Published in final edited form as:
Methods. 2011 December ; 55(4): 387–404. doi:10.1016/j.ymeth.2011.12.006.

Toward structure determination using membrane-protein nanocrystals and microcrystals

Mark S. Hunter[a,*,I] and **Petra Fromme**[a]

[a]Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ USA 85287

Abstract

Membrane proteins are very important for all living cells, being involved in respiration, photosynthesis, cellular uptake and signal transduction, amongst other vital functions. However, less than 300 unique membrane protein structures have been determined to date, often due to difficulties associated with the growth of sufficiently large and well-ordered crystals. This work has been focused on showing the first proof of concept for using membrane protein nanocrystals and microcrystals for high-resolution structure determination. Upon determining that crystals of the membrane protein Photosystem I, which is the largest and most complex membrane protein crystallized to date, exist with only a hundred unit cells with sizes of less than 200 nm on an edge, work was done to develop a technique that could exploit the growth of the Photosystem I nanocrystals and microcrystals. Femtosecond X-ray protein nanocrystallography was developed for use at the first high-energy X-ray free electron laser, the LCLS at SLAC National Accelerator Laboratory, in which a liquid jet brought fully-hydrated Photosystem I nanocrystals into the interaction region of the pulsed X-ray source. Diffraction patterns were recorded from millions of individual PSI nanocrystals and data from thousands of different, randomly oriented crystallites were integrated using Monte Carlo integration of the peak intensities. The short pulses (~ 70 fs) provided by the LCLS allowed the possibility to collect the diffraction data before the onset of radiation damage, exploiting the diffract-before-destroy principle. During the initial experiments at the AMO beamline using 6.9-Å wavelength, Bragg peaks were recorded to 8.5-Å resolution, and an electron-density map was determined that did not show any effects of X-ray-induced radiation damage [Chapman H.N., *et al.* Femtosecond X-ray protein nanocrystallography, Nature 470 (2011) 73–81]. Many additional techniques still need to be developed to explore the femtosecond nanocrystallography technique for experimental phasing and time-resolved X-ray crystallography experiments. The first proof-of-principle results for the femtosecond nanocrystallography technique indicate the incredible potential of the technique to offer a new route to the structure determination of membrane proteins.

Keywords

membrane proteins; structure determination; femtosecond nanocrystallography; protein nanocrystals; X-ray crystallography; XFEL

*Corresponding author, hunter33@llnl.gov.
I Current address: Lawrence Livermore National Laboratory, 7000 East Avenue, Mail Stop L-211, Livermore, California 94551, USA

1. Introduction

1.1. Structural biology and the quest to solve life's mysteries

The structure and function of molecules are intimately related at the atomic and molecular levels. The desire to determine the structure of matter in order to infer the mechanisms and pathways has been a large motivational force in the disciplines of physics, chemistry, and biology. For instance, the structures of inorganic and organic molecules offer insights into catalysis and reaction pathways, and the information can be used to elucidate the action of drugs [1], or the potential environmental impact of an organometallic molecule [2], as two examples.

The biomacromolecules are the facilitators of life. Thereby the structure determination of biomacromolecules is one important clue to help understand the complexity observed in life. Whether the desire is to understand cellular respiration, nutrient uptake and transport by a cell, or various other cellular functions, one focus is to determine high-resolution structures of the molecules involved. The structure forms the basis to elucidate the reaction mechanisms and understand how the structure relates to the function and the dynamics of the molecules. Studying the structures of proteins by X-ray crystallography [3, 4], electron microscopy [5], electron crystallography [6], and nuclear magnetic resonance (NMR) [7, 8] is well established, with the majority of the 70,000+ structures determined by protein X-ray crystallography [9].

The impact of structural biology on the biological sciences becomes tangible when noting the examples of success for structure determination. One area of study that has seen considerable success through the use of structural biology is oxygenic Photosynthesis [10]—proteins involved in Photosynthesis are also a primary focus in the development of methods for femtosecond nanocrystallography. One breakthrough of X-ray protein crystallography was the structure determination of cyanobacterial Photosystem I [11].

1.1.1. Structure of cyanobacterial Photosystem I—Photosystem I (PSI) is a large membrane-intrinsic protein complex that is responsible for light-induced charge-separation, which transfers an electron from the luminal side to the stromal side of the thylakoid membrane. The electron transfer catalyzed by PSI provides the electrons utilized in reducing $NADP^+$ to NADPH in oxygenic photosynthesis [12].

PSI exists in a number of oligomeric forms in nature. In plants, PSI exists as a monomer surrounded by four antenna proteins LHC-(I–IV), whereas in cyanobacteria, PSI exists predominantly as a trimer, but the monomer can be found when cells are grown under high-light conditions; the trimer-tomonomer ratio depends on the light conditions [13]. Cyanobacterial PSI is one of the largest and most complex membrane-protein complex to be crystallized [14] and its structure has been determined to 2.5-Å resolution [11].

The trimeric PSI, as shown in Fig. 1, has a mass of 1,056,000 Da, with each monomer consisting of 12 protein subunits and 127 non-covalently bound cofactors. Trimeric PSI has a diameter of approximately 220 Å and a height of approximately 90 Å, with a stromal hump that extends 40 Å into the stroma of the chlorplasts/ cytoplasm of the cyanobacteria. The cofactor composition in the monomer is 96 chlorophyll molecules, 22 carotenoid molecules, three 4Fe-4S clusters, three lipids, two phylloquinone molecules, and one Ca^{2+} ion [11].

Overall, the structures of the major proteins involved in oxygenic photosynthesis have been solved using X-ray crystallography. X-ray crystallography does not have size restrictions on the protein of interest, unlike NMR and single-particle electron microscopy, which is one of

the reasons for its ubiquitous role in structural biology [15]. Additionally, X-ray crystallography has proven to be an invaluable tool for the determination of the structure of many small molecules, and its extension to the large macromolecules of the cell was inevitable.

1.2. X-ray protein crystallography

1.2.1. Conventional X-ray protein crystallography—X-ray crystallography is the workhorse of structural biology, having accounted for the majority of the biomacromolecular structures determined to date. The power of X-ray crystallography resides in the ability of reproducibly ordered molecules in a lattice to scatter electromagnetic radiation coherently. Coherent scattering implies that the scattered waves have a defined phase relationship, which allows for the addition of the amplitudes of the scattered waves, creating a pronounced effect for the scattered signal [16].

The scattering of X-rays by molecules depends on the atomic number of the atoms incorporated, which corresponds to the number of electrons present. In the forward direction, the limit of the atomic scattering factor as the scattering approaches zero is Z [16]. However, biological macromolecules are composed mostly of H, C, N, O and minor amounts of S and P, with relatively low atomic numbers, causing the scattered intensity to be small. Additionally, unlike in the case of small inorganic or organic molecules, crystals of biomacromolecules contain a large percentage of water, commonly between 30 and 70% of the unit cell volume [17], creating large unit cells with few crystal contact sites. Consequently, in order to record X-ray diffraction to large angles and high resolution, large, well-ordered crystals of proteins are necessary for protein-crystallography experiments, due to the weak scattering of the constituent elements as well as the large unit cell solvent contents.

Photosystem I is used as a model protein for the work presented in this review, and as an example of the power and perils of X-ray crystallography, consider the time associated with the structure determination of PSI from the cyanobacterium *Thermosynechococcus elongatus*. The first micron-sized crystals of PSI extracted from *T. elongatus* were reported in 1988 [18]. The first structural model of PSI, based on crystal diffraction to 6-Å resolution, was determined in 1993 [19, 20], followed by a 4-Å structure in 1996 [21] and an improved structure at 4-Å in 1999 [22, 23]. In 2001, the structure of Photosystem I was unraveled at 2.5-Å resolution [11] from cryogenically cooled crystals that were incubated in sucrose before freezing. Therefore, it took 13 years from the growth of the first microcrystals to the determination of the first near-atomic-resolution structure of Photosystem I based on large, well-ordered single crystals under cryogenic conditions [12].

However, PSI is by no means the only membrane protein that has presented challenges during structural determination. The membrane-protein complexes involved in photosynthesis and respiration are studied using X-ray crystallography, and are examples of extremely complex, molecular machines; the proteins involved in these electron-transport chains include the b6f complex [24], the cytochrome c oxidase [25, 26, 27], the bc_1 complex [28, 29], Photosystem II [30, 31, 32], the LHCII [33], and the plant PSI-LHCI complex [34, 35, 36], amongst others. As is the case with PSI, the structures of these proteins have been solved (if solved) using X-ray crystallography by efforts that extend over more than a decade.

Photosystem I remains the largest membrane-protein complex solved to molecular resolution, and the use of X-ray crystallography to determine the structure of such a complicated molecule is a *tour de force* for the technique. However, the work on the project took the tireless effort, of many dedicated researchers, for over a decade.

A question that needs to be addressed pertains to the causes of the structure determination of a protein taking *a decade*, or longer. In order to address this question, the difficulties with X-ray protein crystallography will need to be discussed from the general point of view. Besides the problems associated with expression and purification of proteins, the major difficulties of X-ray crystallography are: the growth of large and simultaneously well-ordered crystals of proteins, X-ray-induced radiation damage, and the cryogenic cooling of protein crystals.

1.2.2. Protein crystallization and the necessity for large crystals—Obtaining a suitable crystal for X-ray protein crystallography is the least understood step of the structure determination of a protein [3]. The various properties of proteins allow for more complicated mechanisms of crystal growth with respect to small molecules. Protein molecules are much more complicated in regard to composition, structure, degrees of freedom, surface properties, as well as surface potential, than small organic molecules. Additionally, due to the much larger sizes, proteins have far different transport properties in solution, causing slower crystal growth relative to small organic crystals [37].

Protein crystal nucleation and growth are driven by supersaturation of the protein in solution. One of the difficulties associated with protein crystallization is the need for a much larger supersaturation—the supersaturation is defined as $\ln(c/s)$, where c is the concentration and s is the concentration at saturation—than what is typically needed in the case of small organic molecules [37]. The high supersaturation is necessary for the formation of the critical nuclei of the new phase, but high supersaturation is not desirable for the growth of large, well-ordered single crystals. Additionally, multiple protein-replete phases are possible in the case of proteins, such as amorphous precipitates, that are not ordered. High supersaturation will favor the formation of amorphous precipitates, as opposed to crystals due to kinetic factors [38]. An additional result of the high supersaturations used for protein crystallization is that the growth of protein microcrystals is much more common than the growth of macroscopic protein crystals [15]. The difficulties associated with the growth of protein crystals are major issues for protein crystallography because of the weak scattering of the constituents; therefore, much work and extensive optimization is often necessary to grow large protein crystals (if conditions are found at all).

In general, membrane proteins provide greater resistance to structure determination than their soluble counterparts due to difficulties associated with crystallization; although the crystallization of membrane proteins is, in principle, similar to that of soluble proteins, a major difference is that membrane proteins are amphiphilic, having both hydrophobic and hydrophilic domains [39]. Despite their extremely high impact, only four medically relevant human membrane-protein structures have been determined to date, that of a G protein-coupled receptor, the β_2-andronergic receptor [40, 41, 42], Aquaporin-5 [43], and leukotriene C_4 synthase [44], and the dopamine receptor [45]. Much work has been devoted to improving membrane-protein crystallization, including the use of *in meso* crystallization matrices [46, 47], such as the lipidic cubic phase [48] or the sponge phase [49, 50], and the use of amphipols [51], amongst others. The use of *in meso* crystallization matrices has shown much promise as an adaptable media for membrane-protein crystallization [52] and the growth of well-ordered membrane-protein crystals, although challenges remain in the optimization of the growth of membrane-proteins that are suitably large and well ordered for conventional X-ray crystallography.

Even when a large protein crystal grows, the crystal will often suffer from large internal disorder, as measured by the mosaicity. In the mosaic model of crystals, a crystal is considered to be composed of smaller mosaic domains, and mosaicity, or mosaic spread, is a term used to describe the degree of angular misalignment of mosaic domains within a crystal [16]. Working from the mosaic domain model, Bragg reflections can be broadened by the

finite size of mosaic blocks, the angular misalignment of mosaic blocks, and additionally by variations in the unit cell parameters in different mosaic blocks [53]. However, the mosaic domain model is an incomplete picture in the case of membrane proteins, and if PSI crystals are considered, it is seen that only four salt bridges make up the crystal contacts, which would be a primary cause of misalignment. Misalignment of the unit cells could also occur due to excess surface area with respect to volume, because the molecules at the interface will have a reduced number of crystal contacts. Therefore, two effects can be seen as influencing the mosaicity of the crystal: for large crystals, it is the propagation of slight misalignments in different domains, whereas for small crystals, the increased surface area with respect to volume causes misalignment of a large number of molecules. Regardless of the cause of the internal disorder, the broadening of the reflections by mosaic spread causes the scattered intensity to be subtended over a larger solid angle as shown in Fig. 2. The broadening can result in severe consequences, such as being unable to resolve the high-resolution diffraction spots, making it impossible to evaluate the structure to the maximum resolution recorded in the diffraction pattern.

The theoretical dependence of the scattered intensity on the size of a crystal can be readily understood though the kinematic theory of diffraction. The Darwin equation can be used to calculate the integrated scattered intensity in an X-ray crystallography experiment, when using the oscillation method, and shows the scattered intensity with the following form [54]:

$$I = I_0 r_e^2 \frac{\left(1 - \cos^2(2\theta)\right) \lambda^3 V_x}{2\sin(2\theta) \quad \omega V^2} |F_{hkl}|^2 \tag{1}$$

in which I_0 is the incident intensity, r_e^2 is the classical electron radius, $(1-\cos^2(2\theta))$ is the polarization factor and assumes unpolarized radiation, $2\sin 2\theta$ is the Lorentz factor, λ is the wavelength, ω is the angular speed of rotation, V_x is the crystal volume, V is the unit cell volume, and F_{hkl} is the structure factor associated with the particular reflection. From Eq. 1, it becomes clear that the integrated intensity is proportional to the crystal volume relative to the unit cell volume. Consequently, the larger the crystal, without regarding the internal order, the stronger the scattered intensity.

As mentioned already, protein crystallization experiments may result in the production of a shower of microcrystals, without subsequent optimization experiments producing the large, well-ordered protein crystals necessary for data collection at conventional protein crystallography beamlines [55]. However, smaller protein crystals can be beneficial when issues such as high mosaicity or twinning plague the data [56], but the tradeo is weaker scattering. In addition to the weaker scattering provided by a smaller protein crystal, if the crystal is smaller than the X-ray beam focus, then an increased background relative to the signal will also be recorded. One technology development that has increased the usefulness of protein microcrystals for structure determination is the microfocussed beamline [57].

The microfocussed beamline compensates for the weaker scattering of a smaller crystal by increasing the flux density at the sample, I_0 on the RHS of Eq. 1. A $(20\text{-}\mu m)^3$ crystal will scatter 1000× weaker than a $(200\text{-}\mu m)^3$ crystal, which could be compensated for through a 1000× increase in the flux density on the sample by decreasing the focus spot radius by a factor of $1000^{1/2} \approx 31.6$. The reduced beam focus will also result in a smaller background scattering from the mother liquor. However, radiation damage becomes a limiting factor in data collection at microfocused beamlines, as discussed in the following section.

