## Title

Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam



## Journal

Proceedings of the National Academy of Sciences of the United States of America, 111(35)

## ISSN

0027-8424

## Authors

Sawaya, Michael R
Cascio, Duilio
Gingery, Mari
et al.

## Publication Date

2014-09-02

## DOI

10.1073/pnas.1413456111

Peer reviewed

# Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam

Michael R. Sawaya[a,b,1], Duilio Cascio[a,b,1], Mari Gingery[a,b,1], Jose Rodriguez[a,b], Lukasz Goldschmidt[a,b], Jacques-Philippe Colletier[c,d,e], Marc M. Messerschmidt[f,2], Sébastien Boutet[f], Jason E. Koglin[f], Garth J. Williams[f], Aaron S. Brewster[g], Karol Nass[h], Johan Hattne[g], Sabine Botha[h], R. Bruce Doak[h,i], Robert L. Shoeman[h], Daniel P. DePonte[f], Hyun-Woo Park[j,3], Brian A. Federici[j,k], Nicholas K. Sauter[g], Ilme Schlichting[h], and David S. Eisenberg[a,b,l,4]

[a]UCLA–DOE Institute for Genomics and Proteomics, [b]Department of Biological Chemistry, and [l]Howard Hughes Medical Institute, University of California, Los Angeles, CA 90095-1570; [c]Université Grenoble Alpes, [d]Centre National de la Recherche Scientifique, and [e]Commissariat à l'Energie Atomique, Institut de Biologie Structurale, F-38044 Grenoble, France; [f]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025; [g]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [h]Max Planck Institute for Medical Research, 69120 Heidelberg, Germany; [i]Department of Physics, Arizona State University, Tempe, AZ 85287; and [j]Department of Entomology and [k]Graduate Program in Cell, Molecular and Developmental Biology, University of California, Riverside, CA 92521

Contributed by David S. Eisenberg, July 23, 2014 (sent for review April 22, 2014)

**It has long been known that toxins produced by *Bacillus thuringiensis* (Bt) are stored in the bacterial cells in crystalline form. Here we describe the structure determination of the Cry3A toxin found naturally crystallized within Bt cells. When whole Bt cells were streamed into an X-ray free-electron laser beam we found that scattering from other cell components did not obscure diffraction from the crystals. The resolution limits of the best diffraction images collected from cells were the same as from isolated crystals. The integrity of the cells at the moment of diffraction is unclear; however, given the short time (~5 μs) between exiting the injector to intersecting with the X-ray beam, our result is a 2.9-Å-resolution structure of a crystalline protein as it exists in a living cell. The study suggests that authentic in vivo diffraction studies can produce atomic-level structural information.**

XFEL | Cry3A insecticidal toxin | serial femtosecond crystallography

The advent of X-ray free-electron lasers (XFELs) has made it possible to obtain atomic resolution macromolecular structures from crystals with sizes approximating only 1/60th of the volume of a single red blood cell. Brief, intense pulses of coherent X-rays, focused on a spot of 3-μm diameter, have produced 1.9-Å-resolution diffraction data from a stream of lysozyme crystals, each crystal no bigger than 3 $\mu m^3$ (1). A stream of crystals, not just one crystal, is required to collect the many tens of thousands of diffraction patterns that compose a complete data set. No single crystal can contribute more than one diffraction pattern because the XFEL beam is so intense and the crystals so small that the crystals are typically vaporized after a single pulse. Impressively, a photosystem I crystal no bigger than 10 unit cells (300 nm) on an edge produced observable subsidiary diffraction peaks between Bragg reflections, details which would be unobservable from conventionally sized crystals (2). With this new ability to collect diffraction patterns from crystals of unprecedentedly small dimensions, it is conceivable that high-resolution diffraction data could be collected from crystals in vivo. The structure obtained in this manner would be unaltered from that occurring naturally in a living cell, free from distortion that might otherwise potentially arise from nonphysiological conditions imposed by recrystallization. A practical advantage would also be gained by eliminating the need for a protein purification step, whether the in vivo grown crystals were naturally, or heterologously expressed (3).

The nascent field of serial femtosecond crystallography (SFX) has published results on nine different macromolecular systems since its inception in 2009 (Table 1). One system in particular, cathepsin B, marks an advancement toward in vivo crystallography (3, 9). The crystals for this study were *not* grown in artificial crystallization chambers as has been the protocol of conventional macromolecular crystallography since the 1950s. Instead, crystals were grown in cells. Specifically, they were grown in Sf9 insect cells, heterologously expressing *Trypanosoma brucei* cathepsin B. These in vivo-grown crystals were used for the XFEL diffraction experiment. To this end, the cells were lysed and the crystals were extracted before injecting them in the XFEL beam for data collection. This last purification step seems to be the only major departure from our goal of obtaining high-resolution structural information from crystal inclusions in vivo, without requiring the crystal to be extracted from the cell that assembled it. Here we attempt to go one step further than previous studies—to record diffraction from crystals within living cells.

## Significance

**In vivo microcrystals have been observed in prokaryotic and eukaryotic cells. With rare exception, however, the ~100,000 biological structures determined by X-ray crystallography to date have required the macromolecule under study to be extracted from the cells that produced it and crystallized in vitro. In vivo crystals present a challenge for structure determination and pose the question of the extent to which in vivo macromolecular structures are similar to those of extracted and recrystallized macromolecules. Here we show that serial femtosecond crystallography enabled by a free-electron laser yields the structure of in vivo crystals, as they exist in a living cell, and in this case the in vivo structure is essentially identical to the structure of extracted and recrystallized protein.**

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

The authors declare no conflict of interest.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4QX0, 4QX1, 4QX2, and 4QX3).

