Atomic structure of granulin determined from native nanocrystalline granulovirus using an X-ray free-electron laser

Cornelius Gati[1], Dominik Oberthuer[1], Oleksandr Yefanov[1], Richard D. Bunker[2**], Francesco Stellato[1], Elaine Chiu[2], Shin-Mei Yeh[2], Andrew Aquila[1,3], Shibom Basu[5****], Richard Bean[1,3], Kenneth R. Beyerlein[1], Sabine Botha[4*****], Sébastien Boutet[6], Daniel DePonte[1,7], R. Bruce Doak[8], Raimund Fromme[5], Lorenzo Galli[1], Ingo Grotjohann[9], Daniel R. James[8], Christopher Kupitz[5******], Lukas Lomb[4], Marc Messerschmidt[6*******], Karol Nass[1], Kimberley Rendek[9],Robert L. Shoeman[4], Dingjie Wang[8], Uwe Weierstall[5,8], Thomas A. White[1], Garth J. Williams[6********], Nadia A. Zatsepin[5,8], Petra Fromme[5],John Spence[5,8], Kenneth N. Goldie[10],Johannes A. Jehle[11], Peter Metcalf[2*], Anton Barty[1], Henry N. Chapman[1,12,13*]**

[1] Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany [2] School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142, New Zealand [3] European XFEL GmbH, Albert-Einstein-Ring 19, Hamburg, 22761, Germany [4] Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Jahnstrasse 29, Heidelberg, D-69120, Germany [5] School of Molecular Sciences and Biodesign Center for Applied Structural Discovery, Arizona State University, Tempe AZ, 85287-1604, USA [6] Linac Coherent Light Source, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA, USA [7] SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, 94025, USA [8] Department of Physics, Arizona State University, Tempe AZ, 85287, USA [9] School of Chemistry & Biochemistry, Arizona State University, Tempe AZ, 85287, USA [10] Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University Basel, Basel [11] Institute for Biological Control, Julius Kuehn Institute (JKI), Heinrichstr. 243, 64287 Darmstadt, Germany [12] Department of Physics, University of Hamburg, Jungiusstrasse 6, Hamburg, 20355 Germany [13] Center for Ultrafast Imaging, University of Hamburg, Luruperstrasse, Hamburg, 20355 Germany

Submitted to Proceedings of the National Academy of Sciences of the United States of America

To understand how molecules function in biological systems, new methods are required to obtain atomic resolution structures under physiological conditions. Intense femtosecond-duration pulses from X-ray free-electron lasers (XFELs) can out-run most damage processes, vastly increasing the dose before the specimen is destroyed. This in turn allows structure determination from crystals much smaller and more radiation sensitive than previously considered, allowing data collection at room temperature and avoiding structural changes due to cooling. High-resolution structures obtained from XFEL data mostly use crystals far larger than 1 μm^3 in volume, while the X-ray beam is often attenuated to protect the detector from damage caused by intense Bragg spots. Here, we describe the 2 Å resolution structure of native nanocrystalline granulovirus occlusion bodies (OBs) that are less than 0.016 μm^3 in volume using the full power of the Linac Coherent Light Source and a dose up to 1.3 GGy per crystal. The crystalline shell of granulovirus OBs consists, on average, of about 9,000 unit cells, representing the smallest protein crystals to yield a high-resolution structure by X-ray crystallography to date. The XFEL structure shows little to no evidence of radiation damage and is more complete than a model determined using synchrotron data from recombinantly produced, much larger, cryo-cooled granulin microcrystals. Our measurements suggest that it should be possible, under ideal experimental conditions, to obtain data from protein crystals with only 100 unit cells in volume using currently available XFELs and suggest that single molecule imaging of individual biomolecules could almost be within reach.

XFEL | nanocrystals | structural biology | bio-imaging

Introduction

Imaging of biomolecules using radiation of short enough wavelength to resolve individual atoms is limited by radiation damage which destroys the very structure being measured. Energy absorption is unavoidable because the ratio of elastic scattering to damaging photoabsorption is an inherent property of atoms. Absorbed dose is therefore proportional to the scattered intensity, and thus the measured signal, used to determine the structure. The maximum tolerable dose that the sample can withstand fundamentally limits atomic resolution structure determination (Henderson 1995; Garman 2010). The tolerable dose, and therefore damage, is a function of spatial resolution, with fine detail being destroyed first (Howells et al. 2009). Radiation damage is typically overcome by distributing the dose over many identical copies of the same molecule, either using crystals containing many aligned copies of the same molecule, or by aligning images of individual molecules in the case of single particle electron microscopy. At room temperature, free radical production after a dose as small as 1 kGy (gray = J kg^{-1} absorbed energy) causes rapid decay of biological samples (Garman 2010). This decay can be reduced by cooling the sample to below ∼120 K. Doses up to about 30 MGy at synchrotron beam lines are routinely used for atomic structure determination from cryo-cooled protein crystals larger than about 100 μm^3. Similar dose limits apply in single particle cryo-electron microscopy, where individual images must

Significance

The room temperature structure of natively formed protein nanocrystals consisting of 9,000 unit cells has been solved to 2 Å resolution using an unattenuated X-ray free-electron laser beam, representing, by far, the smallest protein crystals used for protein structure determination by X-ray crystallography to date. Radiation damage limits structure determination from protein crystals using synchrotron techniques, while femtosecond X-ray pulses from free-electron lasers enable much higher tolerable doses, extracting more signal per molecule allowing the study of sub-micrometer crystals. Radiation sensitive features, such as disulfide bonds, are well resolved in the XFEL structure despite the extremely high dose (1.3 GGy) used. Analysis of signal obtained in this experiment indicates that structure determination from even smaller protein crystals could be possible.

