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

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

[a]Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany; [b]School of Biological Sciences, The University of Auckland, Auckland 1142, New Zealand; [c]European XFEL GmbH, Hamburg 22761, Germany; [d]School of Molecular Sciences, Arizona State University, Tempe, AZ 85287-1604; [e]Biodesign Center for Applied Structural Discovery, Arizona State University, Tempe, AZ 85287-5001; [f]Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Heidelberg 69120, Germany; [g]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025; [h]SLAC National Accelerator Laboratory, Menlo Park, CA 94025; [i]Department of Physics, Arizona State University, Tempe, AZ 85287; [j]Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University Basel, Basel CH-4058, Switzerland; [k]Institute for Biological Control, Julius Kuehn Institute (JKI), 64287 Darmstadt, Germany; [l]Department of Physics, University of Hamburg, Hamburg 20355, Germany; and [m]Center for Ultrafast Imaging, University of Hamburg, Hamburg 20355, Germany

Edited by Gregory A. Petsko, Weill Cornell Medical College, New York, NY, and approved January 12, 2017 (received for review June 15, 2016)

To understand how molecules function in biological systems, new methods are required to obtain atomic resolution structures from biological material under physiological conditions. Intense femtosecond-duration pulses from X-ray free-electron lasers (XFELs) can outrun most damage processes, vastly increasing the tolerable dose before the specimen is destroyed. This in turn allows structure determination from crystals much smaller and more radiation sensitive than previously considered possible, allowing data collection from room temperature structures and avoiding structural changes due to cooling. Regardless, high-resolution structures obtained from XFEL data mostly use crystals far larger than 1 μm^3 in volume, whereas 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 (LCLS) 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, cryocooled granulovirus 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 | bioimaging | SFX

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 (1). The tolerable dose, and therefore damage, is a function of spatial resolution, with fine detail being destroyed first (2, 3). 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 150 kGy (gray = J·kg^{-1} absorbed energy) causes rapid decay of biological samples (4). 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 cryocooled protein crystals larger than about 100 μm^3 (3, 5, 6). Similar dose limits apply in single particle cryo-electron microscopy, where individual images must have sufficient contrast to allow alignment before averaging to improve signal-to-noise (7, 8).

Serial femtosecond crystallography (SFX) overcomes radiation damage by using ultrabright X-ray pulses that are shorter in

Author contributions: C.G., J.A.J., and P.M. designed research; C.G., D.O., O.Y., R.D.B., F.S., E.C., S.-M.Y., A.A., S. Basu, R.B., K.R.B., S. Botha, S. Boutet, D.P.D., R.B.D., R.F., L.G., I.G., D.R.J., C.K., L.L., M.M., K.N., K.R., R.L.S., D.W., U.W., T.A.W., G.J.W., N.A.Z., P.F., J.C.H.S., K.N.G., J.A.J., P.M., and H.N.C. performed research; J.A.J., P.M., and H.N.C. contributed new reagents/analytic tools; C.G., D.O., O.Y., R.D.B., F.S., E.C., L.G., K.N.G., and P.M. analyzed data; and C.G., D.O., O.Y., R.D.B., K.N.G., J.A.J., P.M., A.B., and H.N.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org [PDB ID code 5G0Z (SFX structure) and 5G3X (SYN structure)].

[1]Present address: MRC Laboratory of Molecular Biology, CB2 0QH Cambridge, United Kingdom.

[2]Present address: Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland.

[3]Present address: Swiss Light Source, Paul Scherrer Institut, 5232 Villigen-PSI, Switzerland.

[4]Present address: Institute for Biochemistry and Molecular Biology, University of Hamburg, 20146 Hamburg, Germany.

[5]Present address: Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Heidelberg 69120, Germany.

[6]Present address: Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, WI 53211.

[7]Present address: National Science Foundation BioXFEL Science and Technology Center, Buffalo, NY 14203.

[8]Present address: SwissFEL, Paul Scherrer Institute, 5232 Villigen-PSI, Switzerland.

[9]Present address: iHuman Institute, ShanghaiTech University, Shanghai 201210, China.

[10]Present address: NSLS-II, Brookhaven National Laboratory, Upton, NY 11973.

[11]To whom correspondence may be addressed. Email: henry.chapman@desy.de or peter.metcalf@auckland.ac.nz.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1609243114/-/DCSupplemental.