[1]M.R.S., D.C., and M.G. contributed equally to this work.

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

[3]Present address: Department of Natural and Mathematical Sciences, California Baptist University, Riverside, CA 92504.

[4]To whom correspondence should be addressed. Email: david@mbi.ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1413456111/-/DCSupplemental.

**Table 1. SFX publications from XFEL sources to date**

| Publication date | System | Product | Resolution (Å) | Title of publication | Authors | Reference |
|---|---|---|---|---|---|---|
| Feb 2011* | Photosystem I | Structure | 8.7 | Femtosecond X-ray protein nanocrystallography | Chapman et al. | 2 |
| Dec 2011* | Lysozyme | Structure | 8.7 | Radiation damage in protein serial femtosecond crystallography using an X-ray free-electron laser | Lomb et al. | 4 |
| Jan 2012* | Photosystem I-Ferredoxin | Data | 11 | Time-resolved protein nanocrystallography using an X-ray free-electron laser | Aquila et al. | 5 |
| Jan 2012* | Cathepsin B | Data | 7.5 | In vivo protein crystallization opens new routes in structural biology | Koopman et al. | 3 |
| Jan 2012* | Photosynthetic Reaction Center | Structure | 7.4 | Lipidic phase membrane protein serial femtosecond crystallography | Johansson et al. | 6 |
| Jun 2012 | Photosystem II | Structure | 6.6 | Room temperature femtosecond X-ray diffraction of photosystem II microcrystals | Kern et al. | 7 |
| Jul 2012 | Lysozyme | Structure | 1.9 | High-resolution protein structure determination by serial femtosecond crystallography | Boutet et al. | 1 |
| Nov 2012 | Thermolysin | Data | 4.0 | Nanoflow electrospinning serial femtosecond crystallography | Sierra et al. | 8 |
| Jan 2013 | Cathepsin B | Structure | 2.1 | Natively inhibited *Trypsanosoma brucei* cathepsin B structure determined by using an X-ray laser | Redecke et al. | 9 |
| Apr 2013 | Photosystem II | Structure | 5.7 | Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature | Kern et al. | 10 |
| May 2013 | Lysozyme | Structure | 3.2 | Anomalous signal from S atoms in protein crystallographic data from an X-ray free-electron laser | Barends et al. | 11 |
| Sept 2013 | Ribosome | Data | <6 | Serial femtosecond X-ray diffraction of 30S ribosomal subunit microcrystals in liquid suspension at ambient temperature using an X-ray free-electron laser | Demirci et al. | 12 |
| Dec 2013 | Photosynthetic Reaction Center | Structure | 3.5 | Structure of a photosynthetic reaction center determined by serial femtosecond crystallography | Johansson et al. | 13 |
| Dec 2013 | Serotonin receptor | Structure | 2.8 | Serial femtosecond crystallography of G protein-coupled receptors | Liu et al. | 14 |
| Jan 2014 | Lysozyme + Gd | Structure | 2.1 | De novo protein crystal structure determination from XFEL data | Barends et al. | 15 |
| This study | Cry3A toxin, isolated crystals and whole cells | Structure | 2.8, 2.9 | 2.9 Å-Resolution protein crystal structure obtained from injecting bacterial cells into an X-ray free-electron laser beam | Sawaya et al. | This study |

*The available XFEL energy was limited to 2 keV (6.2 Å wavelength) when these experiments were conducted.

Our target for in vivo crystal structure determination is the insecticidal Cry3A toxin from *Bacillus thuringiensis* (Bt). The bacterium naturally produces crystals of toxin during sporulation (16). Presumably, the capacity for in vivo crystallization evolved in Bt as a mechanism to store the toxin in a concentrated, space-efficient manner. Since the 1920s, farmers have used the crystalline insecticidal proteins to control insect pests; its production as a natural pesticide is now a commercial enterprise. Attempts to structurally characterize the toxins date back to more than 40 y ago with the first report of diffraction from isolated crystals that were packed together in powder form to obtain a measurable signal; X-ray sources available at the time were relatively weak (17). More than 20 y later, the structure was determined at 2.5-Å resolution by single crystal diffraction using a synchrotron X-ray source (18). However, to achieve this result, the authors dissolved the naturally occurring microcrystals and recrystallized the toxin using the hanging drop vapor diffusion method. To date, more than a dozen Bt toxin structures have been reported from various strains [Protein Data Bank (PDB) ID codes 1cby, 1ciy, 1i5p, 1ji6, 1w99, 2d42, 2c9k, 2rci, 3eb7, 2ztb, 3ron, 4d8m, 4ato, 4ary, and 4arx], but none using naturally occurring crystals, and all of the crystals had lost their native context.

In pursuit of in vivo diffraction, we took advantage of the Bt subsp. *israelensis* strain 4Q7/pPFT3As to produce the largest in vivo crystals achievable. This strain contains the plasmid pPFT3As, which increases expression of Cry3A by 12.7-fold over wild type by using strong promoters and an mRNA stabilizing sequence (19). The level of Cry3A production is such that the cell essentially distorts to take on the shape of the enclosed crystal. The calculated average crystal volume is 0.7 $\mu m^3$ (19), almost accounting for the volume of the cell. To explore the possibilities for in situ data collection of in vivo microcrystals, we injected both the crystals in cells and crystals that we isolated from cells in the XFEL beam and collected SFX diffraction data. Our experiments revealed that the cell wall and other cellular components are not an obstacle to achieving 2.9-Å-resolution diffraction, and analogous studies in other systems might be similarly successful.