1.2.3. Radiation damage and the unending quest to mitigate it—X-ray diffraction is inherently a destructive imaging technique because X-rays are a form of ionizing radiation. Unfortunately, the ionizing nature of the X-rays leads to radiation damage, which

along with sample heterogeneity ultimately limits the resolution of a biomacromolecular structure determined using X-ray crystallography [58]. The problem that ultimately arises with the use of conventional X-ray sources to determine a structure is that the smaller the object one wishes to determine the structure of, the smaller the scattering cross-sectional area. Therefore, to obtain enough scattered intensity to high resolution, the total fluence incident on the sample must be increased [59], increasing the risk of X-ray-induced radiation damage and ultimately decreasing the effective resolution of the reconstructed object [60].

There are two types of X-ray-induced radiation damage during protein crystallography: global damage and local damage [61]. Global radiation damage appears quantitatively in the data set and is not associated with any one particular element of the electron density, and can become manifest as a reduced total scattered intensity, increased unit cell parameters, and increased mosaicity of the crystal (or some combination), amongst other effects, but has a dependence on the experimental conditions [58]. Local damage is damage affecting specific sites such as amino acids within the protein structure itself, and is often ascertained when looking at the electron density maps of the structure of interest [61]. An interesting use was found for site-specific radiation damage, in which the information is used to solve the crystallographic phase problem using a technique called Radiation-damage Induced Phasing (RIP) [62].

When protein crystals are irradiated by X-rays, an X-ray can either interact with the sample, or it can pass through without being affected by the presence of the sample. In the case of 12.4 keV X-rays ($\lambda = 1$ Å), the wavelength typically used for macromolecular crystallography, only approximately 2% of incident X-rays will interact with a 100-μm thick crystal. Of the interacting 2%, 84% will interact through the photoelectric effect, causing the ejection of photoelectrons. Only 8% of the interacting X-rays will produce elastic scattering events, whereas the remaining 8% will produce inelastic scattering [63]. The relatively low ratio of the scattering-cross-sectional area of the light elements for X-rays with respect to photo-ionization and inelastic scattering is one of the major weaknesses of X-ray crystallography with respect to related electron microscopy techniques [64].

The primary radiation damage occurs due to the inelastic events, *i.e.* the photoelectric effect and Compton scattering. The photoelectric effect will cause the ejection of energetic (tens of keV) electrons from the K-shells of many of the light elements (and of other shells in heavier metals). The energy deposited by the primary events will cause a cascade of additional photoelectrons, with energies of a few to tens of eV. Possible effects of the primary and secondary damage are the breakage of chemical bonds, oxidation-reduction reactions, the generation of free radicals, and the production of gaseous species. The accumulation of all of the damage at the microscopic level can lead to strain in the crystal and cause distortion of the long-range order of the molecules in a crystal, leading to tertiary damage, or global damage [58].

Cryogenic cooling of protein crystals was developed [65] to mitigate X-ray-induced radiation damage. The vitreous-solvent matrix that is formed does not reduce the number of radicals, but limits the diffusion of the destructive radicals throughout the unit cell and crystal [64]. However, once the use of 3rd generation synchrotrons became more widespread for structural biology, the high X-ray doses lead to significant damage even under cryogenic conditions; therefore, data sets were still being recorded that were produced from damaged species. That the damage still occurred at cryogenic temperatures could be attributed to the high amount of radicals produced; ultimately there is an absolute limit to the dose a protein sample can tolerate, after which it requisitely deteriorates due to the detrimental effect of the large amount of radicals. The effects of the damage would ultimately be manifest in the degradation of the diffraction pattern, which is the global damage [66]. From cryogenic

electron microscopy, the tolerable dose for a protein crystal—defined as the absorbed energy necessary to reduce the total scattered intensity to a value of 0.5 the initial value—was calculated to be 20 MGy [67]. More recently using synchrotron radiation, the number was empirically determined to be 42 MGy [68], but the authors recommended that data collection should cease once the scattered intensity decreased by a factor of ln(2) = 0.693 from the initial scattered intensity, which corresponded quite well with a 30 MGy limit.

Unfortunately, specific local damage, situated on individual amino acid residues, or at specific sites, can start well before the global damage thresholds are reached [58, 68]. In the case of Photosystem II, individual structure elements that are prone to oxidation-reduction chemistry—one prominent example is the Mn_4CaCl of Photosystem II that catalyzes the oxidation of water—could be damaged with much lower doses [69]; thereby, the allowed dose will depend very specifically on the biomacromolecule of interest; metal containing proteins; proteins with more radiation-damage susceptible amino acids in the catalytic site, solvent exposed, or at crystal contacts; as well as proteins with solvent exposed disulfide bridges, should have the data collection strategy carefully considered at reduced dose limits [58]. A (very general) rule of thumb for proteins that can be used is that for every 1 Å of resolution in the structure, 10 MGy can be absorbed, such that for a 3-Å structure, 30 MGy can be absorbed [59]. It should be noted, however, that photoreduction of metals can occur from a dose that is one to two orders of magnitude lower than the 30 MGy limit [69], and consequently, the rule of thumb from above should be applied cautiously and only for proteins that do not contain significant amounts of metals or cofactors that would increase the damage.

Ultimately, the use of cryoprotectants to avoid cubic ice formation can prove problematic. The addition of the cryoprotectant will change the properties of the solution and could lead to dissolution of the protein crystal, or in some cases, lead to the cracking or breaking of a protein crystal. A suitable cryo-solution must first be found for the crystal, typically the mother liquor plus an anti-freeze chemical such as glycerol, sucrose, and the lighter PEGS, amongst others; the anti-freeze agent should not dilute the concentration of the precipitant components, but replace the water of the solution instead [70]. The concentration of the anti-freeze agent must be tested to determine the necessary concentration to avoid cubic ice formation, which involves the collection of diffraction patterns of the cryogenically cooled solution in a loop [71]. If the protein crystal is not grown in conditions with suitable levels of cryoprotectants in the motherliquor, the time allowed for incubating the crystal in the cryo-solution must be determined empirically. Short incubations-such as through dragging a crystal through a cryo-solution-can minimize the degradation of the crystal by the cryoprotectant but may not allow for osmotic equilibration, whereas longer incubation times allow for osmotic equilibration but also have the potential for more damage to be done to the crystal by the cryoprotectant [70]. As seen from the above discussion, determining a proper cryogenic cooling protocol may be a daunting undertaking.

As can be seen, much effort is devoted to the minimization of radiation damage. Cryogenic cooling of protein crystals and the use of microfocused beams can allow for less of the crystal to be inundated with damaging X-rays, which allows for the frequent shifting of the crystal during data collection to unexposed regions. In sum, the main difficulty with X-ray protein crystallography emanates from the fact that X-ray diffraction is inherently a destructive structural determination technique.

1.3. Objective and Hypotheses for Nanocrystallography

The predominant hypothesis of the new concept of nanocrystallography was that the membrane protein microcrystals, which are more commonly found in protein crystallization screens than protein macrocrystals, could be used to determine high-resolution structures of

membrane proteins. However, the membrane protein microcrystals could not be used with the currently available third-generation X-ray sources in a traditional single-crystal experiment, as X-ray-induced radiation damage becomes the major problem. Consequently, the main objective of the nanocrystallography project is to determine whether microcrystals of membrane proteins can be used to obtain molecular, or even atomic, resolution diffraction patterns. For the work, the intrinsic membrane protein Photosystem I was chosen as a model system.

1.3.1. Challenges Associated with Nanocrystallography—Questions and challenges that immediately arise from consideration of nanocrystallography involve protein crystallization, sample introduction, data collection, data analysis, as well as radiation physics and chemistry. Decades of work on protein crystallization were done to take the crystallization conditions that produced microcrystals and to generate macrocrystals from adjacent areas of the crystallization phase diagram. One major challenge was in determining the reproducibility and quality of protein microcrystals and to investigate the production of protein microcrystals. Since microcrystals of proteins were always (only) considered a stepping stone toward the generation of macrocrystals, the prevalence of formation was never thoroughly evaluated. An additional question to address was how small the membrane protein crystals can be, while still exhibiting enough order to be useful in a crystallographic experiment, and how could the nanocrystals be observed and measured. Additionally, the data quality of diffraction patterns and the merging of data from different crystals needed to be assessed. Most importantly, the way to arrive at high-resolution data that was not encumbered by data from damaged samples needed new, innovative ideas.

sectionPreparation of nanocrystals and microcrystals of PSI One of the first experiments initiated was to test the production of microcrystals of PSI, in order to gain insight into the prevalence, reproducibility, and quality of membrane-protein microcrystals.

1.4. Photosystem I purification and crystallization

Photosystem I (PSI) is isolated from the thermophilic cyanobacterium *Thermosynechococcus elongatus* and crystals are grown as described previously [14]. The cells are grown under low-light conditions allowing for a higher yield of trimeric PSI. The cells are harvested and lysed using a microfluidizer, and the PSI-containing thylakoids are cleaned by three centrifugation washing steps. The protein is solubilized by addition of 0.6% (m/v) N-dodecyl-beta-maltoside (βDDM) and purified using ultracentrifugation and anion-exchange chromatography. The eluent solutions for the chromatography consist of 20 mM MES pH 6.4, 0.02% (m/v) βDDM, and concentrations of $MgSO_4$ varying between 100 and 150 mM.

Photosystem I is crystallized by decreasing the ionic strength of the solution. This method is often referred to as the reverse of ”salting in” of the protein. The decrease of the ionic strength leads to a depletion of counter ions from charged groups at the surface of the protein, which lowers the solubility of the protein. Under low ionic-strength conditions, direct contacts of negatively and positively charged surface groups of adjacent protein molecules are facilitated, which could lead to the formation of crystal contacts.

Fractions containing the PSI-trimer peak from all FPLC runs are collected and concentrated using a 400-mL, stirred, ultrafiltration-unit concentrator (Stirred-cell model 8400, Millipore, Part No. 5124) using a 100-kDa-cuto membrane (Ultrafiltration Membrane Disc Filters, Pall Life Sciences, Part No. OM100076) at 4°C with gentle stirring, using less than 5 psi of pressure[2]. Small amounts of the filtrate are taken and examined visually at different time points for green color. The sample is concentrated and subsequently diluted through the dropwise addition of a buffer containing 0 mM $MgSO_4$ to a $MgSO_4$ concentration between 4

and 12 mM, using a conductivity meter and a standard curve to determine the initial salt concentration.

The concentration of the PSI is determined through a chlorophyll assay, done in triplicate, by adding a small aliquot of the protein solution (0.5–5 μL, depending on the protein concentration) to an 80% (v/v) acetone in water solution. The chlorophyll is extracted into the acetone, and the absorbance of the chlorophyll-containing, acetone solution between 400 and 800 nm is measured. The following formula is used to calculate the chlorophyll concentration:

$$C_{\mathrm{Chl}} = \frac{A_{664} - A_{710}}{76780\,(M^{-1}cm^{-1}) \cdot l} \cdot \frac{1000\,(\mu L)}{V_{\mathrm{sample}}} \quad (2)$$

where C_{Chl} is the chlorophyll concentration of the sample in molar, A_{664} and A_{710} are the absorbance measurements at 664 nm and 710 nm, respectively, l is the path length in centimeters, 76780 $M^{-1}cm^{-1}$ is the molar absorption coe cient at 664 nm wavelength of chlorophyll in acetone [72], and V_{sample} is the volume of the sample, in microliters, used in the chlorophyll assay. The chlorophyll concentration is taken as the mean value of the three measurements produced by Eq. 2.

After diluting the salt, the protein is slowly concentrated to a chlorophyll concentration of 5–12 mM, with gentle stirring at 4°C, inducing crystallization at low ionic-strength conditions using a technique called ultrafiltration crystallization, shown schematically in Fig. 3. The PSI crystals are allowed to grow overnight.

The Photosystem I crystals are harvested from the membrane and 1-mL aliquots of the crystal suspensions are placed into 1.5-mL, reaction vessels forx fractional-sedimentation experiments. The crystals are washed using a bu er containing 3 mM $MgSO_4$, followed by two washes with a buffer containing 0 mM $MgSO_4$.

The PSI-crystal samples are re-suspended, and after 10 minutes the supernatant is carefully removed from the settled pellet and placed into a new 1.5-mL, reaction vessel, with gentle mixing. After 20 minutes, the supernatant is again removed from the settled pellet and placed into a new 1.5-mL, reaction vessel. This procedure is repeated for 30 min, 40 min, and 60 min settling steps. The settled crystals are combined into 1–3, 1.5-mL reaction vessels, and a 500-μL aliquot of a buffer containing 0 mM $MgSO_4$ is placed above each pellet; the pellets are stored at 4°C.

1.5. Results of the PSI purification and crystallization

The ultrafiltration crystallization method was used because it would *intrinsically* produce a large size distribution of PSI crystals. Images of the PSI crystals contained within the pellet of each fractional-sedimentation time point are shown in Fig. 4. The size distribution of the crystals in the pellet remaining after each settling experiment can be estimated. In the 10-min settled crystals of Fig. 4a, the size distribution is measured to be 2–100 μm in size, whereas in the 20-min PSI crystals of Fig. 4b, the size distribution is measured as 2–30 μm in size. The size distribution of the 30-min settled PSI crystals of Fig. 4c is measured to be 2–20 μm in size, and the size distribution of the 40-min settled PSI crystals of Fig. 4d is measured to be ≤ 2–5 μm in size.

[2]Higher pressures in the initial concentration step leads to clogging of the pores and reduced flow. During the crystallization step, amorphous precipitate is formed on the membrane when higher pressure is used, which clogs the membrane pores

Unfortunately, it was not clear what the size distribution, smallest size, or quality of the PSI samples made were, and the only method available at the time to study these characteristics of the samples was X-ray diffraction. However, we were interested in the smallest crystals, and consequently, only the most intense X-ray sources could be used to measure signal to high resolution. The most intense conventional X-ray sources are the third-generation X-ray sources, but novel sample introduction techniques were necessary due to X-ray-induced radiation damage.

2. Serial crystallography and liquid jets

Third-generation X-ray sources, which refer to storage rings that utilize insertion devices such as undulators and wigglers, have revolutionized X-ray experiments in many fields of science. One prominent example of the influence of third-generation X-ray sources is X-ray protein crystallography. X-ray sources are usually described in terms of the brilliance (photons s^{-1} mm^{-2} $mrad^{-2}$ /0.1% bandwidth) of the source, which is a measure that takes into account many properties of the X-ray beam, such as photon flux, beam divergence, and bandwidth, and can succinctly be interpreted as a measure of the quality of the X-ray beam. Third-generation X-ray sources offer much higher brilliance than any previous X-ray source, with the APS, Spring-8, ESRF, and PETRA III being examples of third-generation X-ray sources [73]; the higher brilliance allows the use of smaller protein crystals. However, X-ray-induced radiation damage remains a major obstacle, and so the smallest PSI crystals grown using ultrafiltration crystallization method described in the preceding section could not be used in a traditional X-ray crystallography experiment, due to their small size.