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 (XFEL) 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, whereas femtosecond X-ray pulses from free-electron lasers enable much higher tolerable doses, extracting more signal per molecule, allowing the study of submicrometer 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 levels obtained in this experiment indicates that structure determination from even smaller protein crystals could be possible.

duration than most damage processes (9–14), 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 (12, 15). 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 3D 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 (16, 17).

Early SFX experiments performed at the Linac Coherent Light Source (LCLS) showed that useful structural information could be obtained from submicrometer-sized crystals (10) 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 cryogenic synchrotron data collection (11, 18, 19) (9.4 keV, 33 MGy). Simulations (20, 21) and experiments (12–14) 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 subfemtosecond pulses (22). 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 (9, 23). Despite being able to measure submicrometer-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 (11, 18, 19, 24, 25), at least when performing experiments in vacuum using the Cornell-SLAC Pixel Array Detector (CSPAD) detector.

Here, we use unattenuated XFEL pulses to study native occlusion bodies (OBs) of *Cydia pomonella* granulovirus (CpGV), which are spheroidal ∼265 × 265 × 445 nm semicrystalline particles that each contain a well ordered ∼0.01 μm^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 (*C. pomonella*) in orchards (26). 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 (27).

OBs contain membrane-bound virus particles embedded in a crystalline lattice of viral protein molecules (Fig. 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 midgut 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 (28, 29). 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

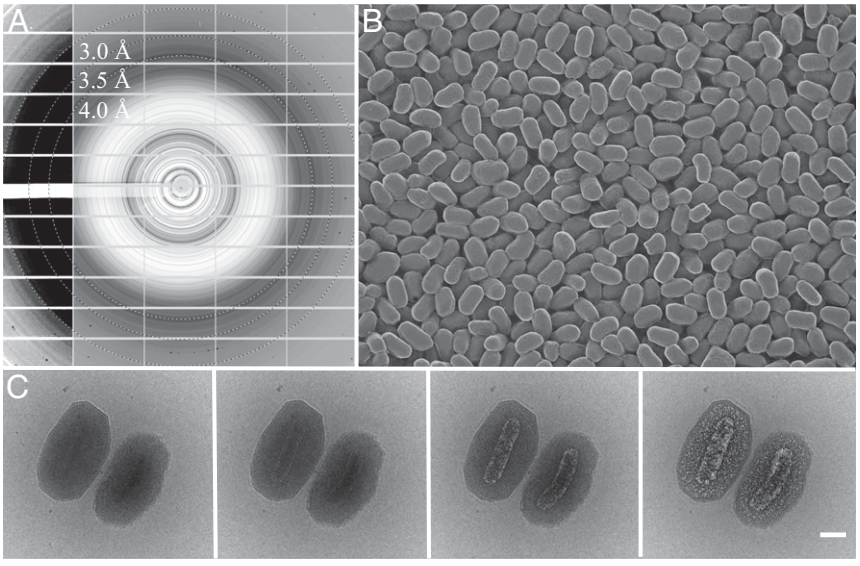

Fig. 1. Granulovirus OBs contain a single virion surrounded by a crystalline protein layer that diffracts to high resolution. (*A*) Powder X-ray diffraction from a pellet of granulovirus OBs at 100 K (*Materials and Methods*). Protein diffraction rings extend to a resolution between 3 and 3.5 Å. The 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 3 Å. (*B*) Freeze etch electron micrograph showing the uniform size distribution of the particles (*Materials and Methods*). (*C*) Cryo-EM. The sequence of four 20 e/$Å^2$ 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. (Scale bar, 100 nm.)

of only 20% is among the lowest known for protein crystals (30). Although 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.

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. Cryocooled granulovirus OBs pelleted in 50% (vol/vol) ethylene glycol diffract X-rays to at least 3.5 Å (Fig. 1). The size distribution of granulovirus OBs was estimated from electron micrographs (Fig. 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 (Fig. 1), we estimate that about 60% of the volume of granulovirus OBs consists of crystalline granulin. The crystalline part of granulovirus OBs is about 1,000th 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.

Results

XFEL Data Collection and Analysis. Granulovirus OBs were produced in *C. pomonella* larvae, purified, and injected into the XFEL beam using a gas dynamic virtual nozzle (GDVN) (31–33) at a concentration of 10^{13} particles per mL in water at a temperature of 20 °C. A total of 1.5 million detector frames were collected (~3.5 h 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 (34). We obtained 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 ~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 Fig. S2). Crystals are randomly distributed in the 3–4-μm diameter liquid stream, and the XFEL beam is focused to a 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. 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 (Fig. 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 the 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 (Table S1).

