The emerging method of femtosecond crystallography (FX) may extend the diffraction resolution accessible from small radiation-sensitive crystals and provides a means to determine catalytically accurate structures of acutely radiation-sensitive metalloenzymes. Automated goniometer-based instrumentation developed for use at the Linac Coherent Light Source enabled efficient and flexible FX experiments to be performed on a variety of sample types. In the case of rod-shaped Cpl hydrogenase crystals, only five crystals and about 30 min of beam time were used to obtain the 125 still diffraction patterns used to produce a 1.6-Å resolution electron density map. For smaller crystals, high-density grids were used to increase sample throughput; 930 myoglobin crystals mounted at random orientation inside 32 grids were exposed, demonstrating the utility of this approach. Screening results from cryocooled crystals of β2-adrenoreceptor and an RNA polymerase II complex indicate the potential to extend the diffraction resolution obtainable from very radiation-sensitive samples beyond that possible with undulator-based synchrotron sources.

Significance

The extremely short and bright X-ray pulses produced by X-ray free-electron lasers unlock new opportunities in crystallography-based structural biology research. Efficient methods to deliver crystalline material are necessary due to damage or destruction of the crystal by the X-ray pulse. Crystals for the first experiments were 5 μm or smaller in size, delivered by a liquid injector. We describe a highly automated goniometer-based approach, compatible with crystals of larger and varied sizes, and accessible at cryogenic or ambient temperatures. These methods, coupled with improvements in data-processing algorithms, have resulted in high-resolution structures, unadulterated by the effects of radiation exposure, from only 100 to 1,000 diffraction images.


See Commentary on page 16898.

1To whom correspondence may be addressed. Email: HodgsonK@stanford.edu or acohen@slac.stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl; doi:10.1073/pnas.1418733111/DCSupplemental.
pioneering experiments demonstrated the utility of serial femtosecond crystallography (SFX) and the use of crystals of less than 5 μm in size, often termed “nanocrystals” (NCs), for macromolecular structure determination to high resolution (16, 17). As NCs may be a ubiquitous but generally overlooked outcome of commercial crystallization screens that fail to produce larger crystals (18), FX may open up many systems to crystallographic analysis. However, to develop FX into a generally applicable method, a number of challenges in the areas of sample preparation, data collection, and data processing must be overcome.

Obtaining a sufficient supply of crystals in an appropriate carrier solution is a first hurdle to conducting a SFX experiment. In addition to the GDVN (2, 3, 14, 16, 17, 19, 20), other injectors such as a nanoflow electrospinning injector (21) and a lipidic cubic phase (LCP) injector (22), have been developed that have a reduced flow rate and lower sample consumption. However, because injectors deliver a continuous stream of solution containing a random distribution of crystals, and the X-ray pulses are extremely short, often only a small percentage of pulses hit a crystal and produce a useful diffraction pattern. Carrying out these experiments at room temperature avoids the difficulties associated with cryoprotection, and datasets obtained at ambient temperatures can provide insight on the functional motions of protein molecules (23). However, there are different and often more complex optimization steps associated with specific injector technologies. Solutions containing a mixture of crystal sizes may require filtering to avoid clogging in the injector nozzle, and delicate crystals may be damaged from the pressures and shear forces of the delivery process itself (24). For experiments conducted in vacuo, stream formation may be disrupted by solution bubbling, drying, or freezing as it exits the injector and enters the vacuum chamber. Drop-on-demand methods that deliver single drops containing crystals to individual X-ray pulses have the potential to significantly reduce sample consumption are in development, such as acoustic and micropiezo activated technologies, but implementation has been complicated by a variety of factors, including difficulties imposed by viscous solutions and unpredictable trajectories of drops that contain crystals of varied shapes and sizes.

Here, we describe an alternative strategy for FX experiments that leverages the well-established benefits of the highly automated goniometer-based setups used at state-of-the-art microfocus synchrotron beam lines, and expands these technologies to take full advantage of the unique capabilities of XFEL sources. Key to this approach is the coupling of highly automated instrumentation with specialized sample containers and customized software for efficient data collection with minimal sample consumption. High-density sample containers, such as microfluidic chips or microcrystal traps (25) for room temperature studies or grids for experiments at cryogenic temperatures, hold samples in known locations. These sample holders enable very rapid and precise positioning of crystals into the X-ray interaction region for consistent production of diffraction patterns. To optimize data completeness and resolution, data may be collected using a range of crystal sizes with a variety of X-ray beam sizes, and different regions of larger crystals may be exposed in different orientations. When small crystals align with the sample holder in a preferred orientation, exposing each crystal at varied angles to the holder surface may take advantage of this effect to enhance completeness. When the X-ray pulse mean diameter is greater than about 8 μm or is highly attenuated, the protein crystal may remain intact after exposure to the X-ray pulse and still diffract, but usually to a lower resolution as a result of radiation damage. In these cases, it is possible to rotate the crystal and collect additional diffraction patterns to use as an aid in indexing and scaling the partially recorded reflections of the initial still diffraction pattern.