Reserved for Publication Footnotes

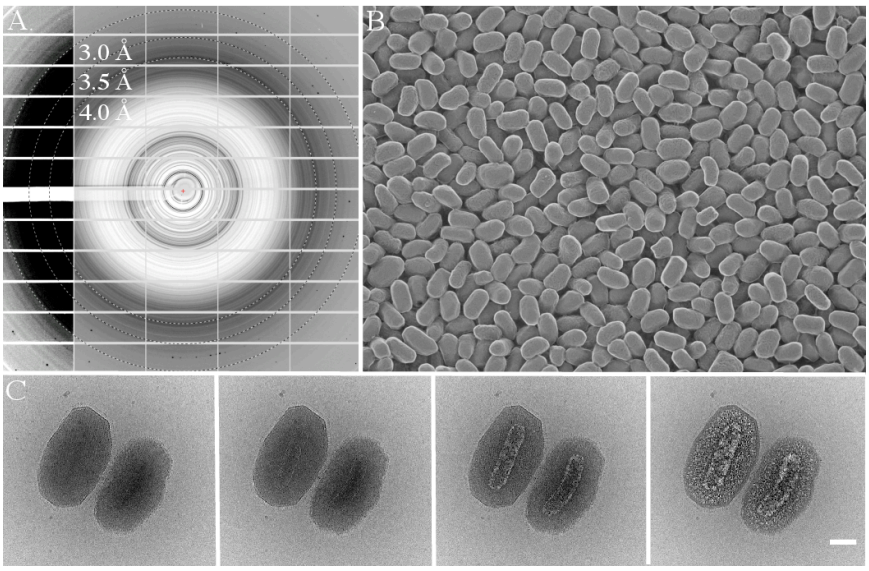

Fig. 1. Granulovirus occlusion bodies contain a single virion surrounded by a crystalline protein layer that diffracts to high resolution. a) Powder X-ray diffraction from a pellet of granulovirus OBs at 100K (Methods). Protein diffraction rings extend to a resolution between 3 and 3.5 Å. The detector panels on the left with enhanced contrast show evidence of diffraction at even higher resolution. Resolution rings are shown at 4, 3.5 and 3Å.b) Freeze etch electron micrograph showing the uniform size distribution of the particles (Methods).c) Cryo electron-microscopy. The sequence of four 20 electron/Å^2 exposures shows the effects of radiation damage on granulovirus OBs. The crystalline lattice is visible only in the first image and hydrogen gas bubbles produced by radiolysis eventually reveal the virion. The increased damage rate near the virion may be a result of the greater cross-section of phosphorous atoms in the DNA and lipid membrane, as well as the higher solvent content. Scale bar 100 nm.

have sufficient contrast to allow alignment before averaging to improve signal-to-noise (Henderson 1990).

Serial femtosecond crystallography (SFX) overcomes radiation damage by utilizing ultra bright X-ray pulses that are shorter in duration than most damage processes (Neutze et al. 2000; Chapman et al. 2011; Boutet et al. 2012; Barty et al. 2012; Lomb et al. 2011; Nass et al. 2015) thereby vastly increasing the tolerable dose and thus the amount of information gathered before the specimen is destroyed. Commonly, crystals in a liquid stream are injected into the X-ray beam and each crystal diffracts for only a few femtoseconds before being destroyed (Barty et al. 2012; Stan et al. 2016). Each crystal is exposed to at most one X-ray pulse, and diffraction patterns from thousands of crystals in random orientations are combined to form a complete three-dimensional dataset. After indexing and integration, hundreds or even thousands of observations of each individual reflection from different crystals contribute to the final dataset through an averaging and modeling process resulting in a single set of reflection intensities for crystallographic structure determination (Kirian et al. 2010; White et al. 2016).

Early SFX experiments performed at LCLS showed that useful structural information could be obtained from sub-micrometer sized crystals (Chapman et al. 2011), using long-wavelength X-ray pulses (2 keV photon energy) at a dose of 700 MGy, and from micrometer sized crystals at an energy and dose similar to typical cryogenic synchrotron data collection (Boutet et al. 2012; Redecke et al. 2013) (9.4 keV, 33 MGy). Simulations (Caleman et al. 2015; Chapman et al. 2014) and experiments (Barty et al. 2012) suggest that it should be possible to outrun radiation damage with short enough pulses using much higher X-ray intensities. Calculations with pulse durations as short as 10 fs predict the possibility of high-resolution structure determination from nanocrystals at doses greater than 10 GGy, and 1 TGy doses should be tolerable with sub-femtosecond pulses (Son et al. 2011). In this context, it has been proposed that the combination of intense X-ray laser pulses and the serial data collection method could be used to obtain macromolecular structures by single molecule diffraction (Neutze et al. 2000; Aquila et al. 2015). Despite being able to measure sub-micrometer sized crystals or single molecules, most SFX experiments carried out to date, where structures have been solved to better than 3.5 Å resolution, have used protein crystals of about 1-10 μm in diameter and the full power of the XFEL source must be attenuated to avoid saturating or damaging the detector (Boutet et al. 2012; Redecke et al. 2013; Liu et al. 2014; Zhang et al. 2015; Ginn et al. 2015), at least when performing experiments in vacuum using the CS-PAD detector.

Here, we use unattenuated XFEL pulses to study native occlusion bodies (OBs) of Cydia pomonella granulovirus (CpGV), which are spheroidal ∼265 x 265 x 445 nm semi-crystalline particles that each contain a well ordered ∼0.01μm^3 protein lattice encasing a virion. Granulovirus OBs are stable, do not aggregate and are readily obtainable because the virus is commercially used in horticulture to control codling moth (*Cydia pomonella*) in apple orchards (Gebhardt et al. 2014).

Granuloviruses belong to the *Betabaculovirus* genus of the *Baculoviridae* family, a group of insect viruses characterized by viral OBs, infectious microcrystals that form within infected larvae and persist for long periods in the environment (Rohrmann 2013).

OBs contain membrane bound virus particles embedded in a crystalline lattice of viral protein molecules (Figure 1).

The OB matrix protein of *Betabaculovirus* (granulin) is homologous to that of *Alphabaculovirus* (= polyhedrin) sharing ∼60% amino acid sequence identity). Much remains to be learned about OBs, such as how the crystals grow within cells, what controls their size and shape, how the polyhedrin/granulin lattice interacts with the embedded virus particles and the envelope layer that forms the outer surface and how the OBs disassemble in the alkaline mid-gut of feeding larvae to release the infectious virus. Such studies would ideally be performed at physiological temperatures and conditions on native OBs rather than recombinantly expressed protein recrystallized into large crystals.

Atomic structures for Lepidopteran *Alphabaculovirus* polyhedrins have been determined using native polyhedral OBs purified from larvae and recombinant microcrystals prepared by expressing polyhedrin in insect cells (Coulibaly et al. 2009; Ji et al. 2009). These revealed a cubic body centered lattice with a 103 Å unit cell densely packed with polyhedrin molecules interconnected by extensive crystal contacts and intermolecular disulfide bonds. The solvent content of only 20% is amongst the lowest known for protein crystals (Trillo-Muyo et al. 2013). While these features are consistent with the extraordinary stability of OBs, these structures are incomplete, with over 30 of the 245 amino acids undefined by electron density.