Although the remaining ~99% of the indexed diffraction patterns showed little evidence of diffraction beyond 3 Å resolution, averaged data from thousands of patterns revealed that granulovirus OBs diffract to at least 2 Å resolution, limited by detector distance. In Fig. 2*A* we show 25 × 25 pixel regions of 225 detector frames centered at the predicted location of reflection 21 26 29 after first accounting for an indexing ambiguity (see Fig. S3). This reflection was predicted to occur in 3,176 frames in the

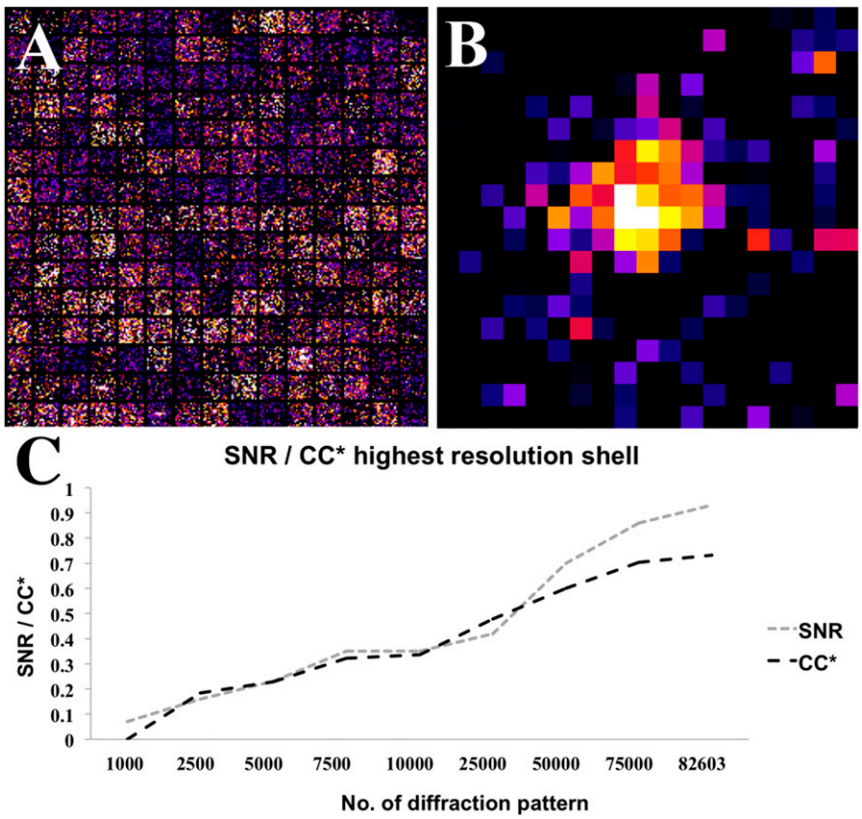

Fig. 2. Averaging thousands of reflections significantly improved overall resolution estimate. (*A*) Two hundred and twenty-five randomly selected single images at the predicted location of the 21 26 29 reflection (corresponding to 2.3 Å resolution). (*B*) Averaged Bragg intensity of the 21 26 29 reflection intensities from 3,176 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 against the number of indexed diffraction patterns included.

Table 1. Crystallographic data collection and refinement statistics

Metric	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	82,603	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 per pulse	1 x 10^{12} photons per 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	1,535,619	77
No. hits/indexed images	487,085/82,603	—
No. total/unique reflections	77,177,221/12,600	182,450/21,007
Resolution, Å	40–2.00	25.51–1.66
Completeness, %[†]	100 (99.95)	100 (100)
Multiplicity	6,008 (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	12,596 (1,252)	21,007 (2,087)
Reflections, R_{free}	1,231 (115)	2,131 (233)
R_{work}, %	14.94 (38.56)	14.91 (24.69)
R_{free}, %	18.99 (44.54)	19.22 (29.62)
No. nonhydrogen atoms	2,127	2,045
Protein	2,030	1,926
Water	97	119
B-factors, Å^2		
Overall/protein/water	39.07, 39.05, 39.35	12.04, 11.52, 20.48
R.m.s bonds, Å/angles, °	0.004/0.71	0.007/0.87
Ramachandran plot, %[‡]		
Favored	96	99
Allowed	3.7	0.9
Outliers	0	0

*The internal virion and surface layer comprise ~40% of the volume of granulovirus OBs.
[†]Values in parentheses are for the highest resolution shell.
[‡]From Molprobity (35).

dataset and corresponds to 2.3 Å resolution. The sum of the pixel regions from all 3,176 detector frames is shown in Fig. 2*B*. A similar improvement occurred in all peaks in this resolution shell, and Fig. 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 (Fig. S4) significantly improved data quality, revealing Bragg peaks not visible in individual patterns.