**Instrumentation**

The experimental setup (Fig. 1 and Fig. S1) combines equipment developed both at the LCLS and by the Structural Molecular Biology group at the Stanford Synchrotron Radiation Lightsource (SSRL). The setup was built into the LCLS X-ray pump/probe instrument (XPP) using its X-ray optical and diagnostics components and energy spectrometer (26). Beryllium compound refractive lenses (27) were used to focus the LCLS self-amplified spontaneous emission beam down to sizes of 5 μm. An SSRL microdiffractometer (goniometer) enables precise and rapid positioning of crystals into the X-ray beam interaction region for data collection (SI Materials and Methods and Table S1). A sample mounted on the microdiffractometer is viewed along the X-ray propagation direction by an in-line camera system. This camera supplies a live video image to the control software, which provides a click-to-center system: clicking on the image of the sample causes the clicked point to move to the rotation center of the microdiffractometer. The Stanford automated mounter (SAM) (28) is used for sample exchange and includes storage locations for 288 sample mounts at cryogenic temperatures and 6 at room temperature. Other equipment includes a beamstop and scatter guard assembly, and a Rayonix MX325HE CCD detector mounted in a positioning cradle. The sample handling and crystallography experiment is controlled through the Blu-Ice/DCS software (29), which can request pulses from LCLS and initiate synchronized data streams with the DAQ software through the use of LCLS-provided EPICS interface. This highly automated system enables the experiment to be performed from the XPP control room without depending on personnel access to the experimental area. Diffraction images were analyzed in quasi-real time to determine unit cell parameters, diffraction resolution, and dataset completeness using the cctbx.xfel package (30).

**Data Collection Protocols**

Macromolecular crystals have a variety of characteristics such as size, shape, mechanical and chemical stress response, and radiation sensitivity. Therefore, a portfolio of sample holders and data collection strategies was developed to meet different experimental requirements and to optimize efficiency. To this end, data collection modes for serial crystallography were developed and incorporated into the Blu-Ice/DCS graphical user interface and experimental control software (29). These modes include click-and-shoot (exposure to a single X-ray pulse at a selected position), raster (automated translation of the sample between exposures), helical (automated translation and rotation of the sample), and a highly efficient multicrystal data collection mode that combines the capabilities of the previous modes, with automated positioning of crystals inside high-density sample holders (such as grids) for rapid multicrystal experiments. For each mode, the photon energy distribution of each X-ray pulse used to collect a diffraction pattern may be recorded using the energy spectrometer (SI Materials and Methods and Fig. S2).

**Raster Data Collection**

The rastering mode of data collection was developed to collect still diffraction images from crystals mounted on meshes, nylon loops, and grids with the goal of extending the resolution obtained from very radiation-sensitive samples. After a sample is robotically mounted on the microdiffractometer, the experimenter uses the software to define a single position (click-and-shoot) or an area of the sample to use for data collection (Fig. S3). Single-shot images are collected within this area, with the translation between shots specified through the user interface. Typically, a focused XFEL pulse with beam sizes of ~8 μm or less will vaporize the sample (Fig. 2), leaving a hole larger than the beam size. An appropriate translation distance is selected.
to ensure that each shot interacts with an unaffected part of the sample.

Diffraction quality screening experiments were carried out using the rastering mode of data collection and grids containing micrometer-sized crystals of a multiprotein complex, composed of RNA polymerase II (Pol II), general transcription factor IIB (TFIIB), and a large nucleic acid scaffold (NAS) (SI Materials and Methods). Diffraction patterns of superior resolution (3.3 Å) were observed when using the microdiffractometer setup at LCLS XPP at 100 K than were previously measured at LCLS-CXI or at the synchrotron (Table S2). The lower diffraction resolution observed (4.0 Å) during screening experiments at LCLS CXI using smaller crystals and the GDVN injector suggests that a larger sample volume may be required to obtain higher resolution diffraction, or that the crystals were damaged during injection, potentially a consequence of their high solvent content (24). The best diffraction resolution (3.7 Å) observed at SSRL BL12-2 from these crystals (grown in same crystallization batch as used at LCLS XPP), suggests that Pol II–TFIIB–NAS crystals suffer significant radiation damage during the exposure time needed to deliver an equivalent dose at a synchrotron.

Improved diffraction resolution was also observed for β2-adrenoceptor/nanobody complex crystals collected at 100 K at LCLS XPP using the microdiffractometer. The best diffraction resolution obtained was 2.3 Å (Fig. 3), and overall screening results indicate a dataset of 2.8-Å resolution could be collected given sufficient time. In contrast, the best resolution observed, using crystals from the same crystallization batch, was 2.9 Å at the Advanced Photon Source 23ID, where a 3.2-Å dataset was obtained (31).

**Helical Data Collection**

A helical data collection mode was developed to efficiently determine “radiation damage-free” structures of sensitive metalloenzymes that form conventionally sized and ideally long crystals. A single X-ray pulse will initiate photoreduction within the exposed region of the crystal, so dataset completeness is optimized by using a beam size smaller than the crystal size and collecting multiple still diffraction patterns from different areas across the crystal. To begin, the experimenter draws a line defining the length and position of the crystal that is overlaid on the real-time sample video display within Blu-Ice (Fig. 4 and Fig. S4), and optionally defines a position on the end of the crystal to collect a still diffraction pattern to use for a strategy calculation (SI Materials and Methods). During data collection, the crystal is rotated to a

![Fig. 1. Schematic representations of the experimental environment. (A) An overview of the experimental setup illustrating the microdiffractometer on the experimental table, the SAM sample exchange robot with sample storage Dewar, the Rayonix 325 CCD detector in a motorized positioning cradle, and the on-axis cryocooler nozzle (orange) for measurements at cryogenic temperatures. (B) A close-up view of the sample environment showing a sample holder (orange) mounted on the diffractometer and the beamstop/collimator assembly (dark blue) inserted for data collection. (C) A close-up view of the sample environment with the beamstop/collimator assembly retracted, side illumination (yellow), and the backlight (green) inserted and for sample viewing with the in-line camera system, that consists of a drilled mirror (aqua) that directs an image of the sample onto a motorized zoom assembly and CCD (purple).](image)

![Fig. 2. A sample mesh after raster data collection using a 3-μm beam size and a larger step size showing areas vaporized by the X-ray pulse.](image)
Data collection using grids starts with a semiautomated alignment procedure including fiducialization of the grid (SI Materials and Methods and Fig. S6). The locations of the grid ports are calculated and overlaid on the video view of the grid. In cases where the crystal size is close to the size of the grid port, the automated data collection system centers each port containing a crystal into the X-ray interaction region and collects a still diffraction pattern. Alternatively, after translating to a new port, a different area of the port may be identified by the experimenter and centered to use for data collection. If many smaller crystals are in a port, an area may be specified for raster data collection, or an option to raster the entire port area may be selected (Fig. S3).