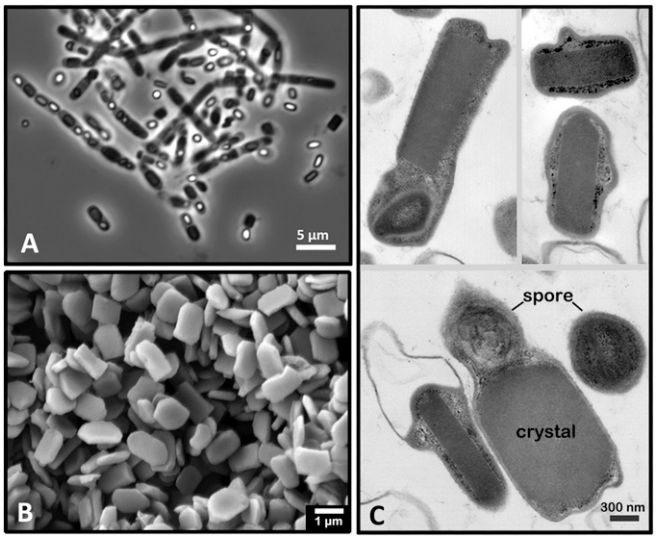


**Fig. 1.** Samples used for XFEL diffraction studies. (*A*) Phase contrast light micrograph of sporulating Bt cells (rod shaped). The dark rectangular shapes inside (and a few outside) cells correspond to the Cry3A toxin crystals. The bright white oval shapes correspond to spores. The micrograph shows that the cells, suspended in pure water, remain intact with no added buffers. (*B*) Scanning electron micrograph of Cry3A crystals isolated from cells. The image shows that the sample is free of large cell debris and that the crystals have a relatively uniform size. (*C*) Transmission electron micrograph of thin-sectioned Bt cells showing that the crystals (rectangular objects with uniform electron density) are so large that the cells are distended to the shape of the crystals. The rounded objects in the cells (and free-floating; *Lower*) are spores.

## Results

**Production and Isolation of in Vivo-Grown Cry3A Crystals.** Cry3A crystals were produced by acrystalliferous Bt subsp. *israelensis* strain 4Q7 containing plasmid pPFT3As harboring the Cry3A gene from DSM 2803, a wild-type isolate of Bt *subsp. morrisoni* (strain *tenebrionis*) (4Q7/pPFT3As). The Cry3A gene in pPFT3As produces rectangular plate crystals in cells several-fold larger than in wild-type strains (19), with approximate dimensions of $1.5 \times 1.0 \times 0.5$ μm (Fig. 1). Crystals were isolated as described in *Materials and Methods*.

**Data Collection.** SFX experiments were carried out in March 2013 at the CXI instrument (Coherent X-ray Imaging) at the SLAC Linac Coherent Light Source (LCLS) (20). The photon energy of the X-ray pulses was 8.52 keV (1.45 Å). Each 40-fs pulse contained up to $6 \times 10^{11}$ photons at the sample position, taking into account a beamline transmission of 60%. The diameter of the beam was ∼1 μm. The in vivo-grown crystals were injected into the XFEL beam using a liquid jet injector and a gas dynamic virtual nozzle (21).The micro jet width was ∼4 μm, and the flow rate was 20–50 μL/min. After emerging from the injector tip, the isolated Cry3A crystals or Bt cells travel in a liquid jet through a vacuum chamber for ∼200 μm before they are intercepted by the X-ray pulse. The crystal concentration was adjusted to compromise between maximizing the hit rate and minimizing the observation of multiple crystals per diffraction image, as described in *Materials and Methods*. Diffraction patterns of these crystals or cells were recorded by a Cornell-SLAC pixel array detector (22). The repetition rate of the XFEL pulses was 120 Hz. The sample to detector distance varied from 110 to 180 mm, and the resolution at the edge of detector varied from 2.3 Å to 3.0 Å, depending on the distance to the sample. A total of 380,688 diffraction images were collected for isolated Cry3A crystals and 736,360 images for the Bt cells (Fig. 2 and Table 2).

**Data Processing and Refinement.** The SFX diffraction data collected from isolated crystals and from whole cells were each processed using two different programs, CrystFEL (23) and cctbx.xfel (26, 27), yielding four data sets total (Table 2). Cry3A models were refined against the four data sets. In all cases the starting model for refinement was the structure of Cry3A (PDB code 1DLC) obtained from recrystallized material (18), with water molecules removed. We found that although the data statistics differ in some aspects (Table 2 and Figs. S1 and S2), the quality of the models obtained from data processed by each of the two programs is similar, as judged by the similarity of the refined models (Fig. S3); the rmsd between α-carbon positions of the models was only 0.10 Å in the case of data collected from isolated crystals and 0.12 Å in the case of data collected from cells.