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 [Protein Data Bank (PDB) ID code 3JVB] (28), 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 with 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 (*SI Materials and Methods*). The overall SFX electron density is well defined, which allowed all but the five N-terminal residues to be modeled. In the cryocooled SYN structure, 24 of the 248 residues could not be modeled (residues 1–12 of the N terminus and loop residues 177–188) (Figs. 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 (Figs. 3 and 4). Granulin shows a high level of structural similarity to WNPV polyhedrin (PDB ID code 3JVB), with a root-mean-square deviation of 0.634 Å across 173 Cα positions (28), 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, 41–43)

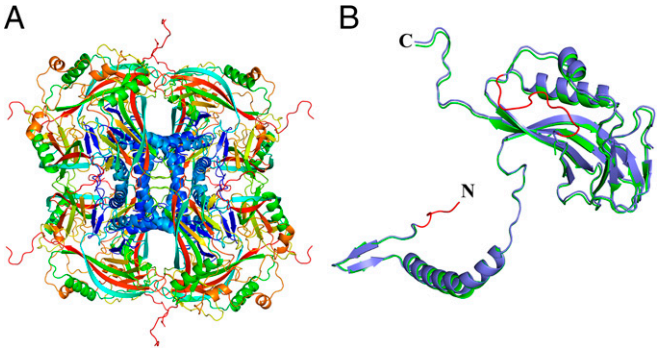

Fig. 3. (*A*) The structure of the biological unit of granulin building blocks forming the crystalline OB. (*B*) Granulin monomer. The SFX structure is displayed in blue and the SYN structure in green. Regions of granulin that were present in the SFX but not in the SYN structure are highlighted in red.

are well defined in the XFEL granulin structure (Figs. 3 and 4) but are completely absent in the WNPV polyhedrin structure.

Discussion

Using submicron-sized (0.01 μm^3) naturally occurring granulovirus OBs, 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 obtained a structure of CpGV granulin with good crystallographic statistics (Table 1), revealing features that are absent in the structure of recombinant granulin (Figs. 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 (Fig. S4), with an improvement of signal-to-noise ratio increasing with the square root of the number of patterns (Fig. 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 ~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 (13, 14) in the vicinity of radiation-sensitive disulfide bonds. Fig. 4 shows that the disulfide bond between the two Cys135 molecules is largely intact in the SFX structure (occupancy 0.68), indicating that the structure is preserved even at sites of atoms that are 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 following question: How small is the smallest crystal that could be studied at an XFEL facility? 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 (*Materials and 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 μm^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 (Fig. S5) 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