**Oscillation Data**

Depending on the beam size, beam intensity, and crystal properties, at cryogenic temperatures the crystal may remain intact after exposure to a single X-ray pulse. In these cases, to aid in processing the resulting still diffraction pattern, the data collection modes include options for collecting additional data from the crystal.

“Oscillation” images are produced by rotating the sample at a constant velocity of 1°/s while exposing the crystal to attenuated X-ray pulses at 120 Hz (Fig. S7). Therefore, a single 1° oscillation image is comprised of 120 overlaid still diffraction patterns. When feasible, 5–11° of oscillation images are collected, centered on the orientation of the first still diffraction pattern. The oscillation data mimic rotation data typically collected at a synchrotron, enabling the use of conventional data analysis programs, such as XDS (33), MOSFLM (34), and POINTLESS (35) to define the unit cell and orientation matrix. This information may then be applied to the corresponding still diffraction pattern to improve estimates of reflection partiality, resolve indexing ambiguities of some space groups that can result in apparent twinning when merging still diffraction patterns from different crystals, and improve scaling of the corresponding still diffraction data.

Hexagonal (P6) myoglobin is an example where an indexing ambiguity between crystals was overcome by the use of oscillation data, which enabled the processing of 739 still diffraction images collected at full intensity (SI Materials and Methods). This approach resulted in a radiation damage-free dataset complete to 1.48 Å and extending to 1.36 Å (Table S4). Following structure solution by molecular replacement (MR) using a monoclinic myoglobin model (PDB ID code 1VXA), the heme group, which was omitted in the MR search model, was clearly visible in the resultant 2Fo − Fc electron density map (Fig. 5C) electron density map. The structure was refined to a total R factor of 15.9% and free R factor of 18.8%; the location of the refined heme group in the original MR electron density is shown (Fig. 5D).
Discussion

The collection of high redundancy data are particularly important for FX experiments, as these datasets are produced from scaled partial observations of each reflection. The goniometer-based instrumentation described here provides an efficient and flexible framework with which to carry out these experiments, using automated strategies tailored to handle a variety of sample requirements, crystal sizes, and experimental goals. These developments, coupled with recent improvements in data-processing algorithms (30), make it possible to derive high-resolution structures, unadulterated by the effects of radiation exposure, using only 100–1,000 still diffraction images. The exposure of over 930 randomly oriented myoglobin crystals using only 32 grids demonstrates the utility of high-density sample containers, such as grids to optimize throughput and sample consumption. These experiments have also shown that the collection of oscillation data following the collection of a single still diffraction pattern offers a very reliable way to resolve indexing ambiguities using conventional software. Although oscillation data can be collected and indexed even from severely damaged crystals, it is necessary that the crystal retains physical integrity after the first still shot, which precludes this approach when a focused beam of ~8 μm or less is used. In cases where an isomorphous dataset from a conventional source is not available, constructing a dataset from the oscillation images collected from each crystal can help with processing the still data. This also provides performance benchmarks for the development of more sophisticated methods for single-shot data scaling and postrefinement.

The mean photon energy used for analyzing each still diffraction image of the myoglobin dataset was determined by calculating the centroid of the energy spectrum of each X-ray pulse measured by the single-shot spectrometer. In the most typical mode of LCLS operation, each X-ray pulse varies in intensity and energy distribution within a window of about 60 eV. Data quality may be improved through the development of algorithms that take the intensity and entire energy spectrum of individual pulses into account. However, variation in quality and nonisomorphism between crystals and even between areas within larger crystals, which can experience uneven stresses across the sample mount, may in some cases have a significantly more harmful effect.

When long crystals are available, helical data collection is a very efficient method of data collection at XFEL sources. Recently, in the case of rod-shaped hydrogenase crystals in tetragonal space group P4₁2₁2, only five crystals and about a half an hour of LCLS beam time were used to obtain the 125 still diffraction patterns used to produce a high quality electron density map to 1.6-Å resolution.

Diffraction quality screening results from Pol II–TFIIIB–NAS and β₂-adrenoreceptor GPCR crystals illustrate the utility of a microdiffractometer setup at LCLS to extend the resolution that can be obtained from very radiation-sensitive samples. However, it should be noted that, for a number of other crystals, improvements in diffraction resolution were not observed at LCLS XPP compared with the diffraction measured at the synchrotron using an equivalent X-ray dose. Furthermore, for a few proteins, exposure to a higher X-ray dose at the synchrotron produced diffraction patterns of superior resolution. It is possible that brighter XFEL pulses may further improve diffraction for these cases. Alternatively, for many conventionally sized crystals, the diffraction obtained on a single image may be limited by the intrinsic order of the crystal rather than radiation damage. As the FX field is in its infancy, systematic testing is necessary to refine the method and sample selection process. Screening crystals of a few microns and larger in size for radiation sensitivity and amenability to cryopreservation at the synchrotron will be valuable steps to identify which samples would benefit most from the FX approach with currently available sources. As many of these will be challenging systems, goniometer-based experimentation provides a reliable and efficient delivery method to conduct FX experiments compatible with fragile crystals of limited quantity at both cryogenic and ambient temperatures.