Table 1. Crystallographic data collection and refinement statistics

	GV-SFX	GV-SYN
Data collection		
Temperature (K)	293	100
Wavelength (Å)	1.56	1.00
Beam size (μm^2)	1.3 x 1.3	5 x 15
Number of crystals included	82603	21
Average particle size (μm)	0.2 x 0.2 x 0.4	5 x 5 x 5
Crystalline fraction[§]	60%	100%
Flux	1 x 10^{12} photons/pulse	1 x 10^{12} photons/s
Max dose per crystal, MGy	1,300	30
Space group	*I* 2 3	*I* 2 3
Unit cell (Å)	*a* = *b* = *c* = 103.4	*a* = *b* = *c* = 102.058
Pulse duration	50 fs	-
No. collected images	1535619	77
No. hits / indexed images	487085 / 82603	-
No. total / unique reflections	77177221 / 12600	182450 / 21007
Resolution (Å)	40 – 2.00	25.51 - 1.66
Completeness (%)*	100 (99.95)	100 (100)
Multiplicity	6008 (1258)	8.7 (7.5)
SNR/<I/σ(I)>	9.57 (0.89)	7.52 (1.49)
$CC_{1/2}$	0.997 (0.297)	0.987 (0.434)
CC*	0.999 (0.677)	0.997 (0.778)
R_{split} (%)	7.91 (141.6)	-
R_{meas} (%)	-	27.52 (166.4)
Wilson B-factor (Å^2)	38.34	8.28
Refinement		
Resolution range (Å)	27.63 - 2.00 (2.07 - 2.00)	25.51 - 1.66 (1.72 - 1.66)
Reflections (refinement)	12596 (1252)	21007 (2087)
Reflections (R_{free})	1231 (115)	2131 (233)
R_{work} (%)	14.94 (38.56)	14.91 (24.69)
R_{free} (%)	18.99 (44.54)	19.22 (29.62)
No. non-hydrogen atoms	2127	2045
Protein	2030	1926
Water	97	119
B-factors (Å^2)		
Overall / Protein / Water	38.34, 36.86, 39.35	12.04, 11.52, 20.48
R.m.s bonds, Å / angles, °	0.003 / 0.58	0.007 / 0.87
Ramachandran plot (%)[†]		
Favored	96	99
Allowed	3.7	0.9
Outliers	0	0

[*] Values in parenthesis are for the highest-resolution shell
[†] From Molprobity (Chen et al. 2010)
[§] The internal virion and surface layer comprise approximately 40% of the volume of granulovirus Obs.

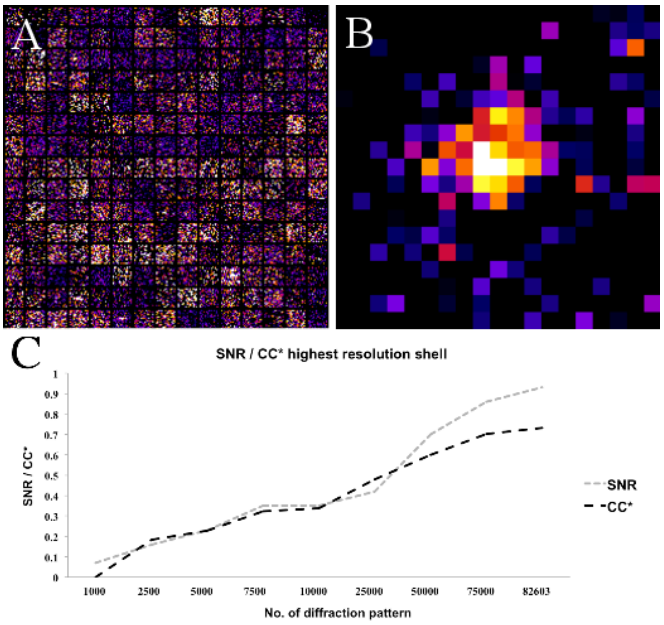

Fig. 2. Averaging of many individual Bragg peak observations over an increasing number of patterns enables high resolution Bragg peaks to rise above noise. A. 225 randomly selected single images at the predicted location of the 21 26 29 reflection (corresponding to 2.3 Å resolution). B. Averaged bragg intensity of the 21 26 29 reflection intensities from 3176 observed reflections after first rotating them into the common frame of reference of the lattice. C. Plot of overall signal-to-noise ratio and CC* metric in the 2 – 2.1 Å resolution shell as a number of indexed diffraction patterns included.

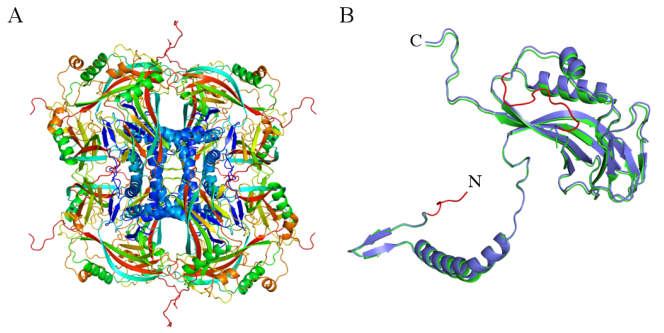

Fig. 3. The structure of the biological unit of granulin building blocks forming the crystalline OB. B. granulin monomer, the SFX structure is displayed in blue, the SYN structure in green. Regions of granulin that were present in the SFX but not in the SYN structure are highlighted in red.

Given the small size and the fact that a membrane bound virus particle occupies much of the volume, the granulin lattice in granulovirus OBs is surprisingly well ordered. Cryo-cooled granulovirus OBs pelleted in 50% ethylene glycol diffract X-rays to at least 3.5 Å (Figure 1). The size distribution of granulovirus OBs was estimated from electron micrographs (Figure S1). The granulovirus OB samples were free of aggregates and had a narrow size distribution and did not contain crystals large enough to risk damaging the detector. By modeling cryo-EM images of individual OBs (Figure 1) we estimate that about 60% of the volume of granulovirus OBs consists of crystalline polyhedrin. The crystalline part of granulovirus OBs is about one thousandth of the volume of most other crystals analyzed so far using femtosecond pulses, and provides a means to determine the practical limits of nanocrystallography using the full flux density of the X-ray FEL.

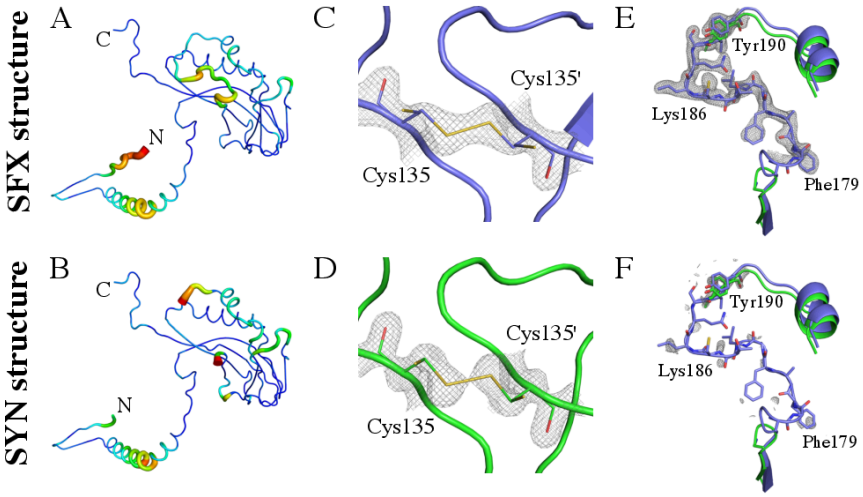

Fig. 4. Comparison of map and model of the SFX (blue) and SYN (green) structures of granulin. A. SFX and B. SYN model with color-map and thickness representing local B-factor C-D. C. SFX and D. SYN structure (blue and green respectively). Intact intermolecular disulfide bond of Cys135 between two molecules at 1.5 σ. E. SFX and F. SYN comparison of density of residues 168-201 (contoured at 1 σ). Electron density defining residues 176-190 was only found for the SFX structure. A. map derived from SFX data, B. map derived from SYN data.