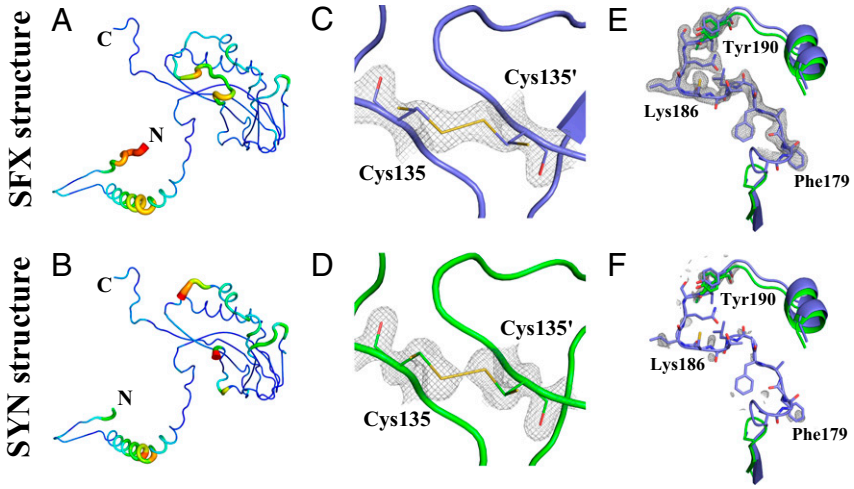

Fig. 4. Comparison of the final $2F_o$-F_c electron density map and model for the SFX (blue) and SYN (green) structures of granulin. (*A*) SFX and (*B*) SYN model with color-map and thickness corresponding to local B-factor. (*C* and *D*) Close-up of Cys135 in (*C*) the SFX (blue) and (*D*) the SYN structure and (green) embedded in its electron (gray mesh, contoured 1.5 σ). Cys135 adopts two distinct side chain conformations (shown as sticks): The major conformer forms an intermolecular disulfide bond with its counterpart in a neighboring molecule and has a refined occupancy of 0.68 and 0.87 for the SFX and SYN structure, respectively. (*E*) SFX and (*F*) SYN comparison of electron density for residues 168–201 (contoured at 1 σ). Electron density defining residues 176–190 was only found for the SFX structure.

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 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 that 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 allow 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.

Materials and Methods

A more detailed description of the sample preparation and diffraction experiments can be found in *SI Materials and Methods*.

Synchrotron experiments of recombinantly grown CpGV granulin in vivo crystals were carried out at the microfocus beamline (X06SA) Swiss Light Source, where a total of 21 crystals (77 frames, 1° oscillation per frame) were merged to a final dataset to 1.66 Å resolution.

SFX experiments were carried out at the CXI instrument of LCLS at the SLAC National Accelerator Laboratory. With a repetition rate of 120 Hz, a total of 487,085 diffraction patterns were collected with a CSPAD detector, of which finally 82,603 were indexed using the CrystFEL software package (17). A final dataset was merged with an estimated resolution of 2 Å.

ACKNOWLEDGMENTS. 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 especially thank Birgit Weihrauch for her support with sample preparation. Data for this paper were collected during LCLS experiment L767 of Nadrian Seeman. C.G. kindly thanks the Partnership for Innovation, Education, and Research (PIER) Helmholtz Graduate School as well as the Helmholtz Association for financial support. This work was supported by the Science and Technology Center Program of the National Science Foundation through BioXFEL under Agreement 1231306, the National Institutes of Health Award R01GM095583. P.M. thanks The Royal Society of NZ Marsden Grant UOA1221 for financial support. Use of the LCLS, SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract DE-AC02-76SF00515.