The use of specifically designed sample holders, the implementation of the CSPAD (37), which can produce images at 120 Hz, or future improved versions of this detector with better dynamic range, and an upgrade of the microdiffractometer will enable data collection rates that match the maximum pulse frequency of 120 Hz delivered by the LCLS XFEL. An upgraded version of the microdiffractometer with faster translation axes...
will target samples spaced 250 μm apart at 120 Hz while accommodating inclinations off the horizontal axis of up to 1 part in 10.

Although these efforts will be important to improve sample throughput, the limited amount of beam time available will remain a serious factor impeding progress of the FX field as a whole. The construction of new sources (38) will help to alleviate this situation. Multiplexing XFEL pulses to multiple experimental instruments is another attractive option to increase access to beam time at LCLS. To take advantage of multiplexing modes (SI Materials and Methods), a dedicated experimental station with similar highly automated remote-access capable goniometer-based instrumentation, is currently being built and will enable the experimenter to preassign data collection strategies, premap sample locations for data collection, and will have the option to use cryogenically cooled samples that may be easily stored and robotically mounted on the microdiffractometer at a moment’s notice.

Materials and Methods

Two types of crystals were selected for screening experiments using the ratiometric mode of data collection: (i) Pol II–TFIIB–NAS, a multiprotein complex, composed of Pol II, transcription factor TFIIB, and a large NAS, and (ii) β2-adrenoreceptor bound to the agonist B1671707 and a nanobody that stabilizes the active state (N6689) (SI Materials and Methods). β2-Adrenoreceptor/nanobody complex crystals were mounted onto micromeshes and flash frozen. A batch of Pol II–TFIIB–NAS crystals were grown inside grids using a specialized crystallization tray and flash frozen. Two types of radiation-sensitive metalloenzyme crystals were selected for data collection:

(i) Cpl (FeFe)-hydrogenase from Clostridium pasteurianum and (ii) sperm whale myoglobin (SI Materials and Methods). Myoglobin crystals were harvested from harnessed blowfish, covered with a thin coat of Paratone-N (Pararad 10132; Hampton Research), and placed inside a grid port, and the grid was flash frozen. Cpl (FeFe)-hydrogenase crystals were grown using the microcapillary batch diffusion method, mounted in cryoloops, and flash frozen (SI Materials and Methods). Diffraction data for all crystals were collected at 100 K using the goniometer-based setup at LCLS XPP (SI Materials and Methods). Datasets were processed using both cctbx.xfel (30) and IPMOSFLM/SCALa (34), and the structures were solved using molecular replacement (SI Materials and Methods).

ACKNOWLEDGMENTS. Portions of this research were carried out at the Linac Coherent Light Source (LCLS), a National User Facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The XPP and CXI instruments were funded through the LCLS Ultrafast Science Instruments project funded by the US Department of Energy, Office of Basic Energy Sciences. Use of the Stanford Synchrotron Radiation Lightsource (SSRL), 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. The SSRt Structural Molecular Biology Program is supported by the US Department of Energy Office of Biological and Environmental Research, and by the National Institutes of Health (NIH), National Institute of General Medical Sciences (including P41GM103993). A.T.B., A.Y.L., M.U., O.B.Z., and W.I.W. were supported in part by a Howard Hughes Medical Institute Collaborative Innovation Award that also provided funds for the purchase of the microcrystalfract.d. N.K.S. acknowledges the support of NIH Grants GM095887 and GM102520 for data-processing efforts. J.M.H. was supported by NIH Grants GM073210, GM082250, and GM094625. This work is supported as a part of the Biological and Electron Transfer and Cata!ysis EfF, an energy Frontiers Research Center funded by the US Department of Energy, Office of Science (DE-SC0012518).

**Supporting Information**

Cohen et al. 10.1073/pnas.1418733111

**SI Materials and Methods**

**SI Microdiffraction**

Core to the microdiffractometer design is an Aerotech ABR1000 air bearing used for the \( \phi \)-rotation axis with rotation speeds of \( >360^\circ/s \). The sample is moved by a range of \( \pm 2.5 \) mm in the \( x \) and \( y \) directions by high-stiffness stages mounted on the air-bearing stage. A lift stage is used for vertical positioning of the \( \phi \) axis to the X-ray interaction region and a linear stage provides horizontal motion by moving the entire lift stage and air-bearing assembly. The horizontal stage offers 10 mm of operational translation (within a sphere of confusion of 1.2 \( \mu \)m) and an overall length adjustment of 100 \( \mu \)m (upgradable to 400 \( \mu \)m) to accommodate long sample holders. During assembly and testing, careful measurement and adjustment of component orientation ensures that the air-bearing \( \phi \) axis, horizontal stage, and sample lift stage motions are in close alignment. Rapid calibration of translations is achieved by home switches or precision absolute encoders. Table S1 lists the operational parameters of the microdiffractometer. The current horizontal positioning stage of the microdiffractometer translates sample holders at a maximum rate of 3 mm/s, enabling it to introduce a new crystal into the X-ray interaction region well before the readout time of the Rayonix MX325HE detector. To optimize thermal stability, the device is contained inside an insulating enclosure (Fig. S1). For measurements at cryogenic temperatures, an on-axis cryocooler nozzle geometry (Fig. 1 and Fig. S1) provides optimal temperature stability along the length of the sample mount. A heated shield just behind the sample holder magnet covers a small hole (20 mm) in the insulated enclosure that serves as a conduit for a zirconium support rod. The support rod, with minimal thermal conductivity, is used to connect the sample magnet and heating shield to the positioning stages. Samples on the microdiffractometer are visualized with an in-line camera system (Fig. 1) with motorized zoom that connects to a 1” \( \times \) 1.4”-diameter fixed elliptical mirror, with a 1.5-mm hole to pass X-rays, that reflects the sample image to a high-resolution camera assembly located below the X-ray path. Samples are illuminated by two sources: a backlight that is pneumatically inserted into the sample mount. A heated shield just behind the sample holder magnet covers a small hole (20 mm) in the insulated enclosure that serves as a conduit for a zirconium support rod. The support rod, with minimal thermal conductivity, is used to connect the sample magnet and heating shield to the positioning stages. Samples on the microdiffractometer are visualized with an in-line camera system (Fig. 1) with motorized zoom that connects to a 1” \( \times \) 1.4”-diameter fixed elliptical mirror, with a 1.5-mm hole to pass X-rays, that reflects the sample image to a high-resolution camera assembly located below the X-ray path. Samples are illuminated by two sources: a backlight that is pneumatically inserted into the sample mount. A heated shield just behind the sample holder magnet covers a small hole (20 mm) in the insulated enclosure that serves as a conduit for a zirconium support rod. The support rod, with minimal thermal conductivity, is used to connect the sample magnet and heating shield to the positioning stages.