One idea presented to avoid X-ray-induced radiation damage using conventional sources is serial crystallography [74]. Serial crystallography proposed the idea that by constantly replenishing a biological sample—in the case of the original paper, it was single biomacromolecules—the X-ray-induced radiation damage was avoided; the idea was that a stream of molecules could be introduced to the source and could diffract before significant radiation damage could occur. Specifically, a jet would be used to introduce the sample to an electron beam—the idea is easily generalized to an X-ray source—and an external alignment mechanism would be used to orient the molecules preferentially, allowing thousands of patterns of identically oriented molecules to be summed up to produce a diffraction pattern.

Sample introduction was key to the serial crystallography idea. The first method tested for sample introduction utilized a Rayleigh jet, in which a glass nozzle was used to produce the jet. A Rayleigh nozzle produces a liquid jet with a diameter equal to the inner diameter of the nozzle and droplets that are approximately twice the inner diameter [75]. However, as the jet diameter needs to be minimal, the inner diameter of the Rayleigh jet must be reduced, leading to issues of clogging. In order to avoid the clogging problems of the Rayleigh jets, a liquid jet system was developed that used a high-velocity accelerating gas to squeeze a liquid into a jet through an aperture [76]. The nozzle—termed a gas-dynamic virtual nozzle—could produce liquid jets from hundreds to thousands of nanometers in thickness [77]. By using a jet to constantly replenish the sample, signal averaging could be used to increase the signal-to-noise ratio of the data and possibly provide high-resolution data that was not plagued by diffraction from damaged species.

3. Serial powder diffraction of membrane-protein nanocrystals

An important initial experiment of the femtosecond nanocrystallography project was to test the reproducibility and quality of membrane protein microcrystals as well as attempting to gain insight into the frequency of occurrence of the submicron crystals. In order to study these aspects of the PSI samples produced using ultrafiltration crystallization, a gas-dynamic virtual nozzle was used to produce a liquid jet of fully-hydrated PSI-crystal sample that was

introduced to the soft X-rays of beamline 9.0.1 of the Advanced Light Source (ALS) of Lawrence Berkeley National Laboratory (LBNL) utilizing the technique of serial crystallography (but replacing external alignment mechanisms with the crystallographic alignment within the crystalline samples) [78, 79].

A full description of the experimental setup can be found in [79]. Briefly, the serial powder diffraction experiments were done by using a liquid injector [75] to introduce a sample of fully-hydrated PSI crystals—grown in 6–8 mM $MgSO_4$—to the 520- or 1560-eV X-rays of beamline 9.0.1 of the ALS. Diffraction patterns were recorded with both an on-axis and off-axis detector, with maximum resolutions in the corners of the detector for the 1560-eV (520-eV) X-rays being 28 Å (84 Å) and 14 Å (42 Å), respectively. For the 1560 eV experiments with an off-axis detector, as many 30-s exposures as possible were taken and averaged together to produce the powder diffraction pattern.

In order to test whether sub-micron PSI crystals exist, 500-nm and 100-nm inline filters were used to restrict the maximum crystal size allowed to proceed into the interaction region. The PSI-crystal samples passed through the 500-nm filter produced measurable diffraction patterns to 28-Å resolution using 1560-eV X-rays on the off-axis detector, the resolution at which the powder lines of PSI crystals would start to overlap due to the large unit-cell volume of the PSI crystals. A PSI crystal passing through a 500-nm filter could maximally have 6000 unit cells, which is nine orders-of-magnitude smaller than PSI crystals used for conventional crystallography. The PSI-crystal samples passed through the 100-nm filter produced measurable diffraction patterns to 51-Å resolution, using 520-eV X-rays, on the off-axis detector; the resolution in the case of the ≤ 100-nm PSI crystals was limited by the small amount of sample used to collect the data. A PSI crystal passing through a 100-nm filter would have less than 100 unit cells. The diffraction patterns collected with an off-axis detector for the 500-nm and 100-nm PSI crystals, crystals referred to as nanocrystals, are shown in Fig. 5 and 6, respectively.

The diffraction patterns shown in Figs. 5 and 6 were radially averaged and the one-dimensional plot is shown in Fig. 7. A very interesting feature of the overlaid plots is that the full-width at half maximum (FWHM) of the peaks is approximately the same for the 100-nm and 500-nm data sets. A Scherrer analysis done utilizing the FWHM of the peaks can estimate the size of the particle producing the powder diffraction patterns [80] and Scherrer analysis of the plots in Fig. 7 indicate that the crystals are the same size in each data set. Originally, it was argued that the (unexpectedly) large peak width in the 500-nm data could be explained by mosaicity effects [79], however, newer data show that the internal disorder of PSI nanocrystals is very small. Therefore the broad powder diffraction rings in the 500-nm patterns could be better explained by the dominance of 100-nm crystals in the 500-nm-filtered sample. This is an area of great interest moving forward.

The serial powder diffraction experiments revealed the potential of membrane-protein nanocrystals and microcrystals for structure determination experiments, but cannot be used to solve the structures of large proteins directly. Although powder diffraction is a suitable method for structure determination for small molecules, peptides, and proteins, the technique becomes less effective as the unit-cell volume of the crystal of interest increases, due to a loss of information in powder diffraction [81]. This is due, in part, to the reduction of three-dimensional reciprocal space information into one-dimension, *i.e.*, the diffraction data is treated as a function of d alone. In addition to the distinct possibility of peak overlap caused by the large unit-cell volume of typical protein crystals, peak broadening due to instrumental factors, such as beam divergence, energy spread, etc., will add to peak broadening caused by finite size effects and mosaicity effects. As the protein of interest transitions from small, soluble proteins, to membrane proteins and protein complexes, the

overlap of the peaks can become an insurmountable factor in the analysis of powder diffraction data. Furthermore, the small spacing of the powder diffraction lines leads to a "continuum" at higher diffraction angles. Therefore, the serial powder diffraction experiments could not be a stand-alone experiment used as a general method for structure determination of membrane proteins using membrane protein nanocrystals and microcrystals, but it will be an effective way to characterize the nanocrystals and microcrystals of membrane proteins moving forward. A single-crystal method would be needed to allow the use of the PSI microcrystals for high-resolution structure determination, and, fortunately, a new type of X-ray source exhibits properties that may allow single-crystal experiments using protein nanocrystals.

4. The fourth-generation X-ray sources

4.1. VUV and X-ray free electron lasers

Recently, fourth-generation VUV and X-ray sources have been designed and built, with noticeable examples being FLASH at Deutsches Elektronen-Synchrotron (DESY) [82], the LCLS at SLAC National Accelerator Laboratory [83], the SPring-8 Compact SASE Source at SPring-8 [84], and the upcoming European XFEL at DESY [85]. The peak brilliance of these fourth-generation X-ray sources, known as the X-ray free electron lasers (XFELs), is up to 10 orders of magnitude higher than the peak brilliance currently available at the most brilliant third-generation X-ray sources [83]. Due to the incredible peak brilliance offered by the XFELs, many new areas of physics and chemistry could be examined with unprecedented spatial resolution.

XFELs involve relativistic electron bunches from a linear accelerator that are compressed in a bunch compressor and fed into a long undulator. In the undulator, X-rays with extremely high brilliance are created by self-amplified spontanteous emission (SASE). To achieve SASE, the XFEL is tuned such that when an electron bunch traverses one period of the undulator, the radiation emitted by the adjacent upstream electron bunch is in phase with that emitted by the downstream electron bunch. The ponderomotive force accelerates those electrons (in the downstream bunch) that are out of phase with the radiation, while decelerating those that are in phase, creating a fine structure of microbunches within the electron bunch. As a result, the spontaneously emitted radiation from the microbunches further amplifies the coherence of the X-ray wavefield, which in turn forces the microbunches downstream in a more well-defined bunch, thus emitting even more coherently. An exponential increase in the spontaneous emission through this resonant process leads to saturation of the XFEL and to the arrival of X-ray bunches of extremely high brilliance at the interaction region [86].

The peak intensity of an XFEL is much higher than that of any third-generation X-ray source and could be of use for X-ray protein crystallography. However, as discussed in Section 1.2, the increased intensity will not solve the problem of X-ray-induced radiation damage. Fortunately, the XFELs offer one important characteristic that may lend itself to mitigation or removal of X-ray-induced radiation damage altogether, the ability to produce ultra-short X-ray pulse durations.

4.2. Diffract before destroy

The X-ray beam produced at an XFEL arrives in pulses with pulse durations related to the electron bunch length in the undulator [87] between several and several-hundred femtoseconds. The time scale of the pulse duration offered by XFELs is on the same (or shorter) time scale as many of the physical and chemical processes involved in X-ray-induced radiation damage, such as the secondary damage pathways. As the secondary events are the main mechanism for the specific damage in X-ray protein crystallography [88], the

short time scale of the X-ray pulses may allow diffraction data to be collected before significant damage degrades the crystal and the diffraction pattern.

In the year 2000, a paper was published showing simulations of the damage pathway of a lysozyme molecule in the gas phase that interacts with an intense X-ray pulse from an XFEL [89]. Ultimately, the large positive charge harbored by the lysozyme molecule led to a Coulomb explosion, but the simulations showed that the explosion occurred on the time scale of 5–10 femtoseconds. The simulation results can be interpreted such that if an X-ray pulse terminates before the onset of the Coulomb explosion, all of the diffraction data collected will be without the degraded signal from the damaged sample, which could add significant background to the diffraction patterns. The results led the authors to propose the diffract-before-destroy hypothesis, in which X-ray-induced radiation damage is avoided by the termination of the X-rays used before the onset of the secondary damage processes.

The lysozyme simulation results were published before any VUV or X-ray free electron laser became operational, and so it was not known whether the simulations would be commensurate with experimental data obtained using an XFEL. The first opportunity to experimentally test the diffract-before-destroy hypothesis was when the first VUV free-electron laser, the Free electron LASer in Hamburg Germany (FLASH), started operation in 2005. FLASH offered wavelengths between 47 and 470 Å [90] with the short, 10–50 fs, pulses expected from the upcoming XFELs, with high peak intensities of up to 10^{16} W/cm^2.

Using objects etched into silicon nitride windows, experiments were carried out showing that the diffraction patterns collected from FLASH could be used to produce an undamaged image of the object to 32-nm resolution [91]. Other results from FLASH showed that the presence of a layer of solvation around the sample of interest will cause the Coulomb explosion to be retarded, and the retarded explosion provided evidence that the pulse durations of the XFELs could be longer than the simulation of the gas-phase biomolecule, due to a sacrificial tampering by the solvent layer [92].

Although FLASH allowed for some interesting experiments, the wavelengths offered were too long to be of direct interest to the crystallography community. However, higher-energy free electron lasers, such as the Linac Coherent Light Source (LCLS) at SLAC National Accelerator Laboratory, offered the potential to operate at crystallography-relevant wavelengths. The first lasing of the LCLS occurred in April 2009, and the Atomic and Molecular Optics (AMO) beamline became available to users at the end of 2009 [93]. The initial operation of the LCLS generated X-rays with up to 1800 eV energy (6.9-Å wavelength) with a repetition rate of 30 Hz and a flux of 10^{12}–10^{13} photons/pulse.

5. Initial femtosecond nanocrystallography experiments

At this point, the serial powder diffraction experiments had shown that nanocrystals and microcrystals of PSI, as small as a few hundred total unit cells, exist and exhibit enough order to diffract X-rays with measurable Bragg peaks. The preparation of PSI nanocrystals and microcrystals was shown to be reproducible and predictable. The growth of PSI nanocrystals likely indicates that nanocrystals of other membrane proteins can be grown and may even be produced more commonly than microcrystals of membrane proteins.

The motivation behind the initial femtosecond X-ray protein nanocrystallography experiments was to provide a proof of principle of the diffract-before-destroy concept for protein crystallography using membrane protein crystals. The experiments were set up so that one diffraction pattern would be collected from one nanocrystal, as described below. Success with the technique of femtosecond X-ray protein nanocrystallography would provide a new avenue for the structure determination of membrane proteins, offering the

ability to use membrane protein crystals of any size for structure determination experiments. Crystals of Photosystem I were chosen as a model system because with a large hexagonal unit cell (a = b = 281 Å, c = 165 Å) and high solvent content (78%), it presents a most stringent test for the analysis of nanocrystals by femtosecond X-ray crystallography.

Three major experiments were done during these femtosecond nanocrystallography studies. The first experiment was to record diffraction data to the highest allowable resolution, of 8.5 Å, and use the data to produce merged intensities, calculate structure factors, and produce an electron-density map of PSI. Complementary experiments were conducted to assess the internal consistency of the LCLS data as well as to compare the LCLS data to conventional X-ray crystallography data for PSI. The second major experiment done was to compare the diffracted intensity for PSI nanocrystals at several different X-ray pulse durations, in order to gain insight into the X-ray-induced radiation damage to the PSI crystallites. The third major experiment was to record the interference patterns around the low-resolution Bragg reflections that are caused by the finite crystal size.

5.1. Preparation of the sample for femtosecond nanocrystallography

The PSI nanocrystal samples used for the first femtosecond X-ray protein nanocrystallography experiments were obtained from two different large-scale PSI preparations, which were completed directly before the LCLS experiments. The PSI nanocrystals from the 40-min settling steps that were used in the experiments were restricted in size by using a 2-μm inline filter; therefore, the PSI nanocrystals contained between 100 and (maximally) 500,000 unit cells for the 2-μm microcrystals. However, there were no visible crystals in the re-suspended sample, and most of the crystals were ≤ 1000 nm in size. It should be noted that even the 2-μm PSI crystallites were between 10^7 and 10^8 smaller than the PSI crystals used in conventional crystallography.

Based on crystal-density calculations, the PSI sample was diluted to 1 mg/mL PSI and loaded into the sample line of a dual-line system. The liquid injector was run with a flow rate between 10 and 12 μL/min, and the jet was centered in the X-ray beam by observing a streak on the back detector, due to scattering of the X-rays by the continuous column of the jet. Data for the PSI samples were collected over the course of 48 continuous hours of experiments. Millions of diffraction patterns were recorded, corresponding to 25 TB of data.

5.2. Diffraction experiments

Details of the first femtosecond X-ray protein nanocrystallography experiment, in addition to a full list of acknowledgements, can be found in [94]. Details of the data analysis can be found in [95, 96].

Concisely, the soft X-ray femtosecond nanocrystallography experiment was done by using a liquid jet to introduce a stream of fully-hydrated PSI crystallites to the LCLS—with X-ray pulse durations between 10 and 250 fs and a repetition rate of 30 Hz—at the AMO beamine [97] of the LCLS utilizing the CAMP chamber [98] as shown in Fig. 8. The X-ray energy used for the initial experiments was 1.8 keV (λ = 6.9 Å), and diffraction was recorded on two sets of detector modules: a “front” detector, able to record reflections to maximum resolution of 8.5 Å at the corner, and a “back” detector, able to record reflections between 4000 and 100 Å[94]. During the femtosecond nanocrystallography experiment, one pulse of the LCLS would ideally inundate one PSI crystallite, producing a diffraction pattern recorded on both the front and back detectors that is read out and digitized before the subsequent LCLS pulse arrives.

5.3. Data processing and analysis of snapshot diffraction patterns

A detailed description of the data analysis of the PSI data from the LCLS is provided in [96], but a brief description is given here. The diffraction patterns were processed by removal of any known bad detector pixels, followed by application of a flat-field correction, normalization of the pixel values with respect to the dark current and gain of the panel, and subtraction of the background. The background was defined as a moving-window selection of 50 frames, with the background for a given pixel defined as the median value over the range [96].

