 in the centre of the beam and a peak power density of 6.5×10^{15} W/cm^2. Backgound scattering from the residual gas in the vacuum chamber did not exceed the readout noise of the detectors nor the noise of the diffuse photon background (which was below 1.3 photons/pixel). This is quite remarkable, considering that the number of photons in the pulse was nearly 100 billion times higher than the background.

Purified Mimivirus was transferred into a volatile buffer (250 mM ammonium acetate, pH 7.5) and the suspension was aerosolised with helium in a gas dynamic nebuliser[14]. The aerosol of hydrated virus particles was sampled into a differentially pumped injector via an inlet nozzle coupled to a skimmer. The aerosol (in helium atmosphere) passed through a variable volume relaxation chamber from where the equilibrated and adiabatically cooled particles entered a differentially pumped aerodynamic lens stack. Particles focused by the aerodynamic lens were intercepted randomly by the LCLS pulses. Diffraction patterns of the free-flying virus particles were exceptionally clean.

Image reconstruction was performed with the open source Hawk software[16], available under the GNU General Public License from http://xray.bmc.uu.se/hawk. The background-corrected diffraction patterns of mimivirus particles and the Hawk configuration files used in the reconstructions are available from the Hawk web site.

References

1. Neutze, R. et al. Potential for femtosecond imaging of biomolecules with X-rays. *Nature* 406, 752-757 (2000).

2. Chapman, H. N., et al., Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nature Physics* 2, 839-843 (2006).

3. Bergh, M., et al. Feasibility of imaging living cells at sub-nanometer resolution by ultrafast X-ray diffraction. *Quar. Rev. Biophys.* 41, 181-204 (2008).

4. Nagler, B. et al. Turning solid aluminium transparent by intense soft X-ray photoionization. *Nature Phys.* 5, 693-696 (2009).

5. Emma, P. et al., First lasing and operation of an Ångström-wavelength free-electron laser. *Nature Photonics*, 4, 641-647 (2010).

6. La Scola, B. et al., A giant virus in amoebae. *Science* 299, 2033 (2003).

7. Xiao, C. et al., Structural Studies of the Giant Mimivirus, *PLoS Biology* 7, e1000092 (2009).

8. Claverie, J.M, Abergel C., Mimivirus: the emerging paradox of quasi-autonomous viruses. *Trends Genet.* 26, 431-437 (2010).

9. Raoult, D. et al., The 1.2-megabase genome sequence of Mimivirus. *Science* 306, 1344-1350 (2004).

10. La Scola, B. et al., The virophage as a unique parasite of the giant Mimivirus. *Nature* 455, 100–4 (2008)

11. Claverie, J.M, Abergel C. Mimivirus and its virophage. *Annu. Rev. Genet.* 43, 49–66 (2009).

12. Struder, L. et al. Large-format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources. *Nucl. Instr. Meth. Phys. Res. A* 614, 483-496 (2010).

13. Bozek, J.D. AMO instrumentation for the LCLS X-ray FEL. *Eur. Phys. J. Special Topics* 169, 129–132 (2009).

14. DePonte et al., Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Physics D-Appl. Phys.* 41, 195505 (2008).

15. Bogan, M.J. et al. Single particle X-ray diffractive imaging. *NanoLetters*, 8, 310-316 (2008).

16. Maia, F.R.N.C., Ekeberg, T., van der Spoel, D., Hajdu, J. Hawk: the image reconstruction package for coherent X-ray diffractive imaging. *J. Appl. Cryst,* 43 [doi:10.1107/S0021889810036083] (2010).

17. Luke, D.R., Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems* 21, 37-50 (2005).

18. Marchesini, S. et al., X-ray image reconstruction from a diffraction pattern alone. *Phys. Rev. B* 68, 140101(R) (2003)

19. Thibault, P., Elser, V., Jacobsen, C., Shapiro, D., Sayre, D., Reconstruction of a yeast cell from X-ray diffraction data. *Acta Cryst.* A62, 248-261 (2006)

20. Shapiro, D. et al. Biological imaging by soft x-ray diffraction microscopy. *Proc. Nat. Acad. Sci. USA* 102, 15343–15346 (2005).