**SI Single-Shot Spectrometer**

The photon energy distribution for each X-ray pulse used to collect a diffraction pattern was recorded during the hydrogenase and myoglobin experiments using an energy spectrometer composed of a thin cylindrically bent silicon membrane that directs and disperses a small fraction of the X-ray pulse (\( <1\% \) of the incident intensity) to a Ce:YAG screen (1). A 12-bit 120-Hz CCD camera with microscope optics was used to record the dispersed spectrograph. This method is one of the available options for characterizing the X-ray pulse energy at LCLS XPP (2–5). The spectrometer data were calibrated by cross-referencing its measured XFEL average spectrum with the scanning average measurement using a channel-cut monochromator. A typical spectrum obtained from a single X-ray pulse and typical averaged spectrum using measurements from 120 X-ray pulses is shown in Fig. S2. The single-pulse photon energies used for analyzing still diffraction images was determined by calculating the centroid of the calibrated energy spectrum.

**SI Data Collection Strategy Calculation**

To optimize completeness, the helical data collection mode includes an option to calculate a data collection strategy. A location on the end of the crystal is defined by the experimenter to collect a still diffraction pattern to use for the strategy calculation. The optimal starting orientation and predetermined planarity and calculated using a script running MOSFLM (6) that is based on the WebIce (7) strategy option originally implemented for multocrystal data collection at SSRl BL12-2. A still diffraction pattern (or groups of patterns from oscillation data collection) is analyzed, and the best rotation range to use for data collection calculated, taking into consideration the beam size and the length of the crystal. A strategy interface within Blu-Ice enables data collected from previous crystals to be added to or excluded from the strategy calculation with the purpose to minimize the number of crystals required to obtain a complete dataset (Fig. S4). The resulting strategy is imported into the data collection interface by a push of a button from within the control software.

**SI Sample Mounting Grid**

Each grid is fabricated out of thin (100 \( \mu \)m) polycarbonate plastic, with multiple holes (or ports) that can hold either larger crystals in known locations or groups of smaller crystals. The prototype grid layout (Fig. S5) contains 75 ports: 15 ports of 125- \( \mu \)m diameter, 15 ports of 200- \( \mu \)m diameter, and 45 ports of 400- \( \mu \)m diameter. Custom grids with different hole shapes, sizes, and layouts may be easily fabricated and their layouts incorporated into the data collection software. The grid is compatible with most automated sample mounting robots and can reliably fit inside the port of a SSRL cassette and SAM robot grippers. Grids may be manually filled with crystals and flash frozen. Alternatively, a number of automated methods have been developed to load grids with crystallization solutions and specialized containers have been designed for growing crystals within grids using vapor diffusion techniques or with LCP within glass sandwich plates.

**SI Grid Alignment**

Data collection using grids starts with a semiautomated alignment procedure. After the grid is mounted onto the microdiffractometer, it is rotated edge-on in the in-line camera view, and the edge of the grid is clicked by the experimenter from within the software video display to move it into the X-ray interaction region (and center of rotation of the microdiffractometer). The grid is then rotated by 90°. The centers of four holes, two on each side of the grid, are next clicked in the video display to set the grid rotation and orientation (Fig. S6). From this information, the location of all of the grid holes is calculated and overlaid on the video view of the grid. This process may be completely automated using video analysis to orient the grid and identify the hole locations. Similar strategies may be carried out for other high-density sample containers of consistent dimensions.

**SI Pol II–TFIIB–NAS**

To study the conformational changes that take place during loading of a transcription bubble into the RNA polymerase active site, a full-length DNA scaffold 54 nt long with 25 noncomplementary base pairs (TB-25), constituting an artificial transcription bubble, was assembled into a complex with RNA polymerase II (Pol II)
and general transcription factor IIIB (TFIIB). Pol II and TFIIB were purified as described previously (8). To assemble the Pol II–TFIIB–TB-25 complex, Pol II was first incubated with 2.5 M excess DNA/RNA scaffold in the presence of 1 mM MgCl₂. Size exclusion chromatography (SEC) was used to remove excess DNA/RNA scaffold from the sample. The eluted Pol II–TFIIB fraction was incubated with 2 M excess TFIP in the presence of high salt (1.5 M NaCl) for 60 min followed by a second SEC step to remove excess TFIP and lower the salt concentration to 100 mM NaCl. The presence of the TB-25 was corroborated using ethidium bromide staining. The complex was concentrated to 8 mg/mL for crystallization trials.