**Comparison of Structures of Cell-Grown Crystals and Reconstituted Crystals.** The crystal structure of Cry3A determined using the in vivo-grown crystals showed no significant structural differences from that previously determined from recrystallized Cry3A (18). The rmsd of 584 α-carbon positions is small, 0.14 Å. In fact, the crystals are isomorphous (28) (recrystallized: a = 117.1, b = 134.2, c = 104.5; in vivo-grown: a = 116.9 ± 1.0, b = 135.8 ± 0.7, c = 105.2 ± 0.5; Table 2). The diffraction limit of the recrystallized Cry3A toxin was higher (2.5 Å) than that of the cell-grown crystals (2.8 Å) at LCLS. Recrystallization outside the boundaries of the cell permitted growth of much larger crystals, which more than compensated for the relatively lower brilliance of the second-generation synchrotron source (Deutsches Elektronen-Synchrotron storage ring, DORIS) that was used to collect the data (18). In addition, the in vivo-grown microcrystals may have suffered from increased disorder owing to an ∼10% impurity of unprocessed proprotein, containing an additional 57 residues at the N terminus (29). The impurity was lacking in the recrystallized material owing to exogenous papain treatment of the starting material before crystallization (18). Furthermore, with only 10% of the toxin remaining uncleaved in vivo (29), it was not surprising to find that no electron density was observed for these 57 residues in the maps calculated from any of our data sets.

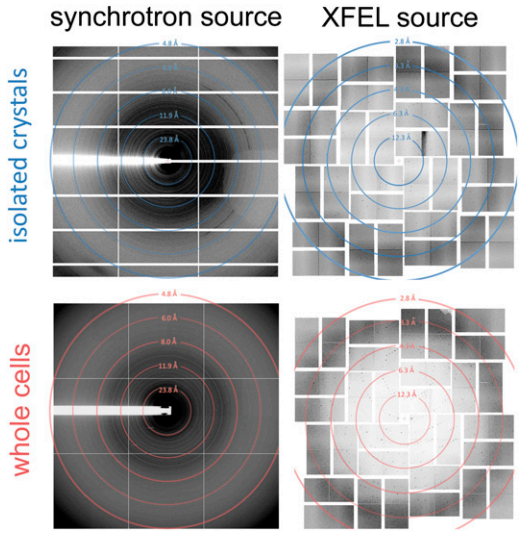


**Fig. 2.** Diffraction images from isolated Cry3A crystals (*Upper*) and cells (*Lower*). The XFEL experiment (LCLS, CXI station) permitted single crystal diffraction to be observed (*Right*), whereas synchrotron sources produced powder diffraction patterns (*Left*). The powder pattern from Cry3A crystals was collected at the Advanced Photon Source Northeastern Collaborative Access Team (APS NECAT) beamline 24-ID-C on a Dectris Pilatus 6M detector. The powder pattern from Bt cells was collected at APS NECAT beamline 24-ID-E using an ADSC Q-315 detector.

**Table 2. Cry3A XFEL data collection and refinement statistics using isolated crystals and whole Bt cells**

| Parameter / Software | Sample: Crystals isolated from Bt cells | | Sample: Whole Bt cells | |
|---|---|---|---|---|
| | cctbx.xfel | CrystFEL | cctbx.xfel | CrystFEL |
| Data collection | | | | |
| Space group | $C222_1$ | $C222_1$ | $C222_1$ | $C222_1$ |
| a (Å) | 116.9 ± 1.0 | 117.1 ± 0.9 | 117.1 ± 1.1 | 117.3 ± 1.1 |
| b (Å) | 135.8 ± 0.7 | 135.4 ± 1.0 | 134.8 ± 0.8 | 135.3 ± 1.2 |
| c (Å) | 105.2 ± 0.5 | 105.6 ± 0.9 | 104.9 ± 0.6 | 105.3 ± 1.1 |
| α, β, γ (°) | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 |
| Wavelength(Å)* | 1.454 ± 0.002 | 1.456 ± 0.002 | 1.457 ± 0.002 | 1.457 ± 0.002 |
| Resolution (Å) | 56.7–2.8 (2.88–2.80) | 88.6–2.8 (2.90–2.80) | 52.4–2.9 (2.99–2.90) | 88.64–2.9 (3.00–2.90) |
| Total patterns | 380,650 | 380,688 | 736,312 | 736,360 |
| Indexed patterns | 78,642 | 76,308 | 30,008 | 30,952 |
| Indexing rate (%)† | 20.7 | 20.0 | 4.1 | 4.2 |
| Total observations | 14,279,911 | 23,731,501 | 4,383,931 | 9,174,339 |
| Multiplicity‡ | 717.8 (1.5) | 1128.0 (691.5) | 252.5 (1.3) | 484.3 (444.5) |
| Unique reflections | 19,894 | 21,038 | 17,360 | 18,944 |
| Completeness (%) | 95.6 (55.7) | 100.0 (100.0) | 92.3 (39.2) | 100.0 (100.0) |
| $R_{split}$(%)§ | 12.2 (75.3) | 15.9 (41.2) | 21.6 (90.6) | 24.4 (48.9) |
| $CC_{1/2}$(%)‡ | 97.7 (27.3) | 91.6 (57.7) | 90.6 (13.6) | 81.7 (40.1) |
| I/σ(I)¶ | 101.2 (1.8) | 9.5 (1.6) | 59.2 (2.3) | 5.2 (0.8) |
| Wilson B ($Å^2$)‡ | 32.4 | 86.4 | 38.0 | 84.6 |
| Refinement | | | | |
| Resolution (Å) | 56.8–2.8 (2.95–2.80) | 44.3–2.8 (2.94–2.80) | 52.4–2.9 (3.08–2.90) | 44.3–2.9 (3.06–2.90) |
| Total reflections | 19,894 | 20,560 | 18,583 | 18,934 |
| $R_{work}$ (%) | 16.5 (28.2) | 17.8 (25.8) | 16.8 (24.8) | 17.7 (26.5) |
| $R_{free}$ (%) | 19.2 (30.8) | 19.7 (26.4) | 20.1 (28.2) | 19.4 (28.5) |
| Protein atoms | 4,659 | 4,659 | 4,659 | 4,659 |
| Water atoms | 26 | 0 | 0 | 0 |
| Protein *B*-factors($Å^2$)‡ | 38.5 | 75.7 | 38.4 | 83.9 |
| Water *B*-factors($Å^2$)‡ | 18.7 | N/A | N/A | N/A |
| rmsd bond lengths (Å) | 0.010 | 0.010 | 0.010 | 0.010 |
| rmsd bond angles (°) | 1.0 | 1.0 | 1.0 | 1.0 |
| Errat score (%)‖ | 97.2 | 96.8 | 96.8 | 97.1 |
| Verify3D score (%)** | 96.1 | 95.9 | 96.9 | 96.8 |
| PDB ID code | 4QX0 | 4QX1 | 4QX2 | 4QX3 |