21. Ding, Y. et al. Measurements and Simulations of Ultralow Emittance and Ultrashort Electron Beams in the Linac Coherent Light Source. *Phy. Rev. Letts.* 102, 254801 (2009).

22. Young, L., et al., Femtosecond electronic response of atoms to ultra-intense X-rays. *Nature* 466, 56-U66 (2010).

23. Hau-Riege, S. P. et al. A sacrificial tamper slows down sample explosion in flash diffraction experiments. *Phys. Rev. Letts.* 104, 064801 (2010).

24. Chapman H. N. et al., Femtosecond X-ray protein nanocrystallography, *Nature,* under review (2010).

25. Huldt, G., Szoke, A., Hajdu, J., Diffraction imaging of single particles and biomolecules. *J. Struct. Biol.* 144, 219 –227 (2003).

26. Fung, R. et al. Structure from fleeting illumination of faint spinning objects in flight. *Nature Phys.* 5, 64 (2009).

27. Loh N. T. D., Elser V., Reconstruction algorithm for single-particle diffraction imaging. *Phys. Rev. E* 80, 026705 (2009).

28. Bortel, G., Faigel, G., Tegze, M., Classification and averaging of random orientation single macromolecular diffraction patterns at atomic resolution, *J. Struct. Biol.* 166, 226-233, (2009).

29. Maia, F.R.N.C., Ekeberg, T., Tımneanu, N., van der Spoel, D., Hajdu, J. "Structural variability and the incoherent addition of scattered intensities in single-particle diffraction", *Physical Reviews E* 80, 031905 (2009).

30. Perkins, T. T., Smith, D. E., Chu, S. Single polymer dynamics in an elongated flow. *Science* 276, 2016-2021 (1997).

Author contribution J.H. and H.N.C. conceived the experiment. C.A., V.S., J.-M.C., M.S., O.J., A.R. I.A. and D.O. prepared and characterised the samples. J.D.B., C.B., K.O.H. created the LCLS beam line. J.H, H.N.C., J.S., L.G., A.B., N.C., A.A., A.M., J.K., K.O.H. and S.P.H.-R. developed the imaging concept. S.W.E., R.H., D.R., A.Ru., C.Sch, L.F., N.K., P.H., B.R., B.E., A.H., Ch.R., D.P., G.W., L.S., G.H., H.G., J.U., I.S., S.H., G.Sch., F.Sch., H.S., K.-U.K., R.A., K.-D.S., F.K., Ma.B., S.Sch., D.Ru., M.A., T.G., H.H., L.G., G.P., H.G., and B.N. designed and set up the CAMP instrument and/or developed and operated the pnCCD detectors. S.Ba. and M.Bar. coordinated instrumentation, worked on engineering, and prepared filters and calibration samples. J.K., S.H.-R., A.B., H.N.C., J.Sch., A.M., N.C. characterised the focus. D.W., B.I., M.M.S., M.S. and J.H. built the sample injector. D.P.D., J.H., M.S., D.W., U.W., R.K., M.H., R.B.D., J.H.C.S. designed and built the nanospray nebuliser. J.H., M.S., M.M.S, F.R.N.C.M., J.A., A.A.M., A.R., M.J.B., Ch.H., R.S., S.Bou., I.A., O.J., D.S., A.B., J.S., D.P.D., A.A., M.L., J.K. T.A.W., A.M., R.L.S., L.L., M.Bar., L.M., J.C.H.S., P.F., I.S., U.W., R.K., M.H., B.D., M.F., G.W., M.Bot., S.W.E., B.E., L.F., R.H., N.K., L.L., D.R., B.R., A.Ru., R.L.S., L.S., I.S., C.S., J.U. and H.N.C. characterised the imaging apparatus and carried out the experiment. T.E., F.R.N.C.M., M.M.S., A.B., T.A.W., A.M., A.A., S.M., M.M., J.H., J.C. analysed the data. T.E., F.R.N.C.M., A.B. and S.M. preformed image reconstructions. J.H. and T.E. wrote the first draft of the manuscript, which was discussed and improved by contributions from all authors.

Author information: The authors declare no competing financial interests.