Once the diffraction data were fully processed, indexing and merging of the data were performed. However, for each X-ray pulse, the Ewald sphere would slice through the reciprocal lattice of the PSI crystal in a random position. Due to the sparse population of reciprocal points in the reciprocal space of the crystal system, most scattering would not occur due to the intersection of the Ewald sphere with a reciprocal lattice point. Instead, much of the scattering is due to intersection of the Ewald sphere with the shape transform laid down upon the reciprocal-lattice point. A peak finder was used to establish peaks, and the peak locations were given to MOSFLM [99], using the DPS algorithm [100], or DirAx [101] for indexing. Indexing the patterns would give the orientation of the crystal relative to the lab frame, and could be used for further processing. Each orientation matrix was further refined by a global optimization of the unit-cell constants and three Euler angles [96].

The structure factors extracted from the LCLS data were compared to structure factors extracted from a conventional crystallographic data set of PSI collected at beamline 8.2.2. of the ALS. The conventional data and LCLS data were collected on samples prepared quite differently, as the conventional data originated from a large single crystal of PSI that had been incubated in 2 M sucrose and cryogenically frozen using liquid propane. The conventional data were collected specifically with large working distance and a 100× attenuated beam, so that the low-resolution data were not from saturated detector pixels.

5.4. Femtosecond nanocrystallography results for PSI

A total of 1.85 million “snap-shot” diffraction patterns from PSI were collected at the 70-fs pulse duration. Bragg peaks from the PSI crystallites were recorded to resolution of 8.5 Å on the front detector, as shown in Fig. 9. The diffraction of the nanocrystals was so strong that even saturated peaks appeared occasionally in the diffraction patterns. Unlike in cryo-EM or traditional crystallography, the X-ray diffraction patterns were collected at room temperature, without any crystal mounting, freezing, or cryo-protectants. The diffraction patterns showed well-resolved peaks, and some orientations provided many Bragg peaks per pattern.

A 20% hit rate was achieved by crystal-density calculations. Most of the recorded crystallite diffraction patterns were from one X-ray pulse hitting one nanocrystal, with the patterns on the front and back detectors reading out before the arrival of the next pulse. Fine adjustment and optimization of the crystal density was done based upon the perceived hit rate as determined from the data stream at the LCLS control room.

Of the 1.85 million diffraction patterns of PSI crystallites that were collected at the 70-fs pulse duration, over 112,000 contained ten or more identifiable peaks, of which a total of 28,192 were indexed with unit cell parameters within 5% of the mean values, with a representative image shown in Fig. 10a. Indexing success was strongly correlated to the number of peaks contained within the pattern, as 60% of the patterns containing 75 or more peaks were indexed. However, approximately 42% of indexed patterns were rejected during the intensity-merging steps, due to disagreement between the predicted and observed peak positions. The reason for the disagreement is still unknown, but deviations between the

predicted and observed peak positions likely resulted from slight indexing errors that produced large effects at the higher-resolution reflections. In total, 16,374 patterns were integrated and contributed to the structure-factor amplitudes used to produce the PSI structure shown in Fig. 10c [94].

In order to further evaluate the nanocrystallography data to 8.5 Å, rigid body refinement of the published PSI structure (PDB code 1JB0) against the 70-fs pulse-duration nanocrystal structure factors was performed using the program *REFMAC* [102], yielding R/R_{free} = 0.284/0.327. In addition, rigid body refinement of the published PSI structure was performed against the structure factors extracted from a higher-resolution single-crystal PSI data set, in which good low-resolution data were measured and collected at beamline 8.2.2 of the ALS. The ALS data set was truncated to 8.5-Å resolution, yielding R/R_{free} of 0.285/0.298. Figure 10c shows a representative regions of the $2mF_0$-$2DF_c$ averaged-kick-map electron density at 8.5 Å from the LCLS data sets. The electron density map clearly show the transmembrane helices, as well as electron density from the membrane extrinsic subunits, and the loop regions are also visible at 8.5-Å resolution. The reference electron-density map shown in Fig 10d was generated using data from a large, single crystal of PSI using 12.4-keV X-rays, with a single crystal of Photosystem I that had been cryogenically cooled to 100 K and had 2.00 M sucrose as a cryo-protectant [94].

The refinement statistics for the PSI electron-density map calculated from the LCLS data are shown in Table 1. The data show good statistics for all categories until the highest resolution shell, in which case the data were measured in the corner of the detectors, causing a reduced number of total measurements, likely contributing to the poorer statistics.

5.5. Comparison of the LCLS data to conventional data

One important comparison required for the new PSI femtosecond X-ray protein crystallography data is with conventional X-ray protein crystallography data of PSI collected at a conventional beamline. The comparison will allow the LCLS data to be compared against the standard for X-ray crystallography.

The LCLS structure factors were compared to structure factors determined from PSI data collected conventionally at beamline 8.2.2 of the ALS, on large crystals at cryogenic temperatures, as a measure of the relative accuracy of the LCLS structure factors. The ALS data were collected using a large, single crystal of PSI that was cryogenically frozen in liquid propane, in which the data were collected to a resolution of 2.3 Å, but the data were truncated at 8.5 Å Å. The overall R_{iso} between the LCLS and ALS structure factors was 23.5%, indicating the structure factors from the LCLS and ALS data are comparable. The linear correlation coeffcient, C_{iso}, comparing the LCLS and ALS structure factors for PSI, approached 0.9 around a resolution of 15 Å, and quickly fell o at higher resolutions, again indicating the similarity in the data sets for the resolutions in which a large number of reflections were recorded and processed (for the LCLS data) [96].

Deviations between the LCLS and ALS data sets for PSI were expected. The ALS data were collected on a crystal that was cryogenically frozen and contained 2.00 M sucrose within the unit cell, whereas the LCLS data were collected at room temperature and had no cryo-protectants. The unit cells had slightly different until cell parameters, and in addition, the different wavelengths would produce different scattering and absorption cross-sections. The LCLS data were not processed to take any absorption into account, although at 1.8 keV, the small absorbance correction for the protein crystal and liquid stream may allow for a better comparison between the data sets.

After showing the similarity of the LCLS data and reconstructed PSI electron-density map to the conventional data and resultant electron-density map, it is clear that the LCLS data, to 8.5-Å resolution, is the same—if not better than—the conventional case. That the LCLS data were collected using samples in which the majority of the crystallites could not be seen in a visible microscope is remarkable. This is in stark contrast to the large PSI crystal at a conventional beamline, where crystals larger than 1 mm are used. Yet, the LCLS data is similar to the conventional data to this resolution, indicating that the femtosecond X-ray protein crystallography could help to solve the structures of difficult-to-crystallize membrane proteins.

5.6. The diffract-before-destroy principle and radiation damage

The major reason that small crystals of proteins cannot be easily used in conventional X-ray crystallography experiments is that X-ray-induced radiation damage will ultimately limit the exposure time, thereby limiting the true resolution of the structure. A major premise of the femtosecond nanocrystallography project is that the femtosecond X-ray pulses are as fast as, or faster, than the time scale of the conventional X-ray damage processes, such as disulfide bond reduction, decarboxylation of amino acids, etc. Diffraction patterns were collected with X-ray pulse durations of 10 fs, 70 fs, and 250 fs, in order to compare the effects of the different pulse durations on the quality of the diffraction data.

Examining the electron density maps produced using the structure factors from the LCLS data sets, no specific damage was evident when using the 70 fs pulse durations. However, the onset of damage, when the secondary electrons were spawned, occurs on the 10–100-fs timescale [88]. To assess the damage induced by the XFEL beam, data from 10-fs and 250-fs-duration pulses were collected to directly compare the decay in scattering as a function of time. The 10-fs pulses were operated at a fluence about 10% of the fluence for the 70-fs pulses, corresponding to an absorbed dose of approximately 70 MGy. Fig. 12 shows the integrated structure factors obtained from the three different pulse durations, 10 fs, 70 fs, and 250 fs, using 97,883, 805,311, and 66,063 patterns respectively, normalized with respect to fluence. The plots of the scattering strength of the crystallites versus resolution were generated by selecting and summing Bragg spots from the patterns.

The thickness of the lines in Fig. 12 indicates the uncertainty of each plot. The 10- and 70-fs data sets are indistinguishable, indicating that the 70-fs pulse durations were short enough to avoid global damage to 8.5-Å resolution, relative to the 10-fs pulses. The decrease in integrated scattering intensity for the 250-fs pulses, beyond 25-Å resolution, indicates that the 250-fs pulse duration caused significant radiation damage. The data clearly show the onset of global radiation damage below 8.5-Å resolution occurs at time scales greater than 70 fs.

The dose received by the PSI crystallites was calculated to be as large as 700 MGy using RADDOSE [63] for a fluence of 900 J/cm^2. The dose of 700 MGy is more than a twenty-fold increase over the damage threshold of 30 MGy from conventional X-ray crystallography [68]. Owen *et al.* determined the $D_{1/2}$ limit of 43 MGy based upon empirical evidence of the totaled scaled-scattered-intensity ΣI_j over all reflections j, which is a measure of global damage. Although 43 MGy corresponded to $D_{1/2}$, the authors still concluded that the 30 MGy limit was the safer choice to ensure quality data. 30 MGy corresponded to a decrease in total scaled-scattered-intensity by a factor of $\ln(2) \approx 69.3\%$ of the initial total scaled-scattered-intensity in the study [68]. It should be noted that these damage numbers are for cryogenically-cooled protein crystals, and that cryogenic cooling increases the allowed dose by many orders of magnitude [65]. The LCLS experiments were carried out at room temperature; therefore, the stability of the crystals at room temperature,

despite the increase in the absorbed dose by a factor of 23 when compared to conventional data collection, is even more remarkable.

The dose of 700 MGy in the LCLS experiment corresponds to K-shell ionization in 3% of all carbon atoms present in the protein molecules. The photoionization process can be considered an instantaneous process and can only be outrun with the shortest pulse durations. The loss of the *K*-shell electrons would requisitely cause a decrease in scattered intensity, as the X-rays are scattered from electrons in the sample. Consequently, the lack of radiation damage for the 70-fs-duration pulses is only relative to the shorter pulses, because a decrease in scattered intensity must occur due to the initial photoionization events. However, photoionization of *K*-shell electrons is also a problem in conventional crystallography, but relaxation of the system after the initial excitation by the electrons contributes a much larger effect to the damage than the initial photoionization by itself.

The energy absorbed during the X-ray exposure is released through photoionization and Auger decay, followed by a cascade of lower-energy electrons caused by the secondary electrons on the 10–100-fs time scale [88]. Additionally, simulations using plasma dynamics [103] indicate that each atom of the PSI crystallite was ionized once during a 100-fs interval after initial exposure and that nuclear motion had begun on the 100-fs timescale. A photoelectron produced in these experiments from the *K*-shell ionization of carbon would have a kinetic energy of approximately (1800 eV – 285 eV) 1500 eV, which could produce hundreds to thousands of low-energy electrons. The secondary electrons and low-energy electrons cause the radiation chemistry to initiate, which leads to specific damage at sensitive amino acid residues/sites. However, redox reactions occur in the ≤ 100-fs timescale and the loss of specific groups from amino acid residues occurs once nuclear motion initiates, which is greater than the 100-fs timescale. It is clearly noted that in conventional crystallography, the specific sample needs to be considered when determining the allowed dose, as the decrease in spot intensity does not give insight into the radiation chemistry occurring. Specific amino acids, such as cysteine, glutamate, and aspartate, are more susceptible to radiation damage than others, and if these radiation-sensitive residues appear in the active site of an enzyme or at the crystal contacts, this could lead to an inability to interpret the electron-density map [63, 104]. Additionally, metal-containing cofactors are acutely prone to local X-ray-induced radiation damage. Using the femtosecond nanocrystallography experiment, consideration of the specific biochemical makeup of the protein may not be necessary, as the experiments are done on a timescale at or faster than the onset of nuclear motion.

The results of the simulations presented in [89] created an expectation that the 70-fs pulse duration data would exhibit some damage effects relative to the 10-fs pulse duration data. The increased time could allow for more Auger relaxation, and secondary ionization events, as well as the Coulomb explosion, which is estimated to begin after just a few tens of femtoseconds [88]. There are multiple possible reasons for the lack of visible radiation damage in Fig. 12 for the 10-fs and 70-fs data. The most immediate reason is that the displacement of atoms in 70 fs is in the sub-Å range and is not visible at 8.5-Å resolution, and therefore the data do not attain a high enough resolution for the increased damage to manifest in the intensity plots of the 70-fs pulses. Another explanation is that the liquid jet surrounding the PSI crystallites acts as a sacrificial tamper for the nanocrystals, in which photoelectrons produced in the stream can neutralize some of the positive charge building up in the sample [92]. One additional explanation is that the nominal pulse durations given are from the length of the electron bunch and that the X-ray pulses are shorter in duration than the electron bunches; simulations of transparency effects in neon using data from the LCLS indicate that nominal 80-fs pulse duration data were more consistent with pulse durations between 20 and 40 fs [87]. Thereby, the lack of damage in the 70-fs data set relative to the

10-fs data set could be attributed to the resolution limit of the data, sacrificial tamper effect of the solvent, or shorter pulses than the nominally reported pulses.

The LCLS experiments collected crystallographic data on PSI nanocrystals at room temperature. The PSI nanocrystals prepared for the LCLS experiments did not require any freezing of crystals, making sample handling much easier. However, cryogenic cooling decreases the temperature factor of the molecules while also (potentially) increasing the internal disorder of the crystal, making the benefit of room temperature diffraction diffcult to quantify until higher-resolution data become available.

5.7. Shape transform results and discussion

Although the diffraction patterns collected on the front detector were used for the integration of the majority of the data used to reconstruct the PSI electron-density at 8.5-Å resolution, the back detector contained a wealth of information as well. In at least one regard, the information collected with the back detector was as equally exciting as the front-detector data.

The back detector measured high angular-resolution data and could be used to resolve interference fringes around Bragg reflections that are caused by the small number of unit cells of the PSI crystallites used in the experiment being irradiated by a transversely coherent X-ray beam. The result of the interference effects around the Bragg peaks would be to observe measurable intensities away from the Bragg peaks, which cannot be measured in conventional X-ray protein crystallography experiments using a protein macrocrystal. The “added” intensity contains a large amount of information that is normally lost in conventional crystallography, so the back detector results were greatly anticipated.

The large working distance of the back detector provided the necessary angular resolution to well-resolve subsidiary maxima, as shown in Fig. 11. Thereby, the crystal size could in principle be determined by counting the number of fringes between Bragg peaks (for instance the number of Bragg peaks between the (100) and (200) reflections would give the number of unit cells along **a***). Patterns from crystals containing less than 10 unit cells per edge were recorded, which shows the PSI crystallites were in the size range of 200–1000 nm, consistent with the serial powder diffraction results. However, the smallest crystals may not have produced diffraction patterns that were identified by the hit-finding algorithm, as they may produce weaker Bragg reflections. Thereby, patterns from smaller crystals may also be present in the data sets. Re-examining the data to search for smaller crystals is still a possibility.