1. Henderson R (1995) The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys* 28(2):171–193.
2. Howells MR, et al. (2009) An assessment of the resolution limitation due to radiation-damage in x-ray diffraction microscopy. *J Electron Spectrosc Relat Phenom* 170(1-3):4–12.
3. Holton JM (2009) A beginner's guide to radiation damage. *J Synchrotron Radiat* 16(Pt 2):133–142.
4. Warkentin M, Hopkins JB, Haber JB, Blaha G, Thorne RE (2014) Temperature-dependent radiation sensitivity and order of 70S ribosome crystals. *Acta Crystallogr D Biol Crystallogr* 70(Pt 11):2890–2896.
5. Henderson R (1990) Cryo-protection of protein crystals against radiation damage in electron and X-ray diffraction. *Proc Biol Sci* 241(1300):6–8.
6. Garman EF (2010) Radiation damage in macromolecular crystallography: What is it and why should we care? *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):339–351.
7. Scheres SH (2014) Beam-induced motion correction for sub-megadalton cryo-EM particles. *eLife* 3:e03665.
8. Grant T, Grigorieff N (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* 4:e06980.
9. Neutze R, Wouts R, van der Spoel D, Weckert E, Hajdu J (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* 406(6797):752–757.
10. Chapman HN, et al. (2011) Femtosecond X-ray protein nanocrystallography. *Nature* 470(7332):73–77.
11. Boutet S, et al. (2012) High-resolution protein structure determination by serial femtosecond crystallography. *Science* 337(6092):362–364.
12. Barty A, et al. (2012) Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nat Photonics* 6(1):35–40.
13. Lomb L, et al. (2011) Radiation damage in protein serial femtosecond crystallography using an x-ray free-electron laser. *Phys Rev B* 84(21):214111.
14. Nass K, et al. (2015) Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *J Synchrotron Radiat* 22(2):225–238.
15. Stan CA, et al. (2016) Liquid explosions induced by X-ray laser pulses. *Nat Phys*, 10.1038/nphys3779.
16. Kirian RA, et al. (2010) Femtosecond protein nanocrystallography-data analysis methods. *Opt Express* 18(6):5713–5723.
17. White TA, et al. (2012) CrystFEL : A software suite for snapshot serial crystallography. *J Appl Cryst* 45(2):335–341.
18. Redecke L, et al. (2013) Natively inhibited trypanosoma brucei cathepsin B structure determined by using an X-ray laser. *Science* 339(6116):227–230.
19. Liu W, Wacker D, Wang C, Abola E, Cherezov V (2014) Femtosecond crystallography of membrane proteins in the lipidic cubic phase. *Philos Trans R Soc Lond B Biol Sci* 369(1647):20130314–20130314.
20. Caleman C, et al. (2015) Ultrafast self-gating Bragg diffraction of exploding nanocrystals in an X-ray laser. *Opt Express* 23(2):1213–1231.
21. Chapman HN, Caleman C, Timneanu N (2014) Diffraction before destruction. *Philos Trans R Soc Lond B Biol Sci* 369(1647):20130313.
22. Son S-K, Young L, Santra R (2011) Impact of hollow-atom formation on coherent x-ray scattering at high intensity. *Phys Rev A* 83(3):33402.
23. Aquila A, et al. (2015) The linac coherent light source single particle imaging road map. *Struct Dyn* 2(4):041701.
24. Zhang H, et al. (2015) Structure of the Angiotensin receptor revealed by serial femtosecond crystallography. *Cell* 161(4):833–844.
25. Ginn HM, et al. (2015) Structure of CPV17 polyhedrin determined by the improved analysis of serial femtosecond crystallographic data. *Nat Commun* 6:6435.
26. Gebhardt MM, Eberle KE, Radtke P, Jehle JA (2014) Baculovirus resistance in codling moth is virus isolate-dependent and the consequence of a mutation in viral gene pe38. *Proc Natl Acad Sci USA* 111(44):15711–15716.
27. Rohrmann G (2013) *Baculovirus Molecular Biology* (National Center for Biotechnology Information, Bethesda), 3rd Ed.
28. Coulibaly F, et al. (2009) The atomic structure of baculovirus polyhedra reveals the independent emergence of infectious crystals in DNA and RNA viruses. *Proc Natl Acad Sci USA* 106(52):22205–22210.
29. Ji X, et al. (2010) How baculovirus polyhedra fit square pegs into round holes to robustly package viruses. *EMBO J* 29(2):505–514.
30. Trillo-Muyo S, et al. (2013) Ultratight crystal packing of a 10 kDa protein. *Acta Crystallogr D Biol Crystallogr* 69(Pt 3):464–470.
31. DePonte DP, et al. (2008) Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J Phys D Appl Phys* 41(19):195505.
32. Weierstall U, Spence JCH, Doak RB (2012) Injector for scattering measurements on fully solvated biospecies. *Rev Sci Instrum* 83(3):035108.
33. Gañán-Calvo AM (1998) Generation of steady liquid microthreads and micron-sized monodisperse sprays in gas streams. *Phys Rev Lett* 80(2):285–288.
34. Barty A, et al. (2014) Cheetah: Software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J Appl Cryst* 47(Pt 3):1118–1131.
35. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
36. Eberle KE, Wennmann JT, Kleespies RG, Jehle JA (2012) *A Manual of Techniques in Invertebrate Pathology* (Academic Press, San Diego), 2nd Ed, pp 15–74.
37. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125–132.
38. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* 69(Pt 7):1204–1214.
39. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.
40. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
41. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084):1030–1033.
42. Lomb L, et al. (2012) An anti-settling sample delivery instrument for serial femtosecond crystallography. *J Appl Cryst* 45(4):674–678.
43. Hart P, et al. (2012) *The CSPAD Megapixel X-ray Camera at LCLS*, eds Moeller SP, Yabashi M, Hau-Riege, SP (SPIE Press, San Diego) Vol 8054, pp 85040C.
44. White TA, et al. (2016) Recent developments in CrystFEL. *J Appl Cryst* 49(Pt 2): 680–689.
45. Brehm W, Diederichs K (2014) Breaking the indexing ambiguity in serial crystallography. *Acta Crystallogr D Biol Crystallogr* 70(Pt 1):101–109.
46. White TA, et al. (2013) Crystallographic data processing for free-electron laser sources. *Acta Crystallogr D Biol Crystallogr* 69(Pt 7):1231–1240.
47. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
48. Terwilliger TC, et al. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* 64(Pt 1):61–69.
49. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
50. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352–367.
51. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.