Correspondence and requests for materials should be addressed to J.H. (janos.hajdu@xray.bmc.uu.se)

Requests concerning Mimivirus should be addressed to C.A (Chantal.Abergel@igs.cnrs-mrs.fr).

Figure 1. The experimental arrangement. Mimivirus particles were injected into the pulse train of the Linac Coherent Light Source (LCLS, Stanford, CA) at the AMO experimental station[13] with a sample injector built in Uppsala. This was mounted into the ASG-CAMP instrument[12]. The aerodynamic lens stack is visible on the left in the centre of the injector body. Particles leaving the injector

enter the vacuum chamber, and are intercepted randomly by the LCLS pulses. The far-field diffraction pattern of each particle hit by an X-ray pulse is recorded on a pair of fast pnCCD detectors[12]. The intense direct beam passes through an opening in the centre of the detector assembly and is absorbed harmlessly behind the sensitive detectors. Some of the low-resolution data also go through this gap and are lost in the current setup.

Figure 2: Single shot diffraction patterns on single virus particles give interpretable results. **a** and **b,** Experimentally recorded far-field diffraction patterns (in false colour representation) from individual virus particles captured in two different orientations. **c,** Transmission electron micrograph of an unstained Mimivirus particle with pseudo-icosahedral appearance. **d** and **e,** Autocorrelation functions for **a** and **b**. The shape and size of the autocorrelations corresponds to those of a single virus particle after high-pass filtering with the missing low-resolution data in the horizontal gap. **f** and **g,** Reconstructed images after iterative phase retrieval with the Hawk software package[16]. The size of a pixel corresponds to 9 nm in the images. Three different reconstructions are shown for each virus particle: Averaged reconstruction with unconstrained Fourier modes[19], followed by averaged images after fitting unconstrained low-resolution modes to a spherical or an icosahedral profile. The orientation of the icosahedron was determined from the diffraction data. The results show small differences between the spherical and icosahedral fits. **h** and **i,** The phase retrieval transfer function (PRTF) for reconstructions where the unconstrained low-resolution modes were fitted to an icosahedron. All reconstructions gave similar resolutions. We characterise resolution by the point where the PRTF drops to 1/e (ref. 20). This corresponds to 32 nm full period resolution in both exposures. Arrows mark the resolution range with other cut-off criteria found in the literature (ref. 32 in Full Methods).

Resolution can be substantially extended for samples available in multiple identical copies[1,25-28]. Work is under way at the LCLS to reduce the amount of missing low-resolution data and to increase the photon flux on the sample.

Full Methods

Experimental set up at the LCLS. The experiment was carried out with the "CFEL ASG multi purpose" (CAMP) instrument[12] on the Atomic, Molecular and Optical Science beam line[13] at the Linac Coherent Light Source[5] in Stanford, California. The CAMP instrument supports a large variety of photon-imaging and atomic and molecular physics experiments. Diffraction patterns were recorded in this instrument on a pair of pn-CCD detectors[12]. The active area for each detector half was 76.8 mm x 38.4 mm, containing 1024x512 pixels of 75x75 μm^2 dimensions. The detector pixels have a full well capacity of 282,000 electrons, corresponding to 571 X-ray photons at 1.8 keV photon energy. The electron bunch length was 70 fs long (FWHM). The photon bunch length is believed to be shorter than the electron bunch length[22]. The measured integrated energy of the photon pulse entering the beam line was 1.73 mJ, corresponding to about $6x10^{12}$ photons per pulse from the LCLS. The beam line transmission was 13.5%, and thus the number of photons delivered to the experiment was $8x10^{11}$ (0.24 mJ). The beam diameter at the interaction point was about 10 μm (FWHM), giving a maximum of $1.6x10^{10}$ photons/μm^2 in the centre of this beam. This gives a peak power density of $6.5x10^{15}$ W/cm^2.

"Containerless" sample handling. Purified Mimivirus[31] particles were transferred into a volatile buffer (250 mM ammonium acetate, pH 7.5) and the suspension (10^{12} particles/ml) was aerosolised at a rate of about 5 μl/min, using helium gas in a gas dynamic nebuliser[14]. The aerosol of hydrated virus particles was sampled into a differentially pumped sample injector via an inlet nozzle coupled to a skimmer. Most of the nebulising gas, and vapours of the volatile buffer, were pumped away at this point.