RESULTS

XFEL data collection and analysis

Granulovirus OBs were produced in *Cydia pomonella* larvae, purified and injected into the XFEL beam using a gas dynamic virtual nozzle (GDVN) (DePonte et al. 2008; Weierstall et al. 2012; Gañán-Calvo 1998) at a concentration of 10^{13} particles/ml in water at a temperature of 20 °C. 1.5 million detector frames were collected (∼ 3.5 hours of data collection) during experiment L767 using the CXI microfocus instrument at LCLS, from which 487,184 crystal diffraction patterns were identified as crystal diffraction "hits" in the data processing step using Cheetah (Barty et al. 2014). We were able to obtain a complete dataset to 2 Å resolution merged from 82,603 indexed crystal diffraction patterns, with an overall R_{split} of 7.89%, demonstrating the overall internal consistency of the data (Table 1).

We observe that approximately 1% of the indexed diffraction patterns contained visible diffraction peaks extending to the corners of the detector at 1.9 Å resolution (see example in Figure S2). Crystals are randomly distributed in the 3-4 μm diameter liquid stream and the XFEL beam is focused to a of 1.3 μm FWHM spot, surrounded by a much larger, but less intense halo. As a result many of the observed patterns are probably due to particles that only partially intersect the brightest part of the XFEL beam (Galli et al. 2015). We estimate that for uniformly distributed 200×200×400 nm^3 spheroids at the concentration and flow rate used (3 - 4 μl/min) in these experiments, the expected diffraction intensity distribution would have a larger contribution from exposures with weak illumination, where the crystal is hit by the X-ray beam outside its FWHM. Although granulovirus particles have a narrow size distribution, there is a 1-2% outlier fraction of larger crystals (Figure S1) that may also have contributed to the strongest diffraction peaks at high resolution. The CrystFEL data processing suite provides an estimated resolution limit of each diffraction pattern. To exclude the possibility that high-resolution signal was entirely derived from the small population of larger crystals we identified patterns that extended beyond 2.5 Å. Excluding those patterns did not appreciably alter the structure or other merging figures of merit, leading us to conclude that the structure was indeed derived from the whole population of measured crystals and not dominated by a few strong patterns from outlier larger crystals (Figure S5).

As a result of data processing the diffraction limit was shown to extend to 2 Å resolution even though Bragg peaks in individual frames are weak and thereby difficult to observe visually at this resolution. This reflects the fact that intensity of weak reflections can be accurately measured by improving statistics through high multiplicity data sets even if the signal-to-noise in individual frames is low. Single particle imaging, both with X-rays or electrons, relies on the averaging of many weak measurements to build up signal-to-noise at high resolution even though much less than one photon per pixel may be detected in each diffraction pattern (Barty et al. 2013). Similarly for crystals, averaging data from many single-crystal diffraction patterns improves the precision of the measured structure factors at predicted Bragg peak locations beyond the resolution of detectable peaks provided the location of those peaks can be accurately predicted (Ayyer et al. 2015). Such predictions can be made by indexing the strong (usually low-resolution) peaks in the pattern as discussed below. In Figure 2 a) we show 25 × 25 pixel regions of 225 detector frames centered at the predicted location of reflection 21 26 29. This particular reflection was predicted to occur in 3176 frames in the dataset and corresponds to 2.3 Å resolution. The sum of the pixel regions from all 3176 detector frames is shown in Figure 2 b). A similar improvement occurred in all peaks in this resolution shell and Figure 2 c) shows the dependence of signal-to-noise ratio of reflections in the highest resolution shell on the number of patterns included in the dataset. Averaging over thousands of patterns (Figure S4) makes it possible to obtain a precise reflection intensity estimate by averaging a sufficient number of apparently 'invisible' Bragg peaks, with low I/σ that are not visible in individual patterns. Weak high-resolution peaks that are not identifiable in individual diffraction patterns rise out of the noise on averaging provided weak and negative intensities are included during the analysis.

Structure determination of granulin

The granulin structure was determined from the XFEL data by molecular replacement using a search model based on Wiseana nucleopolyhedrovirus (WNPV) polyhedrin (PDB ID: 3JVB) (Coulibaly et al. 2009), which shares 51.6% overall sequence identity (69.0% overall sequence similarity determined with EMBOSS Needle and the EBLOSUM62 matrix). Automatic model building followed by iterative refinement led to a final model with R_{work}/R_{free} (%) of 14.9 / 19.0 to a resolution of 2.0 Å. This model is referred to below as SFX. The SFX structure was compared to a previously determined structure (referred to as SYN) obtained by molecular replacement using the same 3JVB model and 1.66 Å resolution synchrotron data from 21 recombinant ∼5 μm CpGV granulin crystals collected at beamline X06SA at the Swiss Light Source (Supplementary Materials). The overall

SFX electron density is well defined, which allowed all but the 5 N-terminal residues to be modeled. In the cryo-cooled SYN structure, 24 of the 248 residues could not be modeled (residues 1-12 of the N-terminus and loop residues 177-188) (Figure 3 and 4).

Granulin structure

The central part of the granulin structure consists of a compact β-sandwich, with two additional α-helices H2/3. Perpendicular to the main body, granulin shows an extruding N-terminus, perpendicular to the central part, beginning with a short β-hairpin structure, followed by a long, bent α-helix H1 (Figure 3 and 4). Granulin shows a high level of structural similarity to WPNV polyhedrin (PDB ID: 3JVB) with a root-mean-square deviation of 0.634 Å across 173 Cα positions (Coulibaly et al. 2009), as expected from sequence homology. However, three flexible loop regions in the central part (146-149, 176-190, 200-207), as well as several residues at the N-terminus (6-13 and 41-43) are well defined in the XFEL granulin structure (Figure 3 and 4), but are completely absent in the WNPV polyhedrin structure.