The interference fringes surrounding the Bragg peaks allow the use of image-reconstruction techniques from the coherent-diffractive-imaging field. Using phase retrieval methods [105, 106], a cross section of the crystal size and morphology can be determined from the interference pattern surrounding a Bragg peak, as shown in Figure 11a–d [94]. The information contained within the shape transform may facilitate a direct solution to the phase problem, according to Shannon's theorem [107], because the intensity between Bragg peaks allows for oversampling of the diffraction pattern [108].

Although the maximum size of the crystals was restricted by inline filtering, the shape transforms were observed because the minimum size could not be strictly chosen, so the minimum size was the soluble PSI_T. The size distribution of the PSI sample was not known in advance; however, the size of the crystal was observed to affect the produced diffraction pattern substantially, as shown in Figure 11a–d, with the observation of the shape transforms.

A diffraction grating with N slits will give rise to diffraction features that are finer than the Bragg peaks by a factor of $1/N$. Therefore, a crystal with N unit cells in a given direction will produce diffraction patterns with $N - 2$ subsidiary maxima between the Bragg peaks, providing a method to determine the size of the crystal projection. The central maxima will have a width of approximately $2/N$, whereas the width of all subsidiary maxima will be $1/N$ for a general parallelepiped crystal. An identical shape transform will be laid down upon each reciprocal lattice site, and the diffracted intensity will have the following form for a general parallelepiped [16]:

$$I \propto \frac{\sin^2(\pi N_a \mathbf{a}\cdot\mathbf{s})\sin^2(\pi N_b \mathbf{b}\cdot\mathbf{s})\sin^2(\pi N_c \mathbf{c}\cdot\mathbf{s})}{\sin^2(\pi \mathbf{a}\cdot\mathbf{s})\quad\sin^2(\pi \mathbf{b}\cdot\mathbf{s})\quad\sin^2(\mathbf{c}\cdot\mathbf{s})} \tag{3}$$

where **a**, **b**, and **c** are the unit cell vectors, *i.e.* $\mathbf{a}\cdot\mathbf{s} = h\cdot x_j$, N_a represents the number unit cells in the **a** direction, etc. The larger the number of unit cells, the smaller the shape transforms until the limit of the conventional crystallography case is reached, in which the inter-Bragg peak intensities are not observable. In the case of a crystal with 10 unit cells in the **a** direction, there will be 8 subsidiary maxima in that direction about the Bragg reflection, and the peak height will be proportional to 10^2, as calculated by Eq. 3.

5.7.1. Hit rates and sample settling—When examining the LCLS data for PSI, the low hit rate for the PSI data became a concern. The major issue for the low hit rate of the LCLS experiments is shown in Fig. 13 and is that the nanocrystals and microcrystals of PSI settled out of solution with time;. The plot shows the decrease in sample chlorophyll concentration as a function of time, as measured by absorbance at 280 nm (for peptide bonds) and 680 nm (for chlorophyll *a*). The plot was made by using a HPLC to run a suspension of PSI crystallites through a 15-m long sample line. Settling of the larger crystallites in the sample reservoir lines was likely the major cause of the decreased protein concentration shown in Fig. 13.

A few methods could be available to prevent the larger crystallites settling from solution at a faster rate than the smaller crystallites. One method would be to keep the crystallites in a neutral-buoyancy solution, which prevents any settling of the crystallites. In the case of PSI, a neutral buoyancy solution can be prepared using 1.4 M sucrose in G_0 buffer. A diffculty that arises with the sucrose-containing solution is that the viscosity increases with the amount of sucrose, and consequently, the liquid pressure in the jet needs to be increased. This can lead to an unstable jet or even a jet that will not run with a given inner-diameter fiber optic. As a result, the inner diameter of the fiber optic would need to be increased, making for a thicker, higher flow-rate jet. The thickness of the jet is a problem when higher-energy X-rays are used, as a strong solvent background will be recorded, which may reduce the dynamic range, or worse, fully saturate the detector pixels. One possibility to overcome the diffculties listed above is to use lower sucrose concentrations. The lower sucrose concentrations will allow for a decrease in viscosity while still increasing the settling time of the larger crystallites.

5.7.2. PSI sample consumption during femtosecond nanocrystallography—An issue that arose during the first femtosecond nanocrystallography experiments related to the low hit rate was that the PSI data were collected for 24 hours at a flow rate of 10 μL/min. Approximately 14 mg of PSI was used for the first femtosecond X-ray protein nanocrystallography experiments in Dec 2009. Operating at 30 Hz, the X-rays only interacted with 0.004% of the sample, hitting one out of every 25,000 nanocrystals at a concentration of one crystallite per four interaction volumes. The effciency increased as the LCLS moved to 120-Hz repetition rate, but this is only an improvement by a factor of four

in the effciency of sample use. Once the European XFEL becomes operational, with tens of kilohertz repetition rate, the effciency will be increased by orders of magnitude, so that a larger percentage of nanocrystals within a sample interact with the XFEL beam [73]. Using 12 kHz with the same setup, one out of every 250 crystallites would be utilized.

In order to show the relative effciencies of protein use between the conventional experiments and the LCLS experiments, the amount of protein in each crystal, and the number of crystals used to collect a data set in each case, must be compared. Although may of the PSI crystallites used in the experiment were smaller than 2 μm in size, the assumption that all the PSI crystallites were 2 μm in size will allow for the most conservative comparison between the conventional and LCLS experiments.

There are approximately 10^7 to 10^8 crystallites with 2-μm edges within a 0.5 mm x 0.5 mm x 2 mm PSI crystal. Using the current hit rate from the LCLS experiments of one in 25,000, $5x10^7$ 2-μm PSI crystallites would produce 2000 diffraction patterns, whereas approximately 10,000 patterns were needed for the completeness of the LCLS data to approach 100%, without any type of post refinement—post-refinement may significantly decrease the number of patterns necessary for completion—and pure Monte Carlo integration of the intensities [96]. Now that the LCLS operates at 120 Hz, $5x10^7$ 2-μm PSI crystallites will produce 8000 diffraction patterns, generating the necessary amount of diffraction patterns for a full Monte-Carlo integration of the intensities.

When conventional X-ray crystallography experiments are done using large PSI crystals, up to 100 large crystals—each containing 10^7–10^8 of the 2-μm PSI microcrystals—are grown and shipped to the experiments. Therefore, the total amount of PSI used for the preliminary LCLS experiments, once scaled for the full repetition rate of the LCLS, is less than the amount of PSI used for data collection on single, large crystals, and the effciency of the femtosecond X-ray protein nanocrystallography technique will be increased by at least two orders of magnitude when the European XFEL becomes operational. Additionally, the conventional crystallization experiments will produce PSI crystals that are too small or too internally disordered (mosaicity ≥ 1.0°) to attain high-resolution, high-quality diffraction patterns. As a consequence, only 1–2% of all protein in a conventional PSI-crystallization experiment produces crystals allow high-quality data sets to be recorded. In fact, the femtosecond X-ray protein nanocrystallography experiments used less protein than *one* conventional X-ray crystallography experiment using PSI, especially since high protein concentrations of 80 mg/mL, in combination with micro and macroseeding experiments, are necessary to achieve the growth of large, single crystals of PSI with mosaicity below 1°.

Even though less PSI sample was used, relative to the conventional crystallographic case, than initially thought, decreasing sample consumption is still a major optimization factor for the femtosecond nanocrystallography project. One promising method to decrease the sample usage would be to decrease the flow rate of the jet. As the jet currently runs at approximately 10 μL/min, slowing the flow rate to a few hundreds of nanoliters per minute would offer a substantial improvement. However, the minimal flow rate is limited by the repetition rate of the X-rays. The linear velocity of the jet, which is related to the flow rate and jet diameter, must be large enough to deliver fresh sample to the X-ray interaction region that could not have received any ill effects of the previous pulse before the subsequent X-ray pulse. In the case of 120 Hz repetition rate, the lower limit of the flow rate will be ~10 nL/min, but the exact number will depend on the jet thickness. Jets running high-viscosity liquids could provide the necessary methods of reducing sample consumption while inhibiting crystallite sedimentation. However, much work remains to be done.

5.8. Resolution limits of the first femtosecond nanocrystallography data

The initial femtosecond nanocrystallography experiments were conducted at 1.8 keV X-rays (λ = 6.9 Å) as determined by the initial operating conditions of the LCLS. The resolution limit of approximately 8.5 Å in the PSI femtosecond nanocrystallography data set was due to the long wavelength used in the experiment, as well as the instrumental geometry.

The 8.5-Å resolution limit of the data also contributed to the small number of peaks in the patterns. The indexing success of MOSFLM or DirAx was correlated to the number of peaks contained within the diffraction pattern, which supports that moving to higher resolution will increase indexing success because the total number of measurable reflections in a given resolution sphere is calculated as follows:

$$N_{\text{reflections}} = V_{\text{unit cell}} \cdot \frac{4}{3}\pi\left(\frac{1}{d_{\min}}\right)^3 \quad (4)$$

Using equation 4, the total number of reflections to 8.5 Å for the PSI structure is calculated to be approximately 82000 non-unique reflections corresponding to 6800 unique reflections. However, the number of reflections in the 3.0-Å-resolution sphere is calculated to be 1.9 million non-unique reflections, corresponding to 158,000 unique reflections. Once the snapshot patterns contain a larger number of peaks, the indexing programs may be able to identify the unit cells with less ambiguity. Additionally, the size of the PSI crystallites may need to be increased to compensate for the reduced scattering cross-section at higher X-ray energies, which will decrease the measured intensities of the shape transforms, providing another benefit to auto-indexing.

6. The CXI instrument

The beamline that is of interest for X-ray protein crystallographers is the Coherent X-ray Imaging (CXI) beamline, which became available for users in February 2011 [109]. The CXI instrument will provide X-ray energies between 2 and 30 keV, with repetition rates up to 120 Hz and as much as 10^{12} photons/pulse. The CXI instrument provides X-ray wavelengths as short as 1.3 Å using the first harmonic, which should allow atomic-resolution diffraction patterns to be recorded.

7. OUTLOOK AND CONCLUSIONS

Membrane proteins are involved in many vital cellular functions and pathways, such as photosynthesis, cellular respiration, and signal transduction, amongst many others. However, the number of solved (unique) membrane protein structures is less than 300. Although X-ray protein crystallography is the workhorse of structural biology, there are many diffculties associated with the technique. The two main di culties of conventional X-ray protein crystallography that the femtosecond nanocrystallography technique attempted to address were the necessity for large, well-ordered protein crystals—which could take years to grow and optimize, but the conditions may never be found in the first place—and the X-ray-induced radiation damage deteriorating the quality of the data and reconstructed electron-density map.

The femtosecond X-ray protein nanocrystallography project was initiated in order to determine whether membrane protein crystallites could be utilized to obtain high-resolution X-ray diffraction patterns, with the ultimate goal of producing molecular- and atomic-resolution electron-density maps to solve the structures of these complex proteins. Many experiments were necessary during the coarse of the project. Experiments were related to sample production, such as the development of methods to produce and characterize

membrane protein nanocrystals, testing of the reproducibly of membrane protein crystallite production. Further, some experiments were related to sample quality, such as determining the diffraction potential and minimum size of membrane protein crystallites. Other experiments and theoretical developments were necessary for the development of sample handling and delivery systems, as well as the development of novel data acquisition and analysis schemes. Of course, none of the data to date could have been collected without the theoretical and experimental development of the LCLS itself, or the CAMP chamber and pnCCD detectors.

Although the first results of femtosecond X-ray protein nanocrystallography have been published in 2011 [94], the technique is rapidly improving and maturing. The excitement is rapidly building in the structural biology community to harness the potential of the femtosecond nanocrystallography technique to revolutionize the capabilities of the structural biology community.

Femtosecond X-ray protein nanocrystallography will offer a new path to the structure of diffcult-to-crystallize proteins, one in which a crystal of any size may be suffcient for structure determination. The most exciting aspect of the technique is that femtosecond X-ray protein nanocrystallography will allow the structures of membrane proteins to be determined without the ill-effects of X-ray-induced radiation damage on the sample.

Ultimately, the true power of femtosecond nanocrystallography will be in its ability to harness the full spatio-temporal resolution offered by X-ray free electron lasers. Time-resolved crystallography using Laue diffraction methods at third-generation synchrotrons have shown the abities of a technique that combines ultrafast spectroscopy with the spatial resolution of crystal-lography, and current time-resolved crystallography can track changes in the electron density on the time scales of hundreds of picoseconds to seconds [110, 111, 112, 113, 114, 115]. The short pulses of the XFELs may allow the time-resolved crystallography to be extended to even shorter time scales, such as hundreds of femtoseconds. Combined with the femtosecond nanocrystallography technique, time-resolved crystallography at XFELs may be more amenable to systems that produce non-cyclical or irreversible, light-induced changes. The immediate implications are in studying the photocatalytic cycles of the photosynthetic proteins, as well as watching light-induced undocking of the proteins involved in the electron transfer of photosynthesis. This could lead to a very fundamental understanding of the basic processes of life and could have implications in artificial protein design and engineering.

The potential applications for femtosecond nanocrystallography are far reaching and could have a major impact on many areas of science and engineering. Results obtained from the method could be used in every area from basic science to the rational design of drugs. However, as is true with most new technologies, some major uses of femtosecond nanocrystallograhy will not be realized or appreciated for many years or decades to come. Although much is left to be done in regard to bringing the femtosecond nanocrystallography method to its full potential, the impact of the technique may be felt in many disciplines. The future of structural biology will be bright, and femtosecond X-ray protein nanocrystallography may present the technique necessary for a paradigm shift from the (mostly) static X-ray protein crystal-lography of today to the (mostly) time-resolved X-ray protein crystallography of the future.

Acknowledgments

Figures 5, 6, and 7 were reprinted from Biophysical Journal, 100, M. S. Hunter, D. P. DePonte, D. A. Shapiro, R. A. Kirian, X. Wang, D. Starodub, S. Marchesini, U. Weierstall, R. B. Doak, J. C. H. Spence, P. Fromme,, X-ray diffraction from membrane protein nanocrystals, 198–206, Copyright (2011), with permission from Elsevier.

A full listing of the people involved in the experiments described here, as well as their contributions, is given in [94]. We would like to thank David Shapiro for his work on the powder diffraction experiments using Photosystem I.

Experiments were carried out at the Linac Coherent Light Source national user facilities operated by Stanford University, as well as the Advanced Light Source user facilities operated by the University of California Berkeley, on behalf of the U.S. Department of Energy (DOE), Office of Basic Energy Sciences.

This work was supported by NSF award IDBR 0555845, the Center for Biophotonics Science and Technology (University of California at Davis), the Lawrence Berkeley National Laboratory Seaborg Fellowship award, by the U.S. Department of Energy through the PULSE Institute at the SLAC National Accelerator Laboratory, and by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344, and the Joachim Herz Stiftung.

I acknowledge support from the Helmholtz Association; the Max Planck Society, for funding the development and operation of the CAMP instrument within the ASG at CFEL; DOE, through the PULSE Institute at the SLAC National Accelerator Laboratory, and by the Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344; the US National Science Foundation (awards 0417142 and MCB-1021557); the US National Institutes of Health (awards 1R01GM095583-01 (ROADMAP) and 1U54GM094625-01 (PSI:Biology)); the Joachim Herz Stiftung, the Swedish Research Council; the Swedish Foundation for International Cooperation in Research and Higher Education, Stiftelsen Olle Engkvist Byggmastare.