*The spectral bandwidth of each X-ray pulse for a self-amplified spontaneous emission free electron laser is approx. 0.2%, and the shot-to-shot rms photon wavelength jitter is approx. 0.2%.

†Indexing rate is defined as the number of indexed images per number of patterns collected. It differs from the previous definition given as the number of indexed images per total hits (1). The revision eliminates dependence on subjective choices of "hit" parameters, such as reflection intensity, threshold values, and minimum acceptable spot sizes. By either definition, the indexing rate does not report on diffraction quality. Rather, diffraction quality is reflected in statistics such as $R_{split}$, I/σ(I), and $CC_{1/2}$.

‡Program-specific differences in B-factors and outer shell statistics are due to different acceptance criteria for observations in the outer shells. Refer to *SI Text*.

§We substituted $R_{merge}$ with $R_{split}$ as is appropriate for SFX experiments in which all reflection measurements are partial (23).

¶Methods of estimating I/σ(I) are reported in *SI Text*.

‖Overall Quality Factor (24).

**Percentage of residues with score >0.2 (25).

**Comparison of Data from in Situ Cell and Isolated in Vivo-Grown Crystals.** The resolution limits of the best diffraction images from cells were comparable to those from isolated crystals. However, there is a difference in quality of the data sets (2.8-Å vs. 2.9-Å resolution), largely due to the fewer number of patterns indexed from the in situ cell diffraction experiment. There are less than half as many indexed patterns for the data collected from whole cells compared with isolated crystals by either cctbx.xfel (78,642 from isolated crystals vs. 30,008 from whole cells) or CrystFEL methods (76,308 from isolated crystals vs. 30,952 from whole cells). When equal numbers of indexed images were used, the data set obtained from isolated crystals was slightly better quality than obtained from whole cells (Table S1). The small difference could be due to variation in beam transmission, jet diameter, or scattering from cell components.

Evidence for increased background scattering from cell components in the data collected from whole cells is not obvious. If background scattering from cell components were significant, it might reduce the indexing rate (defined as the ratio of indexed images to total number of images collected) from whole cells compared with isolated crystals. However, this rate is also affected by differences in crystal concentration, for which we do not have an accurate measure. The 16% reduction in indexing rate we observed for whole cells compared with isolated crystals could be due in part to a lower concentration of crystals in the whole cell sample. At first glance, it might seem possible to eliminate the influence of crystal concentration from this ratio by including in the denominator only those images with recorded diffraction events (i.e., "hits"). However, the criteria for defining a hit differs from data set to data set and between indexing algorithms. To find objective evidence of scattering from noncrystalline cell components, we performed a comparison of radial profiles (plots of intensity vs. scattering angle) obtained from analysis of equal numbers of hits from isolated crystals vs. whole cells. It revealed no significant difference in background scattering between the two samples (Fig. S4). It suggests that scattering from cell components in the whole cell sample does not strongly limit the data quality.

## Discussion

**Cell Integrity at the Moment of Diffraction.** The structure of Cry3A toxin described here originates from injecting bacterial cells into an XFEL beam. That is, whole cells were loaded in the injector, and 2.9-Å-resolution diffraction patterns were collected from the crystals injected in the XFEL beam. However, a true in situ experiment would require the crystal to reside inside a cell during the diffraction event. Control experiments have yielded insufficient data to conclude whether this criterion was met (*SI Text*). In the absence of a more definitive experiment, we consider the possibility that lysis occurred before or after the cells reach the XFEL beam position, located only ~200 μm from the nozzle (Fig. 3). At sample flow rates ranging from 20 to 50 μL/min, the cells travel only for 3.0–7.5 μs to reach the XFEL beam from the nozzle exit. Even if all cell walls have ruptured at the instant of exiting the nozzle, a few microseconds may not be sufficient time for the crystals to dislodge from the cells. Even if the crystals do dislodge from cells, a few microseconds are probably insufficient time for the toxin molecules of the cellular crystals to recrystallize into another form. Thus, the crystal diffraction patterns we recorded are not significantly altered from what would be expected in a true in vivo experiment.

**Prospects for in Vivo Diffraction.** There are other systems for which in vivo diffraction may not only be feasible but highly desirable. These include crystals of seed proteins, secretory granules containing the major basic protein from white blood cells called eosinophils and insulin from the islets of Langerhans, enzyme assemblies of urate oxidase and alcohol oxidase produced by peroxisomes, and the protein HEX-1, which composes the proteinaceous core of woronin bodies to prevent cytoplasmic bleeding in filamentous fungi (30). As technological advances increase the attainable intensity of XFEL pulses, atomic resolution structure determination of smaller, less ordered systems, such as carboxysomes, may become feasible.