DISCUSSION

Using submicron sized (0.01 μm^3) naturally occurring granulovirus occlusion bodies, where each particle contains only 9,000 unit cells, we have determined the room-temperature structure of granulin to 2 Å resolution using X-ray pulses that imparted a dose of up to 1.3 GGy per crystal. Despite the high dose applied and the small size of the crystals, we were able to obtain a structure of CpGV granulin with good crystallographic statistics (Table 1) revealing features that are absent in the structure of recombinant granulin determined using synchrotron radiation (Figure 3 and 4).

The final dataset contained 82,603 indexed diffraction patterns. Our analysis clearly shows that averaging over more diffraction patterns improves all figures of merit (Figure S4) with an improvement of signal-to-noise ratio increasing with the square root of the number of patterns (Figure 2 c).

The 82,603 crystals contributing to the final dataset represent a maximum probed crystalline volume of less than 900 μm^3, equivalent to a single cube-shaped crystal of $\sim$10 µm per side. Room temperature measurement from a single large crystal of equivalent volume using conventional (rotation) data collection would be severely limited in resolution and completeness, because of radiation-induced damage accumulating over many exposures of a single crystal. Exposing many small crystals to femtosecond duration X-ray pulses using serial crystallography is possible because of the tolerance of a much higher dose.

Given the high dose per crystal used in the SFX measurements, we looked for evidence of predicted local radiation damage (Lomb et al. 2011; Nass et al. 2015) in the vicinity of radiation-sensitive disulfide bonds. Figure 4 shows that the disulfide bond between the two Cys135 molecules is intact in the SFX structure, indicating that the structure is preserved even at sites of atoms that are more strongly photo-absorbing, such as sulfur, despite the high dose used in the SFX experiment.

This 2 Å structure, obtained from 0.01 μm^3 size crystals raises the question: how small is the smallest crystal that could be studied at an XFEL facility? As is apparent in Fig. 2, the averaging that was carried out did not require the Bragg peak to be visible in individual patterns, since the lattice orientation was determined from much stronger lower-angle Bragg peaks and this orientation was used to predict the location of high-resolution Bragg peaks. As the crystal size diminishes, so too does the intensity of the Bragg spots, and thus the size of the smallest crystal depends on the number of patterns that can be measured and indexed to build up a sufficiently accurate estimate of the integrated intensity. Assuming that indexing can be carried out, two other parameters that determine the ability to accurately estimate structure factors are the incident pulse flux density and the background level of the diffraction pattern (e.g. from scattering from the jet carrier liquid).

To further study this scaling we simulated diffraction patterns from a perfect granulin crystal of 9,000 unit cells (see methods) and compared the average Bragg peak counts from individual diffraction patterns with our measured data. From this comparison we conclude that the incident flux density was of the order of 10^{12} photons per $1\mu m^2$. A second simulation using the same beam parameters and zero background shows that a crystal of 123 unit cells would be sufficient to produce 'indexable' diffraction patterns using the same dose of 1.3 GGy (Figure S7), and that these data could be assembled into a complete high resolution dataset given a sufficient number of measurements. Using the same beam parameters and focus size it might therefore be possible to study crystals with 100 times smaller volume than those studied here, containing as few as 123 unit cells.

Going further, these rough estimates indicate that single-molecule imaging of a protein of molecular weight equal to one unit cell of granulin would require at least a 100-10,000 fold increase in X-ray peak flux density incident on a single unit cell to obtain continuous diffraction pattern with comparable numbers of scattered photons.

Additionally, single molecule diffraction would require reducing the background scattering to much lower levels than achieved in serial crystallography measurements with a liquid microjet. Indeed, the goal is to achieve background-free injection, where photon-counting statistics are predicted to become the dominant source of error. Additionally, detector performance must be sufficiently well controlled to detect the presence of these particles and determine their orientation. A factor of 100 increase in X-ray intensity (which would impart a dose of 100 GGy) can in principle already be delivered in the 100 nm FWHM focus chamber of the CXI instrument under ideal conditions and perfect focus. Our simple scaling calculations for this highly ordered, nearly perfect protein crystal indicate that at least a perfect focus, background-free measurements, and a noise-free detector capable of measuring single photons would be needed to perform single molecule imaging of larger molecular weight macromolecules at LCLS, properties which are yet to be satisfied both individually and simultaneously.

Under the described limitations and assumptions, it should be possible to reach the border between crystallography and the field of single particle X-ray imaging at ambient temperatures. Further increasing the flux density in parallel with improved signal-to-noise ratio of the data will open completely novel routes to structure determination of biological molecules at room temperature. The short pulses available from free electron lasers allows structure determination from crystals much smaller or more radiation sensitive than previously considered possible for study by X-ray diffraction, as well as time-resolved experiments on the ultrafast timescale (Pande et al. 2016).

METHODS

Sample preparation of *in vivo* grown granulovirus occlusion bodies

Occlusion bodies of *Cydia pomonella* granulovirus (Mexican strain) were produced by infection of fourth instar larvae of *C. pomonella* and purified as described in (Eberle et al. 2012) . In brief, OBs from raw larval homogenates or from commercial preparations were filtered through cotton wool and repeatedly washed by centrifugation (18000 g, 30 min) followed by resuspending the pellet in water. The concentrated OB suspension was then purified using a discontinuous glycerol gradient 30-80%. The OB band was collected, resupended in water and pelleted again (3200 xg, 45 min). If the required purity was not achieved, the glycerol gradient purification was repeated once. The collected OBs were three times washed by centrifugation (3200 g, 45 min) followed by resuspending the pellet in water. The OB suspension was stored at -18°C until usage.

Size estimate of granulovirus occlusion bodies

Close inspection of electron micrographs like Figure 1 revealed occasional larger particles with a similar shape and rare elongated particles with similar diameter but up to five times longer than average granulovirus par-

ticles. To investigate the possibility that these larger particles may dominate the SFX dataset we needed unbiased size distribution statistics. We developed a system to dry particles from a thin liquid layer on glow discharged carbon coated electron microscope grids so that most of the particles are oriented with the long axis approximately parallel to the grid surface and made low magnification images like that shown in Figure S1. We used ImageJ (https://imagej.nih.gov/ij/) to trace the outlines of all the particles in these images and Python scripts to estimate particle crystal volumes from fitted ellipse outlines.

Powder Diffraction

Granulovirus in water was pelleted using a bench top centrifuge and the pellet resuspended in an equal volume of ethylene glycol. The resultant slurry was applied to a MiteGen grid and allowed to dry before cooling by insertion into a 100K cryo-stream. The grid was exposed for 15s to a flux of $\sim$1.56 x 10^{12} 0.97Å photons/s focused to 80 x 40 microns using the X06SA beamline at the Swiss Light Source synchrotron. The detector used was a Pilatus 6M at a camera length of 630mm. Repeated exposures did not show evidence of radiation damage.