The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

[1]. Mavromoustakos T, Durdagi S, Koukoulitsa C, Simcic M, Papadopoulos MG, Hodoscek M, Grdadolnik SG. Strategies in the Rational Drug Design. Current Medicinal Chemistry. 2011; 18:2517–2530. [PubMed: 21568895]

[2]. Dubiella-Jackowska A, Wasik A, Przyjazny A, Namiesnik J. Preparation of soil and sediment samples for determination of organometallic compounds. Polish Journal of Environmental Studies. 2007; 16:159–176.

[3]. Drenth, J. Principles of Protein X-ray Crystallography. 3rd edn.. Springer Science+Business Media, LLC; New York: 2007.

[4]. Rupp, B. Biomolecular Crystallography. Garland Science; New York: 2010.

[5]. Bartesaghi A, Subramaniam S. Membrane protein structure determination using cryo-electron tomography and 3D image averaging. Current Opinion in Structural Biology. 2009; 19:402–407. [PubMed: 19646859]

[6]. Ubarretxena-Belandia, I.; Stokes, DL. Advances in Protein Chemistry and Structural Biology. Vol. vol. 81. Elsevier Academic Press Inc.; San Diego: 2010. Present and future of membrane protein structure determination by electron crystallography.

[7]. Markwick PRL, Malliavin T, Nilges M. Structural Biology by NMR: Structure, Dynamics, and Interactions. Plos Computational Biology. 4

[8]. Yee A, Gutmanas A, Arrowsmith CH. Solution NMR in structural genomics. Current Opinion in Structural Biology. 2006; 16:611–617. [PubMed: 16942869]

[9]. Berman H, Bhat T, Bourne P, Feng Z, Gilliland G, Weissig H, Westbrook J. The Protein Data Bank and the challenge of structural genomics. Nature Structural Biology. 2000; 7:957–959.

[10]. Fromme, P., editor. Photosynthetic Protein Complexes. Wiley-Blackwell; 2008.

[11]. Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, Krauss N. Three-dimensional structure of cyanobacterial Photosystem I at 2.5 angstrom resolution. Nature. 2001; 411:909–917. [PubMed: 11418848]

[12]. Fromme P, Mathis P. Unraveling the Photosystem I reaction center: a history, or the sum of many efforts. Photosynthesis Research. 2004; 80:109–124. [PubMed: 16328814]

[13]. Fromme, P.; Grotjohann, I. *Advances in Photosynthesis and Respiration*, chap. 6. Vol. vol. 24. Springer, Dordrecht; Netherlands: 2006. Structural analysis of cyanobacterial photosystem I; p. 47-69.

[14]. Fromme P, Witt HT. Improved isolation and crystallization of Photosystem I for structural analysis. Biochimica Et Biophysica Acta-Bioenergetics. 1998; 1365:175–184.

[15]. Mueller M, Jenni S, Ban N. Strategies for crystallization and structure determination of very large macromolecular assemblies. Current Opinion in Structural Biology. 2007; 17:572–579. [PubMed: 17964135]

[16]. Woolfson, M. An introduction to X-ray crystallography. 2nd edn.. Cambridge University Press; Cambridge: 1997.

[17]. Matthews BW. Solvent Content of Protein Crystals. Journal of Molecular Biology. 1968; 33:491–497. [PubMed: 5700707]

[18]. Witt I, Witt HT, Difiore D, Rogner M, Hinrichs W, Saenger W, Granzin J, Betzel C, Dauter Z. X-ray characterization of single-crystals of the reaction center-I of water splitting photosynthesis. Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics. 1988; 92:1503–1506.

[19]. Krauss N, Hinrichs W, Witt I, Fromme P, Pritzkow W, Dauter Z, Betzel C, Wilson KS, Witt HT, Saenger W. 3-dimensional structure of system-I of photosynthesis at 6 Angstrom resolution. Nature. 1993; 361:326–331.

[20]. Witt HT, Witt I, Krauss N, Hinrichs W, Fromme P, Saenger W. Crystals and structure of Photosystem-1. Biophysical Journal. 1994; 66:A2–A2.

[21]. Krauss N, Schubert WD, Klukas O, Fromme P, Witt HT, Saenger W. Photosystem I at 4 angstrom resolution represents the first structural model of a joint photosynthetic reaction centre and core antenna, system. Nature Structural Biology. 1996; 3:965–973.

[22]. Klukas O, Schubert WD, Jordan P, Krauss N, Fromme P, Witt HT, Saenger W. Photosystem I, an improved model of the stromal subunits PsaC, PsaD, and PsaE. Journal of Biological Chemistry. 1999; 274:7351–7360. [PubMed: 10066799]

[23]. Klukas O, Schubert WD, Jordan P, Krauss N, Fromme P, Witt HT, Saenger W. Localization of two phylloquinones, Q(K) and Q(K)', in an improved electron density map of Photosystem I at 4-angstrom resolution. Journal of Biological Chemistry. 1999; 274:7361–7367. [PubMed: 10066800]

[24]. Kurisu G, Zhang H, Smith JL, Cramer WA. Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. Science. 2003; 302:1009–14. [PubMed: 14526088]

[25]. Iwata S, Ostermeier C, Ludwig B, Michel H. Structure at 2.8-angstrom resolution of cytochrome-c-oxidase from *Paracoccus-denitrificans*. Nature. 1995; 376:660–669. [PubMed: 7651515]

[26]. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, ShinzawaItoh K, Nakashima R, Yaono R, Yoshikawa S. The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 angstrom. Science. 1996; 272:1136–1144. [PubMed: 8638158]

[27]. Koepke J, Olkhova E, Angerer H, Muller H, Peng GH, Michel H. High resolution crystal structure of Paracoccus denitrificans cytochrome c oxidase: New insights into the active site and the proton transfer pathways. Biochimica Et Biophysica Acta-Bioenergetics. 2009; 1787:635–645.

[28]. Iwata S, Saynovits M, Link TA, Michel H. Structure of a water soluble fragment of the 'Rieske' iron sulfur protein of the bovine heart mitochondrial cytochrome bc(1) complex determined by MAD phasing at 1.5 angstrom resolution. Structure. 1996:567–579. [PubMed: 8736555]

[29]. Iwata S, Lee JW, Okada K, Lee JK, Iwata M, Rasmussen B, Link TA, Ramaswamy S, Jap BK. Complete structure of the 11-subunit bovine mitochondrial cytochrome bc(1) complex. Science. 1998; 281:64–71. [PubMed: 9651245]

[30]. Zouni A, Witt HT, Kern J, Fromme P, Krauss N, Saenger W, Orth P. Crystal structure of photosystem II from Synechococcus elongatus at 3.8 angstrom resolution. Nature. 2001; 409:739–743. [PubMed: 11217865]

[31]. Loll B, Kern J, Saenger W, Zouni A, Biesiadka J. Towards complete cofactor arrangement in the 3.0 angstrom resolution structure of Photosystem II. Nature. 2005; 438:1040–1044. [PubMed: 16355230]

[32]. Umena Y, Kawakami K, Shen J-R, Kamiya N. Crystal structure of oxygen-evolving Photosystem II at a resolution of 1.9 Å. Nature. 2011; 473:55–60. [PubMed: 21499260]

[33]. Liu Z, Yan H, Wang K, Kuang T, Zhang J, Gui L, An X, Chang W. Crystal structure of spinach major light-harvesting complex at 2.72-Å resolution. Nature. 2004; 428:287–292. [PubMed: 15029188]

[34]. Amunts A, Drory O, Nelson N. The structure of a plant photosystem I supercomplex at 3.4 angstrom resolution. Nature. 2007; 447:58–63. [PubMed: 17476261]

[35]. Ben-Shem A, Frolow F, Nelson N. Crystal structure of plant photo-system I. Nature. 2003; 426:630–635. [PubMed: 14668855]

[36]. Amunts A, Toporik H, Borovikova A, Nelson N. Structure Determination and Improved Model of Plant Photosystem I. Journal of Biological Chemistry. 2010; 285:3478–3486. [PubMed: 19923216]

[37]. Malkin AJ, Kuznetsov YG, Land TA, Deyoreo JJ, McPherson A. Mechanisms of growth for protein and virus crystals. Nature Structural Biology. 1995; 2:956–959.

[38]. McPherson A, Malkin AJ, Kuznetsov YG. The science of macromolecular crystallization. Structure. 1995; 3:759–768. [PubMed: 7582893]

[39]. Gabrielsen M, Nagy LA, DeLucas LJ, Cogdell RJ. Self-interaction chromatography as a tool for optimizing conditions for membrane protein crystallization. Acta Crystallographica Section D-Biological Crystallography. 2010; 66:44–50.

[40]. Cherezov V, Caffrey M. Membrane protein crystallization in lipidic mesophases. A mechanism study using X-ray microdiffraction. Faraday Discuss. 2007; 136:195–212. [PubMed: 17955811]

[41]. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK. GPCR engineering yields high-resolution structural insights into β_2-adrenergic receptor function. Science. 2007; 318:1266–1273. [PubMed: 17962519]

[42]. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature. 2007; 450:383–387. [PubMed: 17952055]

[43]. Horsefield R, Norden K, Fellert M, Backmark A, Tornroth-Horsefield S, van Scheltinga ACT, Kvassman J, Kjellbom P, Johanson U, Neutze R. High-resolution X-ray structure of human aquaporin 5. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:13327–13332. [PubMed: 18768791]

[44]. Ago H, Kanaoka Y, Irikura D, Lam BK, Shimamura T, Austen KF, Miyano M. Crystal structure of a human membrane protein involved in cysteinyl leukotriene biosynthesis. Nature. 2007; 448:609–612. [PubMed: 17632548]

[45]. Chien EYT, Liu W, Zhao QA, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC. Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. Science. 2010; 330:1091–1095. [PubMed: 21097933]

[46]. Caffrey M. Crystallizing membrane proteins for structure-function studies using lipidic mesophases. Biochemical Society Transactions. 2011; 39:725–732. [PubMed: 21599641]

[47]. Rummel G, Hardmeyer A, Widmer C, Chiu ML, Nollert P, Locher KP, Pedruzzi I, Landau EM, Rosenbusch JP. Lipidic cubic phases: New matrices for the three-dimensional crystallization of membrane proteins. Journal of Structural Biology. 1998; 121:82–91. [PubMed: 9618339]

[48]. Chiu ML, Nollert P, Loewen MC, Belrhali H, Pebay-Peyroula E, Rosenbusch JP, Landau EM. Crystallization in cubo: general applicability to membrane proteins. Acta Crystallographica Section D-Biological Crystallography. 2000; 56:781–784.

[49]. Cherezov V, Clogston J, Papiz MZ, Caffrey M. Room to move: Crystallizing membrane proteins in swollen lipidic mesophases. Journal of Molecular Biology. 2006; 357:1605–1618. [PubMed: 16490208]

[50]. Wadsten P, Wohri AB, Snijder A, Katona G, Gardiner AT, Cogdell RJ, Neutze R, Engstrom S. Lipidic sponge phase crystallization of membrane proteins. Journal of Molecular Biology. 2006; 364:44–53. [PubMed: 17005199]

[51]. Popot JL, Berry EA, Charvolin D, Creuzenet C, Ebel C, Engelman DM, Flotenmeyer M, Giusti F, Gohon Y, Herve P, Hong Q, Lakey JH, Leonard K, Shuman HA, Timmins P, Warschawski DE, Zito F, Zoonens M, Pucci B, Tribet C. Amphipols: polymeric surfactants for membrane biology research. Cellular and Molecular Life Sciences. 2003; 60:1559–1574. [PubMed: 14513831]

[52]. Johansson LC, Wohri AB, Katona G, Engstrom S, Neutze R. Membrane protein crystallization from lipidic phases. Current Opinion in Structural Biology. 2009; 19:372–378. [PubMed: 19581080]

[53]. Nave C. A description of imperfections in protein crystals. Acta Crystallographica Section D-Biological Crystallography. 1998; 54:848–853.

[54]. Holton JM, Frankel KA. The minimum crystal size needed for a complete diffraction data set. Acta Crystallographica Section D-Biological Crystallography. 2010; 66:393–408.

[55]. Cusack S, Belrhali H, Bram A, Burghammer M, Perrakis A, Riekel C. Small is beautiful: protein micro-crystallography. Nature Structural Biology. 1998; 5:634–637.

[56]. Perrakis A, Cipriani F, Castagna JC, Claustre L, Burghammer M, Riekel C, Cusack S. Protein microcrystals and the design of a microdiractometer: current experience and plans at EMBL and ESRF/ID13. Acta Crystallographica Section D-Biological Crystallography. 1999; 55:1765–1770.

[57]. Bilderback DH, Brock JD, Dale DS, Finkelstein KD, Pfeifer MA, Gruner SM. Energy recovery linac (ERL) coherent hard X-ray sources. New Journal of Physics. 2010; 12:035011.

[58]. Meents A, Gutmann S, Wagner A, Schulze-Briese C. Origin and temperature dependence of radiation damage in biological samples at cryogenic temperatures. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:1094–1099. [PubMed: 20080548]

[59]. Howells MR, Beetz T, Chapman HN, Cui C, Holton JM, Jacobsen CJ, Kirz J, Lima E, Marchesini S, Miao H, Sayre D, Shapiro DA, Spence JCH, Starodub D. An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. Journal of Electron Spectroscopy and Related Phenomena. 2009; 170:4–12. [PubMed: 20463854]

[60]. Chapman HN. X-ray imaging beyond the limits. Nature Materials. 2009; 8:299–301.

[61]. Holton JM. A beginner's guide to radiation damage. Journal of Synchrotron Radiation. 2009; 16:133–142. [PubMed: 19240325]

[62]. Ravelli R, Leiros H, Pan B, Caffrey M, McSweeney S. Specific radiation damage can be used to solve macromolecular crystal structures. STRUCTURE. 2003; 11:217–224. [PubMed: 12575941]

[63]. Paithankar KS, Owen RL, Garman EF. Absorbed dose calculations for macromolecular crystals: improvements to RADDOSE. Journal of Synchrotron Radiation. 2009; 16:152–162. [PubMed: 19240327]

[64]. Henderson R. Cryoprotection of proteins crystals against radiation-damage in electron and X-ray diffraction. Proceedings of the Royal Society of London Series B-Biological Sciences. 1990; 241:6–8.

[65]. Hope H. Cryocrystallography of biological macromolecules - A generally applicable method. Acta Crystallographica Section B-Structural Science. 1988; 44:22–26.

[66]. Murray JW, Garman EF, Ravelli RBG. X-ray absorption by macromolecular crystals: the effects of wavelength and crystal composition on absorbed dose. Journal of Applied Crystallography. 2004; 37:513–522.

[67]. Henderson R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Quart. Rev. Biophys. 1995; 28:171–193.