The results of our studies suggest that the cell components would not prevent obtaining high-resolution diffraction patterns from crystalline material in vivo. We acknowledge that intensified background scattering might present a more serious barrier in other less favorable systems where the crystals make up a smaller portion of the cell volume. Additionally, a different cell delivery system might be required to guarantee the integrity of the cell at the moment of diffraction. However, this technical distinction is likely irrelevant from the point of view of structural biology. If the crystal has not changed its organization during the few microseconds between exiting the nozzle and intercepting the beam, then structures obtained in this manner would reveal the protein crystal structure as it exists inside the living cell.

A thread running through the history of cell biology is the increasing recognition that cellular components are structured. The crystallization of proteins, starting in the 19th century, showed that these large molecules have a defined structure. Sumner's crystallization of urease in 1926 extended this recognition of order to enzymes. Later work revealed the organization of DNA and nucleosomes and the existence of elaborate molecular machines consisting of numerous ordered components. With the advent of electron microscopy in the mid-20th century, it became evident that cells, far from being bags of freely diffusing molecules, are compartmentalized and ordered. That these ordered structures are dynamic, constantly changing, does not contradict the existence of much order at any given instant. As shown by this work, free-electron lasers offer the prospect of interrogating the extent and nature of this order.

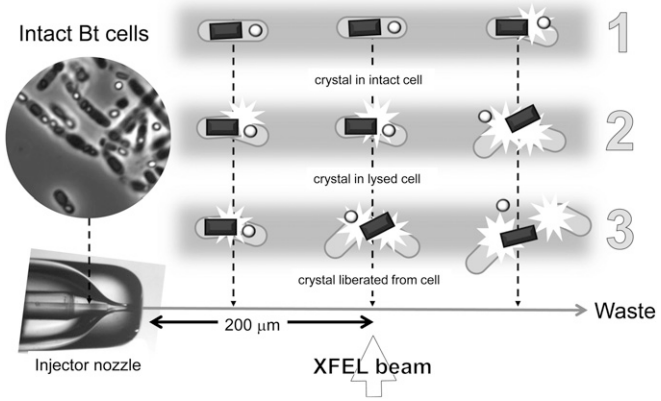


**Fig. 3.** Three scenarios suggesting how the integrity of the cells might vary at the moment of diffraction. The horizontal arrow depicts the flow of sample from injector to waste collection. The XFEL beam intercepts the sample stream ~200 μm from the nozzle. The left, middle, and right columns depict three different time points along the jet trajectory. Depending on the rate of lysis and the flow rate of the jet, the crystals may arrive at the interaction point either (1) inside intact cells, (2) inside lysed cells, or (3) segregated from lysed cells. The time of travel is estimated to be 3–7.5 μs.

## Materials and Methods

**Details of Production and Isolation of in Vivo-Grown Cry3A Crystals.** Crystals were isolated as follows. Five hundred milliliters of glucose-yeast-salts (GYS) liquid growth medium [0.1% glucose, 0.2% yeast extract, 0.05% $K_2HPO_4$, 0.2% $(NH_4)_2SO_4$, 0.002% $MgSO_4$, 0.005% $MnSO_4$, and 0.008% $CaCl_2$] supplemented with 25 μg/mL erythromycin was prefiltered through a 0.22-μm membrane to eliminate dust and suspended contaminants and sterilized by autoclave in a 2-L baffled flask. Media was inoculated with spores (from a lyophilized 3-d lysate) of Bt subsp. *israelensis* strain 4Q7 containing plasmid pPFT3As (4Q7/pPFT3As) (19) and incubated for 3 d at 30 °C with shaking at 250 rpm. Cultures were monitored by phase contrast light microscopy, until sporulation and cell lysis were observed, then spores, crystals, cells, and cell debris were pelleted by centrifugation at 6,000 × *g* for 30 min. The culture pellet was resuspended in 50 mL filtered water and sonicated for 3 min on ice [1 s on, 1 s off (6 min elapsed time); 60% intensity] to lyse remaining cells. The lysate was pelleted at 6,000 × *g* for 30 min at 4 °C, washed in 50 mL filtered water to remove soluble material and some of the spores, then repelleted before being resuspended in 15 mL filtered water. The crystals remained intact and did not dissolve in the absence of ions or buffer (Fig. 1*B*). The crystals can be induced to dissolve if exposed to alkaline conditions (pH 10) as exist in the larval gut. Such were the conditions used to solubilize Cry3A for recrystallization (28). Crystals were separated from other cellular components on sucrose step gradients (11 mL each of filtered 67%, 72%, and 79% wt/vol sucrose solutions) formed in 25 × 89-mm transparent, thin-wall tubes (Beckman). Each gradient was overlaid with 5 mL of lysate and centrifuged in a Beckman SW28 rotor at 35,000 × *g* for 1 h at 4 °C. Crystals formed a wide band above the interface of the 72% and 79% sucrose layers and were recovered from each gradient in 8–10 mL of sucrose solution using a BioComp Gradient Fractionator (BioComp Instruments). Recovered gradient bands were pooled and serially dialyzed six times into 100 volumes of filtered water at 4 °C for ≥1 h to remove sucrose. Dialyzed crystals were pelleted at 6,000 × *g* for 15 min at 4 °C, resuspended in 10 mL filtered water, and stored at 4 °C. After settling, excess liquid was removed to leave a milky-white slurry of suspended crystals.