Early crystallization condition screening results were examined using a selection protocol combining bright-field microscopy, UV fluorescence microscopy, and dynamic light scattering to detect crystallization drops containing nanometer-sized crystals (NCs); and transmission electron microscopy (TEM) to accurately identify protein NCs and determine NC quality (9). A crystallization condition containing 32% (vol/vol) Tacsimate, 100 mM Hepes, pH 7.5, 10 mM DTT, and 2% (vol/vol) glycerol yielded high-quality NCs (0.5–3 μm in size) and was used to generate several milligrams of NC slurry of ~30% (wt/vol). NCs were used with the GDVN injector for XFEL diffraction screening experiments at LCLS-CXI (10). Diffraction patterns were recorded at 120 Hz using 10-kEV X-rays pulses of 40-fs duration with focus size of 200 nm at the X-ray interaction region and a Cornell-SLAC Pixel Array detector (11). The patterns were analyzed in real time with cctbx.xfel (12). Although the hit rate was low, a few patterns were observed with diffraction to 4.0 Å.

Efforts to produce larger crystals using a NC seeding technique (9) ultimately resulted in the production of larger crystals (of up to about 50 μm). Using this technique, a batch of crystals were grown inside grids using a specialized crystallization tray and flash frozen. During diffraction screening experiments, each grid was mounted on the goniometer using the SAM robot and the semiautomated grid alignment procedure described above was performed. Grids containing crystals were centered in the sample camera view. To collect still diffraction images, crystals were either identified by the experimenter through examination of the sample video display and centered into the X-ray interaction region, or a region of the port was selected for raster data collection. During diffraction screening experiments, each grid was mounted in this manner using the goniometer setup at LCLS-XPP with an X-ray beam size of 20 μm × 20 μm, energy of 9.6 keV, and 30-fs pulse length. Although there was only enough beam time available to screen a handful of crystals, a few of these were observed to diffract to 3.3-Å resolution. A larger number of crystals from the same crystallization batch were subsequently screened in grids at SSRL BL12-2 using a 20 μm × 20 μm or 10 μm × 10 μm beam size. The X-ray dose used to collect a single exposure from fresh crystals was incremented until improvements in resolution were no longer obtained, and observed resolution began to decrease (this corresponded to an optimal dose of about 35 MGy per frame). The best diffraction observed was only 3.7 Å. Diffraction quality screening results are summarized in Table S2. To determine the resolution limit, a peak search using the program ADXV (13) was used to locate diffraction peaks with I/σ of greater than 2. Those at the outer edge of the pattern were then visually examined at multiple locations at high zoom, noting the counts in the peak pixels relative to the background pixels and verifying that the peak fits the crystal lattice. The resolution shell that gave a reasonable number of acceptable peaks was then specified as the maximum resolution for that image.

SI β₂-Adrenoreceptor
Crystals of β₂-adrenoreceptor bound the agonist BI167107 and a nanobody that stabilizes the active state (Nb6B9) were obtained as described previously (14). Crystals ranging in size from 20 to 100 μm were mounted on MiTeGen micromeshes (www.mitegen.com/) and screened at LCLS for diffraction quality by collecting still diffraction patterns using an X-ray beam size of 3 μm × 3 μm or 50 μm × 50 μm and energy of 9.0 keV. Crystals were either identified by the experimenter through examination of the sample video display and centered into the X-ray interaction region, or a region of the port was selected for raster data collection. The crystals used were from the same crystallization batch as those used for the initial structure determination at a synchrotron (14), allowing for direct comparison of their diffraction from these different X-ray sources.

SI Cpf1
Cpf1 [FeFe]-hydrogenase from Clostridium pasteurianum was purified with slight modifications to the previously described methods (15). Q-Sepharose Fast Flow resin was substituted for DEAE resin, and Sephacryl S-200 resin was substituted for Sephadex G-100 resin; the hydroxypatite step was excluded. Crystals were grown using the microcapillary batch diffusion method with a precipitating solution containing 25% (wt/vol) PEG 4000, 0.1 M sodium acetate, 0.2 M ammonium sulfate, and 1 mM sodium dithionite to remove trace oxygen (16). A protein solution containing 30 mg/mL protein, 50 mM Tris (pH 8.0), 0.2 M KCl, and 1 mM sodium dithionite was allowed to gradually equilibrate with the precipitating solution, and large brown crystals (≥400 μm) were observed in approximately 2 wk. All sample manipulations were carried out under anaerobic conditions in an MBraun Unilab glove box with 100% nitrogen atmosphere. Before data collection, Cpf1 crystals were cryoprotected by equilibrating them in crystallization solution supplemented with 14% (vol/vol) glycerol in the presence of 1 mM sodium dithionite and then subsequently flash cooled in liquid nitrogen on large nylon loops (0.8–1.0 mm).

Data collection was performed using the helical data collection mode, with an unattenuated 50-μm beam size and a photon energy of 9.017 eV. Five crystals, with lengths ranging from 1 to 3 mm, were used for data collection. Crystals were translated by 70 μm and rotated by 0.5° between exposures, and a total of 162 images were collected. Eleven degrees of oscillation data were collected at the first position exposed on the crystal. Still data were processed using both IPMOSFLM/SCALA and cctbx-xfel and data processing statistics are given in Table S3. The available version of the program XDS (17) could not be used to process the still diffraction patterns, but was used to analyze the oscillation images corresponding to each crystal. In some cases, for example when the crystal was very mosaic or split at the first position exposed, IPMOSFLM was instead used to index and postrefine a group of still images from the same crystal.