[68]. Owen RL, Rudino-Pinera E, Garman EF. Experimental determination of the radiation dose limit for cryocooled protein crystals. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103:4912–4917. [PubMed: 16549763]

[69]. Yano J, Kern J, Irrgang KD, Latimer MJ, Bergmann U, Glatzel P, Pushkar Y, Biesiadka J, Loll B, Sauer K, Messinger J, Zouni A, Yachandra VK. X-ray damage to the Mn4Ca complex in single crystals of Photosystem II: A case study for metalloprotein crystallography. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:12047–12052. [PubMed: 16103362]

[70]. Garman E. Cool data: quantity AND quality. Acta Crystallographica Section D. 1999; 55(10): 1641–1653.

[71]. Garman E, Schneider T. Macromolecular cryocrystallography. Journal of Applied Crystallography. 1997; 30:211–237.

[72]. Porra RJ, Thompson WA, Kriedemann PE. Determination of accurate extinction coe cients and simultaneous-equations for assaying chlorophyll-a and chlorophyll-b extracted with 4 different solvents - verification of the concentration of chlorophyll standards by atomic-absorption spectroscopy. Biochimica Et Biophysica Acta. 1989; 975:384–394.

[73]. Altarelli M. From 3rd-to 4th-generation light sources: Free-electron lasers in the X-ray range. Crystallography Reports. 2010; 55:1145–1151.

[74]. Spence JCH, Doak RB. Single molecule diffraction. Physical Review Letters. 2004; 92:198102. [PubMed: 15169448]

[75]. Weierstall U, Doak RB, Spence JCH, Starodub D, Shapiro D, Kennedy P, Warner J, Hembree GG, Fromme P, Chapman HN. Droplet streams for serial crystallography of proteins. Experiments in Fluids. 2008; 44:675–689.

[76]. DePonte DP, Weierstall U, Schmidt K, Warner J, Starodub D, Spence JCH, Doak RB. Gas dynamic virtual nozzle for generation of microscopic droplet streams. Journal of Physics D-Applied Physics. 2008; 41:195505.

[77]. DePonte DP, Doak RB, Hunter M, Liu ZQ, Weierstall U, Spence JCH. SEM imaging of liquid jets. Micron. 2009; 40:507–509. [PubMed: 19246201]

[78]. Shapiro DA, Chapman HN, DePonte D, Doak RB, Fromme P, Hembree G, Hunter M, Marchesini S, Schmidt K, Spence J, Starodub D, Weierstall U. Powder diffraction from a continuous microjet of submicrometer protein crystals. Journal of Synchrotron Radiation. 2008; 15:593–599. [PubMed: 18955765]

[79]. Hunter MS, DePonte DP, Shapiro DA, Kirian RA, Wang X, Starodub D, Marchesini S, Weierstall U, Doak RB, Spence JCH, Fromme P. X-ray Diffraction from Membrane Protein Nanocrystals. Biophysical Journal. 2011; 100:198–206. [PubMed: 21190672]

[80]. Scherrer P. Determination of the size and inner structure of colloid particles by X-rays. Nach. Ges. Wiss. G+ttingen. 1918; 26:2.

[81]. Margiolaki I, Wright J. Powder crystallography on macromolecules. Aact Crystallographica Section A. 2008; 64:169–180.

[82]. Ayvazyan V, Baboi N, Bohnet I, Brinkmann R, Castellano M, Castro P, Catani L, Choroba S, Cianchi A, Dohlus M, Edwards HT, Faatz B, Fateev AA, Feldhaus J, Flottmann K, Gamp A, Garvey T, Genz H, Gerth C, Gretchko V, Grigoryan B, Hahn U, Hessler C, Honkavaara K, Huning M, Ischebeck R, Jablonka M, Kamps T, Korfer M, Krassilnikov M, Krzywinski J, Liepe M, Liero A, Limberg T, Loos H, Luong M, Magne C, Menzel J, Michelato P, Minty M, Muller UC, Nolle D, Novokhatski A, Pagani C, Peters F, Pfluger J, Piot P, Plucinski L, Rehlich K, Reyzl I, Richter A, Rossbach J, Saldin EL, Sandner W, Schlarb H, Schmidt G, Schmuser P, Schneider JR, Schneidmiller EA, Schreiber HJ, Schreiber S, Sertore D, Setzer S, Simrock S, Sobierajski R, Sonntag B, Steeg B, Stephan F, Sytchev KP, Tiedtke K, Tonutti M, Treusch R, Trines D, Turke D, Verzilov V, Wanzenberg R, Weiland T, Weise H, Wendt M, Will I, Wolff S, Wittenburg K, Yurkov MV, Zapfe K. Generation of GW radiation pulses from a VUV free-electron laser operating in the femtosecond regime. Physical Review Letters. 2002; 88:104802. [PubMed: 11909361]

[83]. Arthur J, Materlik G, Tatchyn R, Winick H. The LCLS - A fourth generation light-source using the SLAC linac. Review of Scientific Instruments. 1995; 66:1987–1989.

[84]. Shintake T, Tanaka T, Hara T, Togawa K, Inagaki T, Kim YJ, Ishikawa T, Kitamura H, Baba H, Matsumoto H, Takeda S, Yoshida M, Takasu Y. Status of SPring-8 compact SASV source FEL project. Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment. 2003; 507:382–387.

[85]. Geloni G, Saldin E, Samoylova L, Schneidmiller E, Sinn H, Tschentscher T, Yurkov M. Coherence properties of the European XFEL. New Journal of Physics. 2010; 12:035021.

[86]. Bonifacio R, Pellegrini C, Narducci LM. Collective instabilities and high-gain regime in a free-electron laser. Optics Communications. 1984; 50:373–378.

[87]. Young L, Kanter EP, Krassig B, Li Y, March AM, Pratt ST, Santra R, Southworth SH, Rohringer N, DiMauro LF, Doumy G, Roedig CA, Berrah N, Fang L, Hoener M, Bucksbaum PH, Cryan JP, Ghimire S, Glownia JM, Reis DA, Bozek JD, Bostedt C, Messerschmidt M. Femtosecond electronic response of atoms to ultra-intense X-rays. Nature. 2010; 466:56–66. [PubMed: 20596013]

[88]. Hau-Riege SP, London RA, Szoke A. Dynamics of biological molecules irradiated by short X-ray pulses. Physical Review E. 2004; 69:051906.

[89]. Neutze R, Wouts R, van der Spoel D, Weckert E, Hajdu J. Potential for biomolecular imaging with femtosecond X-ray pulses. Nature. 2000; 406:752–757. [PubMed: 10963603]

[90]. Chapman H, Ullrich J, Rost JM. Intense X-ray science: the first 5 years of FLASH. Journal of Physics B-Atomic Molecular and Optical Physics. 43

[91]. Chapman HN, Barty A, Bogan MJ, Boutet S, Frank M, Hau-Riege SP, Marchesini S, Woods BW, Bajt S, Benner H, London RA, Plonjes E, Kuhlmann M, Treusch R, Dusterer S, Tschentscher T, Schneider JR, Spiller E, Moller T, Bostedt C, Hoener M, Shapiro DA, Hodgson KO, Van der Spoel D, Burmeister F, Bergh M, Caleman C, Huldt G, Seibert MM, Maia F, Lee RW, Szoke A, Timneanu N, Hajdu J. Femtosecond diffractive imaging with a soft-X-ray free-electron laser. Nature Physics. 2006; 2:839–843.

[92]. Hau-Riege SP, Boutet S, Barty A, Bajt S, Bogan MJ, Frank M, Andreasson J, Iwan B, Seibert MM, Hajdu J, Sakdinawat A, Schulz J, Treusch R, Chapman HN. Sacrificial Tamper Slows Down Sample Explosion in FLASH Diffraction Experiments. Physical Review Letters. 2010; 104:064801. [PubMed: 20366823]

[93]. Emma P, Akre R, Arthur J, Bionta R, Bostedt C, Bozek J, Brachmann A, Bucksbaum P, Coffee R, Decker FJ, Ding Y, Dowell D, Edstrom S, Fisher A, Frisch J, Gilevich S, Hastings J, Hays G, Hering P, Huang Z, Iverson R, Loos H, Messerschmidt M, Miahnahri A, Moeller S, Nuhn HD, Pile G, Ratner D, Rzepiela J, Schultz D, Smith T, Stefan P, Tompkins H, Turner J, Welch J, White W, Wu J, Yocky G, Galayda J. First lasing and operation of an angstrom-wavelength free-electron laser. Nature Photonics. 2010; 4:641–647.

[94]. Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia F, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, Rolles D, Rudenko A, Foucar L, Kimmel N, Weidenspointner G, Holl P, Liang MN, Barthelmess M, Caleman C, Boutet S, Bogan MJ, Krzywinski J, Bostedt C, Bajt S, Gumprecht L, Rudek B, Erk B, Schmidt C, Homke A, Reich C, Pietschner D, Struder L, Hauser G, Gorke H, Ullrich J, Herrmann S, Schaller G, Schopper F, Soltau H, Kuhnel KU, Messerschmidt M, Bozek JD, Hau-Riege SP, Frank M, Hampton CY, Sierra RG, Starodub D, Williams GJ, Hajdu J, Timneanu N, Seibert MM, Andreasson J, Rocker A, Jonsson O, Svenda M, Stern S, Nass K, Andritschke R, Schroter CD, Krasniqi F, Bott M, Schmidt KE, Wang XY, Grotjohann I, Holton JM, Barends TRM, Neutze R, Marchesini S, Fromme R, Schorb S, Rupp D, Adolph M, Gorkhover T, Andersson I, Hirsemann H, Potdevin G, Graafsma H, Nilsson B, Spence JCH. Femtosecond X-ray protein nanocrystallography. Nature. 2011; 470:73–81. [PubMed: 21293373]

[95]. Kirian RA, Wang XY, Weierstall U, Schmidt KE, Spence JCH, Hunter M, Fromme P, White T, Chapman HN, Holton J. Femtosecond protein nanocrystallography-data analysis methods. Optics Express. 2010; 18:5713–5723. [PubMed: 20389587]

[96]. Kirian RA, White TA, Holton JM, Chapman HN, Fromme P, Barty A, Lomb L, Aquila A, Maia F, Martin AV, Fromme R, Wang XY, Hunter MS, Schmidt KE, Spence JCH. Structure-factor analysis of femtosecond micro-diffraction patterns from protein nanocrystals. Acta Crystallographica Section A. 2011; 67:131–140.

[97]. Bozek JD. AMO instrumentation for the LCLS X-ray FEL. European Physical Journal-Special Topics. 2009; 169:129–132.

[98]. Struder L, Eppa S, Rolles D, Hartmann R, Holl P, Lutz G, Soltau H, Eckart R, Reich C, Heinzinger K, Thamm C, Rudenko A, Krasniqi F, Kuhnel KU, Bauer C, Schroter CD, Moshammer R, Techert S, Miessner D, Porro M, Halker O, Meidinger N, Kimmel N, Andritschke R, Schopper F, Weidenspointner G, Ziegler A, Pietschner D, Herrmann S, Pietsch U, Walenta A, Leitenberger W, Bostedt C, Moller T, Rupp D, Adolph M, Graafsma H, Hirsemann H, Gartner K, Richter R, Foucar L, Shoeman RL, Schlichting I, Ullrich J. Large-format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources. Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment. 2010; 614:483–496.

[99]. Leslie AGW. The integration of macromolecular di raction data. Acta Crystallographica Section D-Biological Crystallography. 2006; 62:48–57.

[100]. Steller I, Bolotovsky R, Rossmann MG. An algorithm for automatic indexing of oscillation images using Fourier analysis. Journal of Applied Crystallography. 1997; 30:1036–1040.

[101]. Duisenberg AJM. Indexing in single-crystal diffractometry with an obstinate list of reflections. Journal of Applied Crystallography. 1992; 25:92–96.

[102]. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallographica Section D-Biological Crystallography. 1997; 53:240–255.

[103]. Scott HA. Cretin - a radiative transfer capability for laboratory plasmas. Journal of Quantitative Spectroscopy & Radiative Transfer. 2001; 71:689–701.

[104]. Burmeister WP. Structural changes in a cryocooled protein crystal owing to radiation damage. Acta Crystallographica Section D-Biological Crystallography. 2000; 56:328–341.

[105]. Marchesini S, He H, Chapman HN, Hau-Riege SP, Noy A, Howells MR, Weierstall U, Spence JCH. X-ray image reconstruction from a diffraction pattern alone. Physical Review B. 2003; 68:140101.

[106]. Robinson I, Harder R. Coherent X-ray diffraction imaging of strain at the nanoscale. Nature Materials. 2009; 8:291–298.

[107]. Sayre D. Some implications of a theorem due to Shannon. Acta Crystallographica. 1952; 5:843–843.

[108]. Spence J, Kirian R, Wang X, Weierstall U, Schmidt K, White T, Barty A, Chapman H, Marchesini S, Holton J. Phasing of coherent femtosecond X-ray diffraction from size-varying nanocrystals. Optics Express. 2011; 19:2866–2873. [PubMed: 21369108]

[109]. Boutet S, Williams GJ. The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). New Journal of Physics. 2010; 12:035024.

[110]. Whri AB, Katona G, Johansson LC, Fritz E, Malmerberg E, Andersson M, Vincent J, Eklund M, Cammarata M, Wulff M, Davidsson J, Groenhof G, Neutze R. Light-Induced Structural Changes in a Photosynthetic Reaction Center Caught by Laue Diffraction. Science. 2010; 328:630–633. [PubMed: 20431017]

[111]. Schotte F, Soman J, Olson J, Wulff M, Anfinrud P. Picosecond time-resolved X-ray crystallography: probing protein function in real time. Science. 2004; 147:235–246.

[112]. Moffat K, Szebenyi D, Bilderback D. X-ray Laue diffraction from protein crystals. Science. 1984; 223:1423–1425. [PubMed: 17746054]

[113]. Moffat K. The frontiers of time-resolved macromolecular crystallography: movies and chirped X-ray pulses. Faraday Discussions. 2003; 122:65–77. [PubMed: 12555850]

[114]. Graber T, Anderson S, Brewer H, Chen YS, Cho HS, Dashdorj N, Henning RW, Kosheleva I, Macha G, Meron M, Pahl R, Ren Z, Ruan S, Schotte F, Rajer VS, Viccaro PJ, Westferro F, Anfinrud P, Moffat K. BioCARS: a synchrotron resource for time-resolved X-ray science. Journal of Synchrotron Radiation. 2011; 18:658–670. [PubMed: 21685684]

[115]. Schotte F, Lim M, Jackson T, Smirnov A, Soman J, Olson J, Phillips G, Wulff M, Anfinrud P. Watching a protein as it functions with 150-ps time-resolved X-ray crystallography. Science. 2003; 300:1944–1947. [PubMed: 12817148]

[116]. Schrödinger, LLC. The PyMOL Molecular Graphics System. Version 1.3r1. 2010.