**Preparation of *Bt* Cells Containing Cry3A Crystals.** Crystal-containing 4Q7/pPFT3As cells were grown from spores inoculated into 500 mL filtered GYS medium supplemented with 25 μg/mL erythromycin and incubated for 1.5–2 d at 30 °C, shaking at 250 rpm. Cells were monitored by phase contrast light microscopy until sporulation and then harvested by centrifugation at 6,000 × *g* for 15 min at 4 °C. All remaining steps were done at 4 °C and all liquids filtered through a 0.22-μm membrane. The cell pellet was washed with 50 mL water, repelleted at 6,000 × *g* for 30 min, and resuspended in 15 mL water. Sucrose step gradients (11 mL each of 67%, 72%, and 79% wt/vol sucrose) in 25 × 89-mm transparent, thin-wall tubes were each loaded with 5 mL of washed cells and centrifuged in an SW28 rotor at 35,000 × *g* for 1 h at 4 °C. Cells forming a broad band in the 72% sucrose step were recovered with a BioComp Gradient Fractionator (BioComp Instruments). Gradient bands were extensively dialyzed into water to remove sucrose (six changes of 100 volumes of filtered water at 4 °C for ≥1 h), then pelleted at 6,000 × *g* for 30 min and resuspended in 2 mL water. No salts, sucrose, or other materials were added. The cells and crystals did not seem to lose integrity because of the absence of ions or buffer (Fig. 1*A*). Samples were passed through a 10-μm stainless steel frit (Upchurch Scientific, part A-107X) to remove any large particles that might clog the sample injector. The frit was seated in an HPLC filter holder (Upchurch

Scientific, part A-356), which was adapted to a 3-mL Luer-Lok tip disposable syringe for convenience of filtering. There was virtually no resistance to pushing the sample through the frit.

**Adjustment of Crystal Concentration.** To maximize the chances of crystals intercepting the X-ray pulses, we aimed for a concentration of $10^{11}$ to $10^{12}$ crystals/mL. To estimate the crystal count, we pelleted the isolated crystals by a 30-s spin in a tabletop centrifuge. We started with 25 μL of crystal pellet diluted to 1 mL with water. Estimating 0.7 $\mu m^3$ per crystal, the calculated concentration corresponded to $3.6 \times 10^{10}$ crystals/mL. Even though this concentration is likely to be an overestimate for lack of accounting for the space between crystals, many of the diffraction patterns showed multiple lattices. We interpreted the appearance of multiple lattices per exposure to signify that the crystal slurry was too concentrated. The sample was then diluted to $2.4 \times 10^{10}$ crystals/mL for the remaining runs. In retrospect, we realize that most of the multiple lattices could have been the result of crystals clumping together, a physical attachment that cannot be broken by dilution. The entire 25-μL crystal pellet was consumed over the course of the crystal diffraction experiment (58 min).

**Adjustment of Cell Concentration.** For the in-cell experiments, the cells were pelleted in the same way as the isolated crystals. We used 25 μL of cell pellet/mL. We did not dilute the cells because there were relatively fewer instances of multiple lattices per diffraction image.

**Algorithms for Processing Serial Femtosecond Crystal Diffraction Images.** The data were processed with cctbx.xfel and CrystFEL, as described in *SI Text*.

**Comparison of SFX Data Processed Using Two Independent Algorithms.** Please refer to *SI Text*.

**Atomic Refinement.** Because the crystals isolated from cells and crystals within cells are isomorphous with the published structure determined using conventional means (18), the atomic refinement was started with a rigid body refinement in phenix.refine (31), followed by atomic refinement, manual rebuilding of the models in COOT (32), and TLS refinement. The final cycles of atomic refinement were performed using BUSTER (33), and COOT was used to add solvent molecules. We used coordinates of the Cry3A model (PDB code 1DLC) as a source of external geometric restraints when refining with Buster. Table 2 shows the results of the atomic refinement. The low values obtained for R and $R_{free}$ are likely related to the high quality of the starting model (1DLC), which was determined at a higher resolution (2.5 Å) than the diffraction data obtained and used for refinement in these experiments. The structure validation was performed using the SAVES server (http://nihserver.mbi.ucla.edu/SAVES/), which validates the models using the programs PROCHECK (34), WHAT_CHECK (35), ERRAT (24), and VERIFY 3D (25) to assess the stereochemical quality, nonbonded interactions, and the compatibility of each amino acid in its local environment. In addition, an analysis of of CC* and $CC_{work}$ offers further evidence that we have not overfit our model. Fig. S5 shows a plot of CC* vs. resolution for each of the four refinements. CC* is an estimate of the correlation between the measured data and hypothetical noise-free signal (36). It is derived mathematically from $CC_{1/2}$, which measures the correlation between two randomly chosen halves of the unmerged data set. Plotted with CC* is $CC_{work}$, the correlation between the measured structure factors in the working set and the corresponding structure factors calculated from the model coordinates. If we had overfit our model, it would be indicated by $CC_{work}$ having a larger value than CC*. It would indicate that the model agrees better with the experimental data than does the true signal. In none of the four refinements do these statistics indicate overfitting.