Data collection for synchrotron structure

Recombinantly grown granulin *in vivo* crystals in 50% ethylene glycol were applied to Mitegen 400/50 mesh-grids and cooled in liquid nitrogen before transfer to the cryo-stream at the MD2 microdiffractometer (Maatel) used for data collection.

Diffraction data for recombinant CpGV granulin at 100 K were collected in 1° rotations from beamline PXI (X06SA) of the Swiss Light Source, Switzerland, with synchrotron radiation (10^{12} photons/s, beam size of 5 x 15 μm^2 FWHM) at a wavelength of 1.0 Å (0.02 % bandwidth) using a Mar225 CCD detector (MarResearch). The data were indexed and integrated with *XDS* (Kabsch 2010), scaled and merged with *AIMLESS* (Evans & Murshudov 2013), and converted to structure factor amplitudes with *CTRUNCATE* (Winn et al. 2011). Merging statistics were calculated with *phenix.merging_statistics* (Adams et al. 2010). The structure determined and refined in the same manner as the XFEL structure.

The high-resolution limit of 1.66 Å for the SYN structure, as well as the 2 Å for the SFX structure was determined by paired-refinement and correlation coefficient analysis, ensuring CC_{work} or CC_{free} did not exceed CC* across the full resolution range (Karplus & Diederichs 2012).

Data collection by serial femtosecond crystallography

Serial femtosecond crystallography experiments with CpGV OBs were carried out at the CXI instrument of the Linac Coherent Light Source (LCLS – beamtime ID: L767), at SLAC National Accelerator Laboratory (Menlo Park, USA). The LCLS X-ray beam with an estimated pulse duration of 50 fs and a photon energy of 7.9 keV was focused using X-ray optics in a Kirkpatrick-Baez geometry to a beam size of about 1.3 × 1.3 μm^2 FWHM. The average pulse energy was 2.7 mJ, and the pulse repetition rate was 120 Hz. Granulovirus OBs were injected as a liquid suspension into the X-ray interaction region as a liquid jet formed by a gas-focussed nozzle (Gañán-Calvo 1998; DePonte et al. 2008; Weierstall et al. 2012). In this system the liquid suspension was pushed through a 50 µm inner diameter capillary to subsequently form a jet of about 3-4 µm diameter at the interaction region by the action of a co-axially flowing helium stream. The flow rate of the sample was 10-15 µl/min. Granulovirus OB settling was prevented by the use of an anti-settling sample-delivery instrument (Lomb et al. 2012) equipped with syringe-temperature control (set to 20°C) for injection. The velocity of the jet was sufficient to fully replenish the sample between X-ray pulses. Diffraction data were collected on a Cornell-SLAC Pixel Array Detector (CSPAD) (Hart et al. 2012), a multi-panel detector consisting of 64 individual tiles. The detector was set to a distance of 130 mm, giving a maximum diffraction resolution of 1.88 Å at its corner.

SFX data analysis

Individual 'hits', defined as diffraction patterns containing more than 20 Bragg peaks with a signal-to-noise larger than 6, were identified using the software Cheetah (Barty et al. 2014). The average hit rate fluctuated between 10 and 60% during the experiment, for a total of 487,085 frames containing crystal diffraction. Some of them showed diffraction to the corner of the detector (see Fig. S2 as an example). 82,603 of these patterns were indexed using CrystFEL (White et al. 2012) (version 0.5.2+93d0472).

Granulin crystals belong to space group *I*23, a space group with arbitrary choice of two different orientations of the unit-cell axes that are compatible with a particular set of Bragg peak locations in the diffraction pattern. This ambiguity can be overcome by examining the relative intensities of the Bragg peaks. In SFX, this is complicated by the unknown degree of partiality of the recorded Bragg peaks and other sources of error in the individual intensity measurements. The indexing ambiguity was resolved using the CrystFEL "ambigator" tool (White et al. 2016), which uses an algorithm related to that devised by Brehm and Diederichs (Brehm & Diederichs 2013), comparing the average Pearson correlation between a single diffraction pattern and every other in the dataset. As expected, approximately half of the 82,603 frames (42,143) were initially indexed inconsistently and were corrected according to the average correlation coefficients (Figure S4). The Bragg peaks were integrated using the Monte Carlo approach as implemented in CrystFEL(White et al. 2013). The averaged intensities were converted to structure factor amplitudes by Xdsconv (Kabsch 2010). Table 1 and Figure S5 show the quality of the final dataset.

Molecular replacement was performed using Phaser(McCoy et al. 2007) using a search model of a homologous polyhedrin with 52.5% sequence identity (PDB ID: 3JVB). The initial model was subjected to automatic model building using phenix.autobuild(Terwilliger et al. 2008), followed by iterative cycles of manual model (re-)building using Coot(Emsley et al. 2010) and refinement with phenix.refine(Afonine et al. 2012). Positive difference density in five previously undefined regions was identified, which enabled us to model a loop region of 15 residues (176 – 190; KKGFAFPLTCLQSVY), residues 41 – 43 (IRE), 146 – 149 (AHPD) and 200 – 207 (DVLWPYFY) and additional 8 residues (6 – 13; SLRYSRHD) at the N-terminus. All residues of the granulin amino acid sequence but the first five at the N-terminus were defined by electron density and were included in the final model. PISA(Krissinel & Henrick 2007) was used to calculate the biological assembly. Electron density maps were generated with Phenix (Afonine et al. 2012) and figures were generated with Pymol.

The high-resolution limit of (2 Å) was determined by paired-refinement and correlation coefficient analysis, ensuring CC_{work} or CC_{free} did not exceed CC* across the full resolution range (Karplus & Diederichs 2012), as for the synchrotron structure.

Cryo-electron microscopy

Granulovirus OBs at a concentration of $1x10^{11}$ occlusion bodies/ml (OB/ml) were adsorbed to glow discharged lacey-carbon on 300 mesh copper grids (Ted Pella). After 30 seconds incubation, the grid was blotted for 2 seconds using a Vitrobot Mark IV (FEI Company) then immediately plunge frozen into liquid ethane. The sample was imaged on a Philips CM200 (FEI Company) transmission electron microscope at an accelerating voltage of 200 kV equipped with a TemCam-F416 (TVIPS) CCD detector. Data was collected at a nominal magnification of 50,000 x corresponding to a 2.14 Å pixel size and at a defocus of around -2.5 µm.