The heavier aerosol (in a wet helium atmosphere at 10^{-2} mbar pressure) passed through a variable volume relaxation chamber from where the equilibrated and adiabatically cooled particles entered a differentially pumped aerodynamic lens stack. The pressure dropped from about 10^{-2} mbar to about 10^{-4} mbar at the exit of the lens. Particles focused by the aerodynamic lens entered the interaction zone (10^{-6} mbar) with an estimated velocity of 60-100 m/s, and were intercepted randomly by the LCLS pulses. The diffraction patterns of free-flying virus particles were exceptionally clean. Background scattering from the residual helium gas and from traces of the volatile buffer did not exceed the readout noise of the detectors nor the noise of the diffuse photon background (this was below 1.3 photons/pixel).

Data processing included removal of signal from known bad or saturated pixels, correction for the residual common mode offsets, and application of a flat-field correction. After this, the patterns were used directly without symmetrisation.

Image reconstruction was performed with the open source Hawk software package[16], using the RAAR algorithm[17] and its support constraint, enhanced by additional reality and positivity constraints. A Fourier constraint was applied to match Fourier amplitudes with experimental amplitudes. This was performed by a projection. No explicit Fourier constraints were used for regions of missing data, although these regions were implicitly constrained in Fourier space by the real space constraints. The support was handled by a Shrinkwrap algorithm[18] with the constraint to have a specific area that was estimated from the autocorrelation function.

Weakly constrained modes[19] in the reconstructions were identified and removed, using the formalism of Thibault et al.[19]. This is a linear algebra method to compensate for noise, or the lack of constraints in the missing central region of the pattern. The diffracted amplitudes in the region of missing data can be recovered by iterative phasing algorithms but for patterns where this region is extensive, the recovered amplitudes will

be unreliable[19]. These modes were identified and their constraining power was calculated by performing a singular-value decomposition on the transform from the region of missing data to the support. The singular values identify the modes that are most weakly constrained, as the singular vectors, and determine their constraining power. In the patterns discussed in this paper there are modes with very low constraining power, and these are therefore virtually unconstrained, and whose strength, therefore, had to be estimated in another way. The threshold used for identifying unconstrained modes was 0.999, corresponding to a constraining power of 0.045. The uncertainty in the total image density dropped to below 10% after removing these modes. Missing modes were fitted to match the total density of a spherical or a suitably rotated icosahedral profile. The missing modes were adjusted to give a total density that best matched the target.

The number of weakly constrained modes that were classified as unconstrained differs slightly between the two reconstructions since the support recovered through the Shrinkwrap algorithm[17] is slightly different in the two cases. For the reconstruction starting from Figures 2a, the median number was 8 and the average was 7.75 modes. For the reconstruction starting from Figure 2b the number was slightly higher: median 12 and average 11.85. This difference is due to a larger area missing in the centre of Figure 2b due to more saturated pixels.

Residual phase fluctuations were suppressed by averaging many reconstructions, using different random seeds. The results gave improved image reliability. For reconstructions from Figure 2a, 10,000 iterations were used and 200 reconstructions were obtained from different random starting positions and then used to calculate the PRTF. The support was updated every 20 iterations. All 200 reconstructions had a Fourier error below a threshold of 0.33. For reconstructions from Figure 2b, 40,000 iterations were performed and 94 reconstructions obtained. Out of these only 56 reconstructions had a Fourier error below a threshold of 0.33. The differences underline the deleterious effect of

missing low-resolution data on image reconstruction. Reliable image reconstruction needs a more efficient way of measuring low angle diffraction data, including a wider dynamic range for the detector. This could be achieved by placing a graded attenuator around the central hole of the pnCCD detector pair.

We estimate the image resolution in the reconstruction by computing the phase-retrieval transfer function[2,20]. No consensus has emerged so far on what single PRTF value should be used as the measure of resolution. Values between 0.5 and 0.1 can be found in the literature[32]. We characterise resolution by the point where the PRTF drops to 1/e (ref. 20). Diffraction data extend to higher resolution than the resolution given by the PRTF.