**ACKNOWLEDGMENTS.** We thank M. Capel, K. Rajashankar, N. Sukumar, J. Schuermann, I. Kourinov, and F. Murphy [Northeastern Collaborative Access Team Beamline 24-ID at the Advanced Photon Source, which is supported by National Center for Research Resources Grant 5P41RR015301-10 and National Institute of General Medical Sciences Grant 8 P41 GM103403-10 from the National Institutes of Health (NIH)]. Use of the Advanced Photon Source is supported by the US Department of Energy (DOE) under Contract DE-AC02-06CH11357. We also thank Harold Aschmann and the University of California, Los Angeles (UCLA)–DOE X-ray Crystallography Core Facility, which is supported by DOE Grant DE-FC02-02ER63421; and Heather McFarlane and Daniel Anderson at UCLA for help with cell preparation and filtration. Portions of this research were carried out at the Linac Coherent Light Source, a National User Facility operated by Stanford University on behalf of the DOE Office of Basic Energy Sciences. The CXI instrument was funded by the Linac Coherent Light Source Ultrafast Science Instruments project funded by the DOE Office of Basic Energy Sciences. This work was supported by Keck Foundation Grant 2843398, NIH Grant AG-029430, National Science Foundation Grant MCB 0958111, DOE Grant DE-FC02-02ER63421, NIH Grants GM095887 and GM102520 for data-processing methods (to N.K.S.), NIH Grant AI45817 (to B.A.F.), Howard Hughes Medical Institute, and the Max Planck Society.

1. Boutet S, et al. (2012) High-resolution protein structure determination by serial femtosecond crystallography. *Science* 337(6092):362–364.
2. Chapman HN, et al. (2011) Femtosecond X-ray protein nanocrystallography. *Nature* 470(7332):73–77.
3. Koopmann R, et al. (2012) In vivo protein crystallization opens new routes in structural biology. *Nat Methods* 9(3):259–262.
4. 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.
5. Aquila A, et al. (2012) Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Opt Express* 20(3):2706–2716.
6. Johansson LC, et al. (2012) Lipidic phase membrane protein serial femtosecond crystallography. *Nat Methods* 9(3):263–265.
7. Kern J, et al. (2012) Room temperature femtosecond X-ray diffraction of photosystem II microcrystals. *Proc Natl Acad Sci USA* 109(25):9721–9726.
8. Sierra RG, et al. (2012) Nanoflow electrospinning serial femtosecond crystallography. *Acta Crystallogr D Biol Crystallogr* 68(Pt 11):1584–1587.
9. Redecke L, et al. (2013) Natively inhibited *Trypanosoma brucei* cathepsin B structure determined by using an X-ray laser. *Science* 339(6116):227–230.
10. Kern J, et al. (2013) Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* 340(6131):491–495.
11. Barends TR, et al. (2013) Anomalous signal from S atoms in protein crystallographic data from an X-ray free-electron laser. *Acta Crystallogr D Biol Crystallogr* 69(Pt 5):838–842.
12. Demirci H, et al. (2013) Serial femtosecond X-ray diffraction of 30S ribosomal subunit microcrystals in liquid suspension at ambient temperature using an X-ray free-electron laser. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69(Pt 9):1066–1069.
13. Johansson LC, et al. (2013) Structure of a photosynthetic reaction centre determined by serial femtosecond crystallography. *Nat Commun* 4:2911.
14. Liu W, et al. (2013) Serial femtosecond crystallography of G protein-coupled receptors. *Science* 342(6165):1521–1524.
15. Barends TR, et al. (2014) De novo protein crystal structure determination from X-ray free-electron laser data. *Nature* 505(7482):244–247.
16. Höfte H, Whiteley HR (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol Rev* 53(2):242–255.
17. Holmes KC, Monro RE (1965) Studies on the structure of parasporal inclusions from *Bacillus thuringiensis*. *J Mol Biol* 14(2):572–581.
18. Li JD, Carroll J, Ellar DJ (1991) Crystal structure of insecticidal delta-endotoxin from *Bacillus thuringiensis* at 2.5 A resolution. *Nature* 353(6347):815–821.
19. Park HW, Ge B, Bauer LS, Federici BA (1998) Optimization of Cry3A yields in *Bacillus thuringiensis* by use of sporulation-dependent promoters in combination with the STAB-SD mRNA sequence. *Appl Environ Microbiol* 64(10):3932–3938.
20. Boutet S, Williams GJ (2010) The Coherent X-Ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). *New J Phys* 12:035024.
21. Weierstall U, Spence JC, Doak RB (2012) Injector for scattering measurements on fully solvated biospecies. *Rev Sci Instrum* 83(3):035108.
22. Philipp HT, Hromalik M, Tate M, Koerner L, Gruner SM (2011) Pixel array detector for X-ray free electron laser experiments. *Nucl Instrum Methods A* 649(1):67–69.
23. White TA, et al. (2012) CrystFEL: A software suite for snapshot serial crystallography. *J Appl Cryst* 45(2):335–341.
24. Colovos C, Yeates TO (1993) Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci* 2(9):1511–1519.
25. Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356(6364):83–85.
26. Sauter NK, Hattne J, Grosse-Kunstleve RW, Echols N (2013) New Python-based methods for data processing. *Acta Crystallogr D Biol Crystallogr* 69(Pt 7):1274–1282.
27. Hattne J, et al. (2014) Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nat Methods* 11(5):545–548.
28. Li J, Henderson R, Carroll J, Ellar D (1988) X-ray analysis of the crystalline parasporal inclusion in *Bacillus thuringiensis* var. tenebrionis. *J Mol Biol* 199(3):543–544.
29. Carroll J, Li J, Ellar DJ (1989) Proteolytic processing of a coleopteran-specific delta-endotoxin produced by *Bacillus thuringiensis* var. tenebrionis. *Biochem J* 261(1):99–105.
30. Doye JPK, Poon WCK (2006) Protein crystallization in vivo. *Curr Opin Colloid Interface Sci* 11(1):40–46.
31. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352–367.
32. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
33. Blanc E, et al. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2210–2221.
34. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 26(2):283–291.
35. Hooft RW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. *Nature* 381(6580):272.
36. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084):1030–1033.