For the IPMOSFLM (6) protocol, postrefined orientation and experimental parameters obtained from the oscillation data were used as input for integration of the still diffraction patterns using a fake oscillation angle of 0.0°. The crystals were processed individually and visual examination of images that had large scale factors confirmed that the corresponding areas of the crystals were of poor diffraction quality. These images were manually rejected before scaling and merging, producing a final dataset comprising 110 images. For cctbx-xfel processing, all 162 images were used, and 4 were rejected by the program. To check the quality of the data, both datasets were scaled to a synchrotron dataset collected from a single Cpf1 crystal from the same crystallization experiment. The CCP4 program SCALEIT (18) was used to compare the datasets giving a R factor of 30% and a weighted R factor of 13%. Although it was found that the datasets were essentially isomorphous with the synchrotron data, there were some differences that could possibly be attributed to the presence of radiation damage in the case of the synchrotron dataset, compared with the essentially damage-free femtosecond dataset in the case of the LCLS data.

The Cpf1 structure was solved by molecular replacement (MR) using the program MOLREP with both the MOSFLM/SCALA
and cctbx-xfel datasets. The starting model for MR was the *C. pasteurianum* [FeFe]-hydrogenase (PDB ID code 3CSY) with the Fe-S clusters and associated ligands removed. The MR solution was refined for 15 cycles of XYZ and B refinement using the program REFMAC (19). MR and refinement statistics are listed in Table S3. The $2F_o - F_c$ and $F_o - F_c$ electron density maps calculated following refinement showed the location of the Fe-S clusters and the Fe-Fe prosthetic groups (Fig. 5A and B). Peak heights for the iron atoms were between 10 and 11 $\sigma$, and the peak heights for the sulfur atoms in the Fe-S cluster were between 4.5 and 5.5 $\sigma$. To check the quality of the data and the subsequent electron density maps, several contiguous regions of the Cpl polypeptide chain, amounting to about 5% each time, were removed and the truncated model submitted to simulated annealing refinement using PHENIX, and omit electron density omit calculated. These maps, contoured at 1.0 $\sigma$, showed the presence of the omitted sections of polypeptide.

**SI Myoglobin**

Crystallization of recombinant sperm whale myoglobin in the space group P6 was performed as previously described (20), with the following modifications: the drop was allowed to equilibrate for 1 day; 1,000-fold-diluted solution was used to seed crystals; and the well solution did not include EDTA. Myoglobin crystals were harvested from hanging drops. The grid was covered with a thin coat of Paratone-N (Parabar 10312; Hampton Research). Individual crystals were picked out of the drop using a nylon loop tool, placed inside a drop of Paratone-N, and then placed inside a grid port. The size of the crystal and the size of the grid port were matched as closely as possible.

Data were collected using a 50-µm focus at a nominal energy of 9.5 keV from crystals within 32 grids with the following protocol: (i) the port containing the crystal was centered and single X-ray pulse was used to collect a single diffraction pattern. (ii) Eleven degrees of oscillation data were collected around the still image location using multiple attenuated pulses (Fig. S7). (iii) A second still diffraction pattern was collected. The energy spectrum was measured for each pulse used for data collection, although thus far only the mean pulse energy has been used for data processing. Data from a total of 932 crystals were collected.

Because P6 is a polar space group, the diffraction pattern can be indexed in two different indexing schemes, with the sixfold rotation axis pointing in a positive or negative direction. Attempts to merge data from different crystals arbitrarily indexed will result in apparent merohedral twinning with the winning operation k.h.$^{-1}$. Although the Brehm and Diederichs algorithm (21) has been developed to solve the indexing ambiguity for data consisting of single still diffraction patterns from multiple crystals, it was not available at the time of the experiment. However, using the oscillation data (Fig. S7) that were collected around the position of each still pattern collected, it was quite straightforward to determine the indexing relationship between each of the crystals and that of an arbitrarily chosen reference by comparing the intensities of the reflections related by the above-mentioned operation with the program POINTLESS (22). The oscillation data could be processed like conventional oscillation data, using the program XDS (17) for autoindexing, integration, and postfine-


Cohen et al. www.pnas.org/cgi/content/short/1418733111


Fig. S1. Photograph of the diffractometer-based instrumentation installed at LCLS-XPP.
**Fig. S2.** Heat map of detector information and 2D graph of example spectrometer data. (A) Measured spectrum of a single LCLS self-amplified spontaneous emission pulse. (B) The average spectrum of 120 X-ray pulses for the collection of a single degree of oscillation data.
Fig. S3. Screen capture images of the Blu-Ice GUI illustrating the raster interface for specifying the area of the sample use for data collection. (A) Rastering the entire circular area of a grid hole. (B) Specific areas of the hole are selected for data collection. (C) Interface for specifying collection parameters.

Fig. S4. Screen capture images of the Blu-Ice GUI illustrating the helical data collection mode. (A) Interface for specifying collection parameters; (B) video display of a Cpl crystal. The yellow line defines the crystal length and the overlaid boxes depict the locations for future exposure to a single X-ray pulse. (C) Interface for multicrystal data collection strategy determination.
Fig. S5. Schematic of the polycarbonate high-density sample mounting grid.