Depth of field: The angle spanned by the signal was small enough for the entire particle to fit within the depth of field. Defocus effects are therefore avoided by using a reality constraint. A resolution of 1 nm (same X-ray wavelength) would require measuring at high angle, leading to significant deviation from the projection image. Real-value constraints would not work in the latter case, and this would make the reconstruction more challenging, but by no means impossible (see e.g. ref. 33).

Transmission electron microscopy was performed with a Hitachi H-7100 electron microscope on unstained Mimivirus particles deposited on Formvar®-coated gold grids.

A route for improvements: There is a clear need to achieve higher resolution in single shots. This requires an increased photon flux on the sample as well as a wider dynamic range for detecting photons in the diffraction pattern. The LCLS[5] is capable of delivering very short X-ray pulses[21] to outrun significant sample explosion with more photons per pulse. Tighter focusing and an increased photon output from the FEL have already been achieved and will increase the flux on samples in forthcoming runs. A broad dynamic range in detecting photons is necessary to avoid saturation at low angles. In this first set

of experiments, there were already exposures with significantly higher resolutions than those reported here but these exposures contained too many saturated pixels at low angles, preventing image reconstruction. Reliable image reconstruction needs a more efficient way of measuring low angle diffraction data. A graded attenuator around the central hole of the pnCCD detector pair could help here. An additional pair of detectors placed far behind the first detector pair could record more of the low angle data over a larger area. Maintaining sample integrity during injection is a key requirement in the experiment. More data on a diverse set of samples (such as cells, viruses and macromolecules) will be needed to map out the available parameter space here. Hit rates could be increased by improved injection methods, using a narrower particle beam. A future extension to imaging single macromolecules will need these developments.

Methods References

31. Byrne, D. et al., The polyadenylation site of Mimivirus transcripts obeys a stringent 'hairpin rule'. *Genome Res.* 19, 1233-1242 (2009).

32. Steinbrener, J., Nelson, J., Huang, X., Marchesini, S., Shapiro, D., Turner, J.J., Jacobsen, C., Data preparation and evaluation techniques for x-ray diffraction microscopy. *Optics Express* **18**, 18599 (2010).

33. Seibert, M.M. et al., Femtosecond diffractive imaging of biological cells. *J. Phys. B: At. Mol. Opt. Phys.* 43, 194015 (2010).
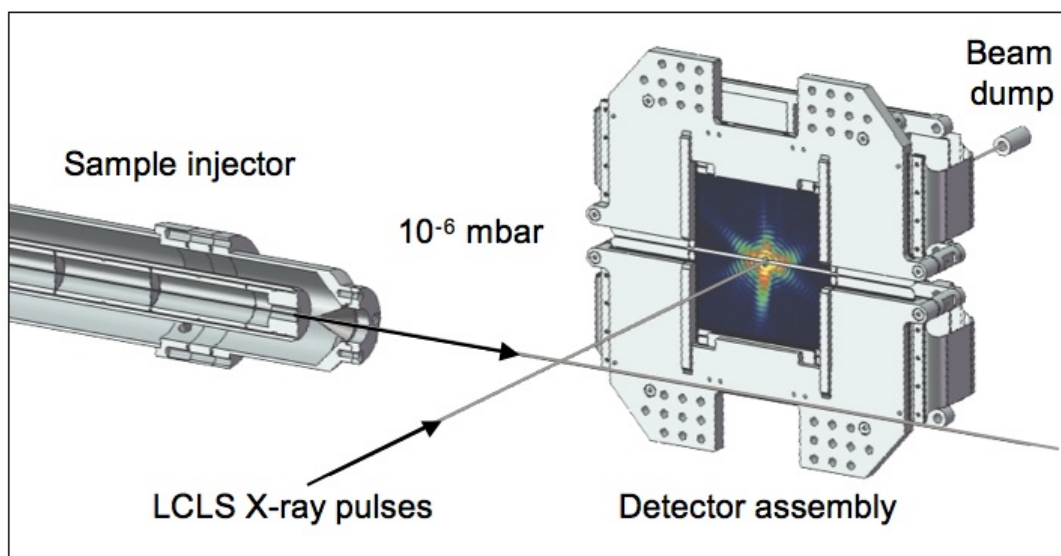
Figure 1

Figure 2