Simulations

Diffraction simulations were performed using the MOLTRANS software package, developed by E. Weckert (DESY – available upon request). MOLTRANS calculates the scattered amplitude from each atom of the structure arriving at each pixel of the detector. The complex-valued wavefield scattered from all atoms in one individual unit cell is first calculated, from which the wavefield from the entire crystal is determined through a coherent summation that takes the relative positions of different unit cells in the crystal into account. The intensity at the detector has to be scaled according to the incident pulse flux density and the detection efficiency. The incident flux density depends on the pulse energy, the beamline transmission from the position of the pulse energy monitor (60%, depending on KB mirror efficiency), and the area of the beam at the sample. Both the pulse energy and the overlap of the beam with the sample varies from shot to shot. Therefore simulations were compared to the brightest patterns in order to find the photon flux in the center of the beam. At the same time the patterns with very narrow Bragg peaks, corresponding to the particles with diameter greater than 400nm, were excluded from this analysis to avoid attributing the effect of the brightest patterns to the small fraction of bigger granulovirus OBs. Such analysis suggests that the X-ray flux density was close to 10^{12} photons/μm^2.

These simulations deliberately omit background scattering and assume an ideal overlap of 10^{12} photons in a 1 μm^2 top-hat X-ray focus with the sample during the exposure. Two different simulations were performed: 1) one dataset calculated for a crystal of 9,000 unit cells, which was used for comparison with the collected data (Figure S7, a), and 2) one dataset calculated for a very small crystal of spherical shape consisting of 123 unit cells (Figure S7, b).

For dataset 2, we generated more than 5000 diffraction patterns from these crystals in random orientation and indexed those (4,350 indexed from 5,043). The fact that indexing was successfully achieved on individual diffraction patterns indicates that such small crystals should be of sufficient size for high-resolution structure determination, given noise-free and background-free detection of the X-ray diffraction patterns.

Adams, P.D. et al., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography*, 66(Pt 2), pp.213–221.

Afonine, P.V. et al., 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst (2012). D68, 352-367 [doi:10.1107/S0907444912001308]*, pp.1–16.

Aquila, A. et al., 2015. The linac coherent light source single particle imaging road map. *Structural Dynamics*, 2(4), pp.041701–13.

Ayyer, K. et al., 2015. Determination of crystallographic intensities from sparse data. *IUCrJ*, 2(1), pp.29–34.

Barty, A. et al., 2014. Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *Journal of Applied Crystallography*, 47(Pt 3), pp.1118–1131.

Barty, A. et al., 2012. Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nature Photonics*, 6(1), pp.35–40.

Barty, A., Küpper, J. & Chapman, H.N., 2013. Molecular imaging using X-ray free-electron lasers. *Annual Review of Physical Chemistry*, 64, pp.415–435.

Boutet, S. et al., 2012. High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science*, 337(6092), pp.362–364. Available at: http://gateway.webofknowledge.com/gateway/Gateway.cgi?GWVersion=2&SrcAuth=mekentosj&SrcApp=Papers&DestLinkType=FullRecord&DestApp=WOS&KeyUT=000306542600057.
Brehm, W. & Diederichs, K., 2013. Breaking the indexing ambiguity in serial crystallography. *Acta Cryst (2014). D70, 101-109 [doi:10.1107/S1399004713025431]*, pp.1–9.
Caleman, C. et al., 2015. Ultrafast self-gating Bragg diffraction of exploding nanocrystals in an X-ray laser. *Optics Express*, 23(2), pp.1213–1231.
Chapman, H.N. et al., 2011. Femtosecond X-ray protein nanocrystallography. *Nature*, 470(7332), pp.73–77.
Chapman, H.N., Caleman, C. & Timneanu, N., 2014. Diffraction before destruction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1647), pp.20130313–20130313.
Chen, V.B. et al., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*, 66(1), pp.12–21.
Coulibaly, F. et al., 2009. The atomic structure of baculovirus polyhedra reveals the independent emergence of infectious crystals in DNA and RNA viruses. *Proceedings of the National Academy of Sciences*, 106(52), pp.22205–22210.
DePonte, D.P. et al., 2008. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics*, 41(19), p.195505.
Eberle, K.E. et al., 2012. Chapter II Basic techniques in insect virology. In *Manual of Techniques in Invertebrate Pathology*. Elsevier, pp. 15–74.
Emsley, P. et al., 2010. Features and development of Coot. *Acta Crystallographica Section D Biological Crystallography*, 66(4), pp.486–501. Available at: http://scripts.iucr.org/cgi-bin/paper?ba5144.
Evans, P.R. & Murshudov, G.N., 2013. How good are my data and what is the resolution? *Acta Crystallographica Section D Biological Crystallography*, 69(Pt 7), pp.1204–1214.
Galli, L., Metcalf, P. & Chapman, H.N., 2015. Implications of the focal beam profile in serial femtosecond crystallography L. Juha, S. Bajt, & R. London, eds. *SPIE Optics+ …*, 9511, pp.95110H–95110H–7.
Gañán-Calvo, A.M., 1998. Generation of Steady Liquid Microthreads and Micron-Sized Monodisperse Sprays in Gas Streams. *Physical Review Letters*, 80(2), pp.285–288.
Garman, E.F., 2010. Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta Crystallographica Section D Biological Crystallography*, 66(4), pp.339–351.
Gebhardt, M.M. et al., 2014. Baculovirus resistance in codling moth is virus isolate-dependent and the consequence of a mutation in viral gene pe38. *Proceedings of the National Academy of Sciences*, 111(44), pp.15711–15716.
Ginn, H.M. et al., 2015. Structure of CPV17 polyhedrin determined by the improved analysis of serial femtosecond crystallographic data. *Nature Communications*, 6, p.6435.
Hart, P. et al., 2012. The CSPAD megapixel x-ray camera at LCLS. In S. R. Stock, ed. SPIE Optical Engineering + Applications. SPIE, pp. 85040C–85040C–11.
Henderson, R., 1990. Cryo-Protection of Protein Crystals against Radiation Damage in Electron and X-Ray Diffraction on JSTOR. In Proceedings: Biological Sciences.
Henderson, R., 1995. The Potential and Limitations of Neutrons, Electrons and X-Rays for Atomic-Resolution Microscopy of Unstained Biological Molecules. *Quarterly Reviews of Biophysics*, 28(2), pp.171–193.
Howells, M.R. et al., 2009. An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. *Journal of Electron Spectroscopy and Related Phenomena*, 170, pp.4–12.
Ji, X. et al., 2009. How baculovirus polyhedra fit square pegs into round holes to robustly package viruses. *The EMBO Journal*, 29(2), pp.1–10.
Kabsch, W., 2010. XDS. *Acta Crystallographica Section D Biological Crystallography*, 66(2), pp.125–132.
Karplus, P.A. & Diederichs, K., 2012. Linking Crystallographic Model and Data Quality. *Science*, 336(6084), pp.1030–1033.
Kirian, R.A. et al., 2010. Femtosecond protein nanocrystallography-data analysis methods. *Optics Express*, 18(6), pp.5713–5723.
Krissinel, E. & Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, 372(3), pp.774–797.
Liu, W. et al., 2014. Femtosecond crystallography of membrane proteins in the lipidic cubic phase. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1647), pp.20130314–20130314.
Lomb, L. et al., 2012. An anti-settling sample delivery instrument for serial femtosecond crystallography. *Journal of Applied Crystallography*, 45(4), pp.674–678.
Lomb, L. et al., 2011. Radiation damage in protein serial femtosecond crystallography using an x-ray free-electron laser. *Physical Review B*, 84(21), p.214111.
McCoy, A.J. et al., 2007. Phaser crystallographic software. *Journal of Applied Crystallography*, 40(4), pp.658–674.
Nass, K. et al., 2015. Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *Journal of Synchrotron Radiation*, 22(Pt 2), pp.225–238.
Neutze, R. et al., 2000. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature*, 406(6797), pp.752–757.
Redecke, L. et al., 2013. Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser. *Science*, 339(6116), pp.227–230.
Rohrmann, G.F., 2013. Baculovirus Molecular Biology.
Stan, C.A. et al., 2016. Liquid explosions induced by X-ray laser pulses. *Nature Physics*, pp.1–7.
Son, S.-K.S., Young, L. & Santra, R., 2011. Impact of hollow-atom formation on coherent x-ray scattering at high intensity. *Physical Review A*, 83(3), p.033402.
Terwilliger, T.C. et al., 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica Section D Biological Crystallography*, 64(1), pp.61–69.
Trillo-Muyo, S. et al., 2013. Ultratight crystal packing of a 10 kDa protein. *Acta Crystallographica Section D Biological Crystallography*, 69(Pt 3), pp.464–470.
Weierstall, U., Spence, J.C.H. & Doak, R.B., 2012. Injector for scattering measurements on fully solvated biospecies. *Review of Scientific Instruments*, 83(3).
White, T.A. et al., 2013. Crystallographic data processing for free-electron laser sources. *Acta Crystallographica Section D Biological Crystallography*, 69(Pt 7), pp.1231–1240.
White, T.A. et al., 2012. CrystFEL: a software suite for snapshot serial crystallography. *Journal of Applied Crystallography*, 45(2), pp.335–341.
White, T.A. et al., 2016. Recent developments in CrystFEL. *Journal of Applied Crystallography*, 49(Pt 2), pp.680–689.
Winn, M.D. et al., 2011. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography*, 67(Pt 4), pp.235–242.
Zhang, H. et al., 2015. Structure of the Angiotensin Receptor Revealed by Serial Femtosecond Crystallography. *Cell*, 161(4), pp.833–844.