Fig. S6. Screen capture images of the Blu-Ice GUI illustrating the semiautomated grid alignment process. (A) The grid is aligned to be edge-on to the in-line camera view and the edge of the grid is moved into the X-ray beam position. (B) Two specific holes on one side of the grid are selected by clicking on their location inside the video display. (C) The grid is translated and the location of two holes on the opposite side of the grid are clicked next. (D) shows the completed calibration of the grid orientation in relation to the diffractometer axes and X-ray beam position.
Table S1. The motion performance of the microdiffractometer

<table>
<thead>
<tr>
<th>Axis</th>
<th>Range</th>
<th>Speed</th>
<th>Encoder/homing resolution</th>
<th>Motion increment</th>
</tr>
</thead>
<tbody>
<tr>
<td>φ</td>
<td>Infinite</td>
<td>90°/s</td>
<td>0.1 milli-degrees</td>
<td>0.1 milli-degrees</td>
</tr>
<tr>
<td>Sample-x</td>
<td>2.5 mm</td>
<td>0.25 mm/s</td>
<td>0.1 μm</td>
<td>0.06 μm</td>
</tr>
<tr>
<td>Sample-y</td>
<td>2.5 mm</td>
<td>0.25 mm/s</td>
<td>0.1 μm</td>
<td>0.06 μm</td>
</tr>
<tr>
<td>Sample-horz</td>
<td>100 mm</td>
<td>3 mm/s</td>
<td>0.2 μm</td>
<td>0.09 μm</td>
</tr>
<tr>
<td>ϕ-vertical</td>
<td>13 mm</td>
<td>0.05 mm/s</td>
<td>0.2 μm</td>
<td>0.09 μm</td>
</tr>
</tbody>
</table>

Table S2. Comparison of diffraction quality screening results for Pol II-TFIIB-NAS

<table>
<thead>
<tr>
<th>Largest crystal dimension, μm</th>
<th>Location</th>
<th>Delivery method</th>
<th>Best resolution observed, Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>LCLS-CXI</td>
<td>GDVN injector at RT</td>
<td>4.0</td>
</tr>
<tr>
<td>&lt;50</td>
<td>SSRL BL12-2 Grids</td>
<td>At 100 K</td>
<td>3.7</td>
</tr>
<tr>
<td>&lt;50</td>
<td>LCLS-XPP</td>
<td>Grids at 100 K</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Table S3. Clostridium pasteurianum hydrogenase data processing

<table>
<thead>
<tr>
<th>Processing software</th>
<th>IPMOSFLM/SCALA</th>
<th>cctbx-xfel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection/processing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell dimensions, Å*</td>
<td>a = b = 111.28, c = 103.81</td>
<td>a = b = 110.08, c = 103.01</td>
</tr>
<tr>
<td>Resolution limits, Å</td>
<td>38.0–1.60</td>
<td>44.7–1.60</td>
</tr>
<tr>
<td>No. of images</td>
<td>110</td>
<td>158</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>93.9 (88.0)</td>
<td>91.7 (62.4)</td>
</tr>
<tr>
<td>CC1/2†</td>
<td>0.48 (0.09)</td>
<td>0.39 (0.23)</td>
</tr>
<tr>
<td>Reflections, observed/unique</td>
<td>408,563/80,047</td>
<td>663,856/76,845</td>
</tr>
<tr>
<td>Molecular replacement and refinement‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFİo, background</td>
<td>9.9 (4.0)</td>
<td>11.4 (3.1)</td>
</tr>
<tr>
<td>Weighted R factor, %</td>
<td>42.8</td>
<td>42.2</td>
</tr>
<tr>
<td>Score§</td>
<td>0.50</td>
<td>0.51</td>
</tr>
<tr>
<td>Rwork, %</td>
<td>39.2 (34.6)</td>
<td>39.9 (35.8)</td>
</tr>
<tr>
<td>Rfree, %</td>
<td>38.1 (38.9)</td>
<td>40.2 (38.7)</td>
</tr>
<tr>
<td>FOM#</td>
<td>0.63</td>
<td>0.55</td>
</tr>
</tbody>
</table>

*Space group P4_2_2_2.
†The correlation coefficient between half datasets is defined in ref. 1.
‡Using MOLREP for molecular replacement and REFMAC for 15 cycles of refinement.
§Rotation function peak height from MOLREP.
#The score is the product of the Correlation Coefficient of intensities and the maximal value of the Packing Function.
#The FOM is the overall figure of merit of the phases.

### Table S4. Myoglobin data processing

<table>
<thead>
<tr>
<th>Processing software</th>
<th>Single-pulse data</th>
<th>Multiple-pulses oscillations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection/processing</td>
<td>IPMOSFLM/SCALA</td>
<td>cctbx-xfel</td>
</tr>
<tr>
<td>Resolution limits, Å</td>
<td>39.15–1.40</td>
<td>39.15–1.36</td>
</tr>
<tr>
<td>No. of images</td>
<td>253</td>
<td>739</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>91.6 (45.7)</td>
<td>91.6 (31.0)</td>
</tr>
<tr>
<td>CC1/2*</td>
<td>0.767 (0.30)</td>
<td>0.956 (0.752)</td>
</tr>
<tr>
<td>Reflections, observed/unique</td>
<td>560,156/37,257</td>
<td>1,299,845/43,737</td>
</tr>
<tr>
<td>Molecular replacement and refinement†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFree/S, background</td>
<td>7.26</td>
<td>8.58</td>
</tr>
<tr>
<td>Weighted R factor, %</td>
<td>35.9</td>
<td>48.4</td>
</tr>
<tr>
<td>Score§</td>
<td>0.577</td>
<td>0.664</td>
</tr>
<tr>
<td>Rwork, %</td>
<td>28.5 (45.0)</td>
<td>15.9 (38.5)</td>
</tr>
<tr>
<td>Rfree, %</td>
<td>31.5 (46.0)</td>
<td>18.8 (38.9)</td>
</tr>
</tbody>
</table>

The values in parentheses correspond to the highest resolution bin.
*The correlation coefficient between half datasets is defined in ref. 1.
†Using MOLREP for molecular replacement and Phenix for refinement.
‡Rotation function peak height from MOLREP.
§The score is the product of the Correlation Coefficient of intensities and the maximal value of the Packing Function.