End Notes

ACCESSION CODES

The atomic coordinates and structure factors have been deposited in Protein Data Bank (PDB) under accession code 5G0Z for the SFX structure and 5G3X for the SYN structure.

ACKNOWLEDGEMENTS

We thank Ilme Schlichting, Thomas Barends, Nadrian Seeman, Jens Birktoft, Jennifer Padilla, Nam Nguyen, Michael J. Bogan, Guangmei Huang, Adrian Turner, Chitra Rajendran, Martin Middleditch, James Dickson, Nobuhiro Morone and John Heuser for their assistance in preparation and during the experiment. We want to specially thank Birgit Weihrauch for her support with sample preparation. Data for this paper was collected during LCLS experiment L767 of Ned Seeman. C.G. kindly thanks the PIER Helmholtz Graduate School, as well as the Helmholtz Association for financial support. This work was supported by the STC Program of the National Science Foundation through BioXFEL under Agreement No. 1231306, the National Institutes of Health Femtosecond Award 617095583. PM thanks The Royal Society of NZ Marsden Grant UOA1221 for financial support. Use of the Linac Coherent Light Source (LCLS), 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. ***AUTHOR CONTRIBUTIONS*** CG, OY, RDB, PM, AB and HNC prepared the manuscript. RDB, EC, S-MY and PM collected synchrotron data of recombinant granulin microcrystals and performed structure refinement. S-MY produced the granulin microcrystals. JJ produced and purified native CpGV nanocrystals. PM did the powder diffraction experiment with native CpGV nanocrystals. FS pre-characterized native CpGV nanocrystals and performed initial data collection. KRB, DD, RB, DW, DRJ, SBasu, RLS, UW, SBotha and RBD participated in sample delivery and together with CG, AA, TB, TAW, SB, GJW, MM, SBasu, SBoutet performed experimental data collection. CG performed data processing and analysis of SFX data. DO and RDB performed structure refinement of the SFX structure and analyzed the structure with PM. OY was responsible for simulation of SFX diffraction pattern and the analysis thereof. KNG performed cryo-EM data collection and processing. PM performed size distribution analysis. ***COMPETING FINANCIAL INTERESTS*** The authors declare no competing financial interests.

We thank Ilme Schlichting, Thomas Barends, Nadrian Seeman, Jens Birktoft, Jennifer Padilla, Nam Nguyen, Michael J. Bogan, Guangmei Huang, Adrian Turner, Chitra Rajendran, Martin Middleditch, James Dickson, Nobuhiro Morone and John Heuser for their assistance in preparation and during the experiment. We want to specially thank Birgit Weihrauch for her support with sample preparation. Data for this paper was collected during LCLS experiment L767 of Ned Seeman. C.G. kindly thanks the PIER Helmholtz Graduate School, as well as the Helmholtz Association for financial support. This work was supported by the STC Program of the National Science Foundation through BioXFEL under Agreement No. 1231306, the National Institutes of Health Femtosecond Award 617095583. PM thanks The Royal Society of NZ Marsden Grant UOA1221 for financial support. Use of the Linac Coherent Light Source (LCLS), 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.

AUTHOR CONTRIBUTIONS

CG, OY, RDB, PM, AB and HNC prepared the manuscript. RDB, EC, S-MY and PM collected synchrotron data of recombinant granulin microcrystals and performed structure refinement. S-MY produced the granulin microcrystals. JJ produced and purified native CpGV nanocrystals. PM did the powder diffraction experiment with native CpGV nanocrystals. FS pre-characterized native CpGV nanocrystals and performed initial data collection. KRB, DD, RB, DW, DRJ, SBasu, RLS, UW, SBotha and RBD participated in sample delivery and together with CG, AA, TB, TAW, SB, GJW, MM, SBasu, SBoutet performed experimental data collection. CG performed data processing and analysis of SFX data. DO and RDB performed structure refinement of the SFX structure and analyzed the structure with PM. OY was responsible for simulation of SFX diffraction pattern and the analysis thereof. KNG performed cryo-EM data collection and processing. PM performed size distribution analysis.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.