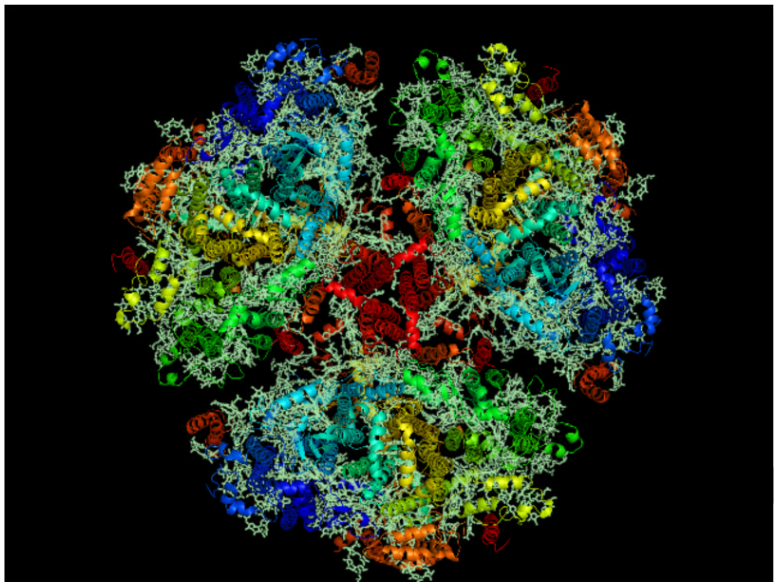


Fig. 1. Structure of cyanobacterial Photosystem I

Membrane normal image of the trimeric Photosystem I from cyanobacteria using RSCB PDB (www.pdb.org, [9]) ID 1JB0 at 2.5-Å resolution [11] created using Py-Mol [116]. Cyanobacterial Photosystem I is the largest membrane protein complex solved to molecular resolution.

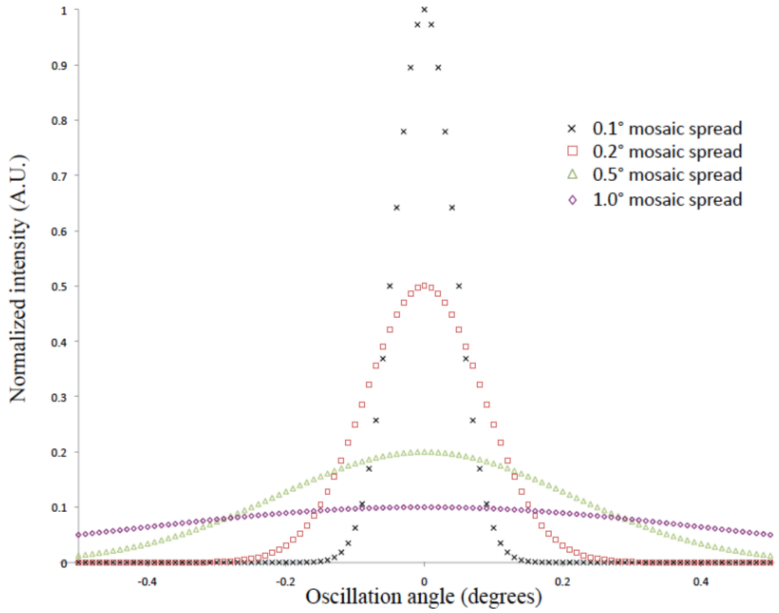


Fig. 2. Effects of mosaicity

Plot of Gaussian peak profiles showing the effect of increase mosaicity. The mosaicity is defined as the full-width at half-maximum of the peak and is related to the variance by FWHM = $2.3548(\sigma^2)^{1/2}$. The peak height has a linear dependence on the standard deviation of the Gaussian peak. The higher the mosaicity, the lower the peak maximum and the broader the peak width, which can have major implications for high-resolution data.

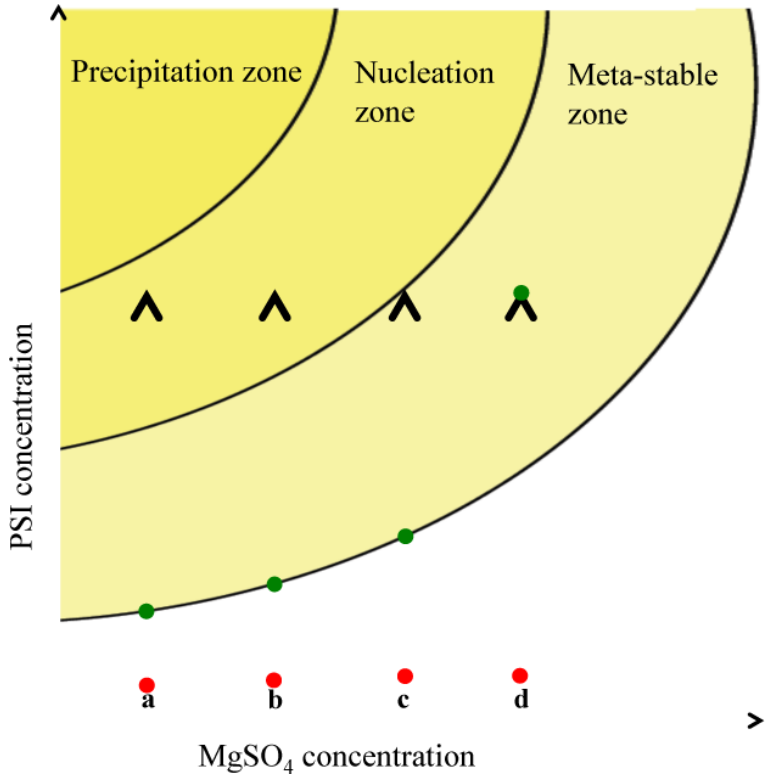
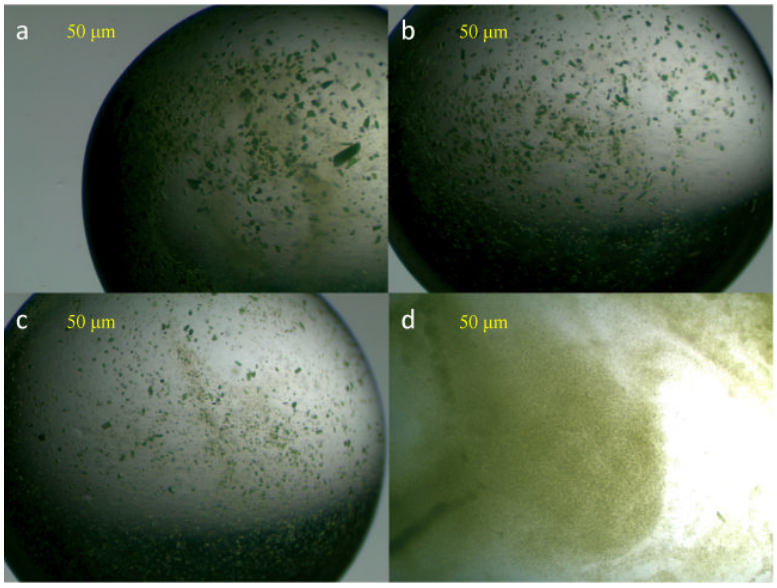


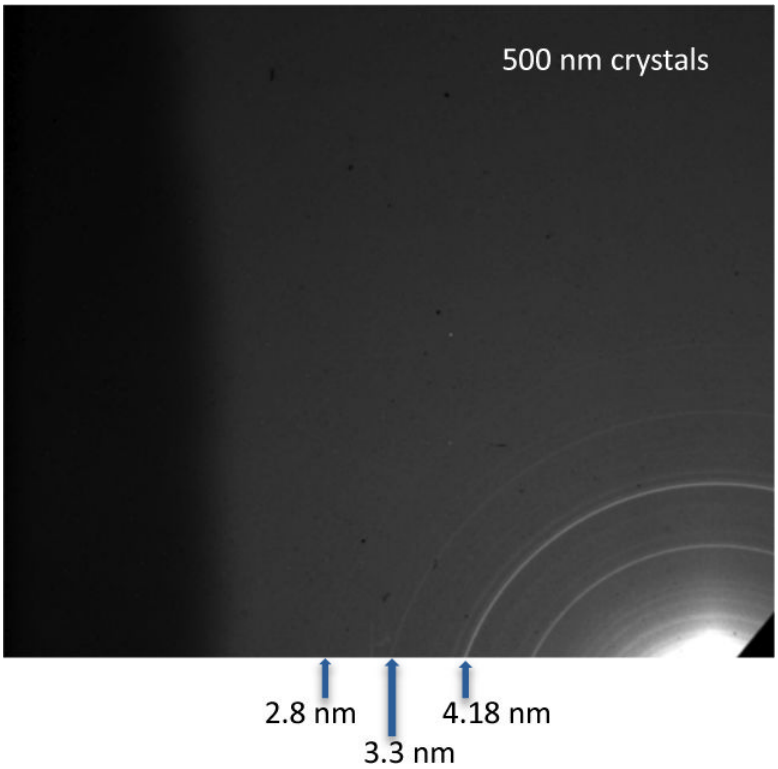
Fig. 3. Schematic phase diagram of Photosystem I

Schematic phase diagram of Photosystem I showing the ultrafiltration crystallization method. Lines a–d indicate different starting $MgSO_4$ concentrations for the crystallization experiments, in which the red circles indicate the starting position of a crystallization experiment and the green circles show the equilibrium positions of the sample in the phase diagram. The crystallization experiments (a,b,c) would lead to a shower of crystals and crystallites, whereas (d) would need further concentrating to produce crystalline sample. (a) would produce the conditions that led to the largest number and fastest growing crystals, whereas (c) would produce the smallest number of crystals that grew slowly comparative to (a).



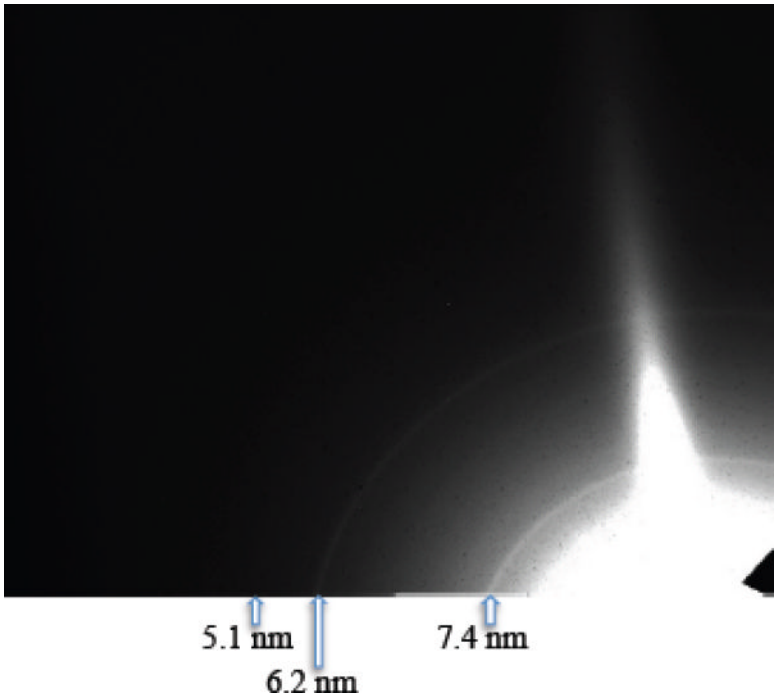
4. Ultrafiltration crystallization results for PSI

Images of the crystals contained within the Photosystem I crystal suspensions after the solution is allowed to settle for 10 min (a), 20 min (b), 30 min (c), and 40 min (d).



5. Powder diffraction from 500-nm PSI crystallites

(a) Diffraction patterns from ≤ 500 nm PSI crystals using an off-axis detector to increase measurable scattering angle. The crystals were irradiated with 1560 eV X-rays (λ = 0.8 nm) and had a resolution of 1.3 nm in the upper-left corner. Less than 6000 unit cells are in each of the ≤ 500 nm PSI crystals, yet powder diffraction is seen to 28-Å resolution, at which point the semi-discrete powder rings come a continuous background due to the large unit cell dimensions of photosystem I. Image and caption modified from [79]



6. Powder diffraction from 100-nm PSI crystallites

Diffraction pattern from 100-nm PSI crystals with an off-axis detector and an X-ray energy of 520 eV (λ = 2.4 nm). The resolution at the corner of the detector is 4.0 nm. Less than 100 PSI trimers are in each crystal. Image taken from and caption modified from [79].

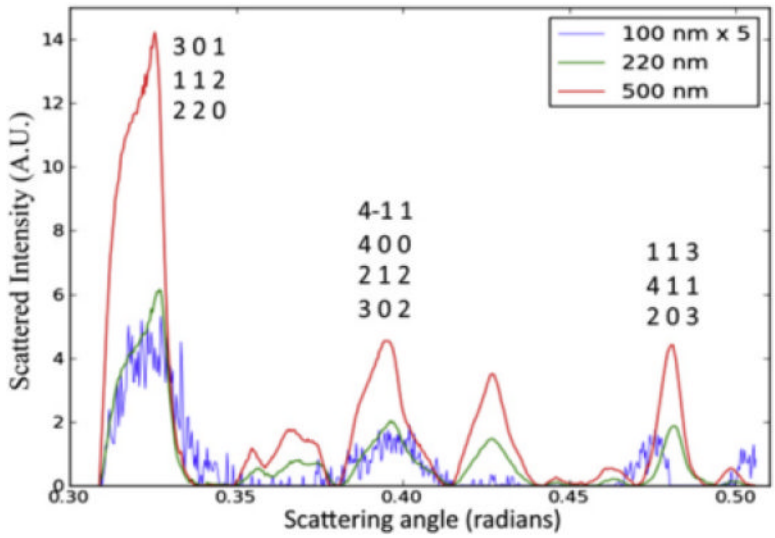


Fig. 7. Radially integrated PSI powder diffraction patterns

A comparison of the scattering power of the 100 nm, 220 nm, and 500 nm crystals at specific Bragg reflections using the scattered intensity as a function of scattering angle. Figure and caption taken from [79].

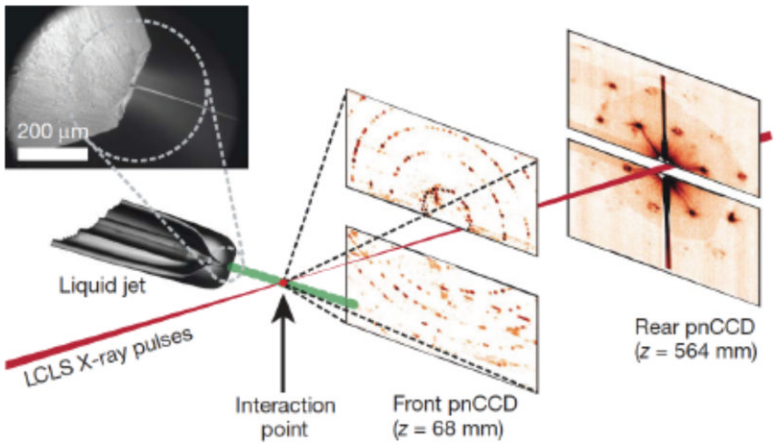


Fig. 8. Femtosecond nanocrystallography

Photosystem I nanocrystals flow in their buffer G_0 in a gas-focused 4-µm diameter jet, perpendicular to the X-ray beam that is focused on the jet. The insert shows an environmental scanning electron micrograph of the nozzle, flowing jet, and focusing gas [76]. Two pairs of high-frame rate pnCCD detectors [98] record low and high-angle diffraction from single X-ray FEL pulses, at a rate of 30 Hz. Crystals arrive at random times and orientations in the beam, and the probability of hitting one is proportional to the crystal concentration. Figure and caption taken from [94] with slight modifications.

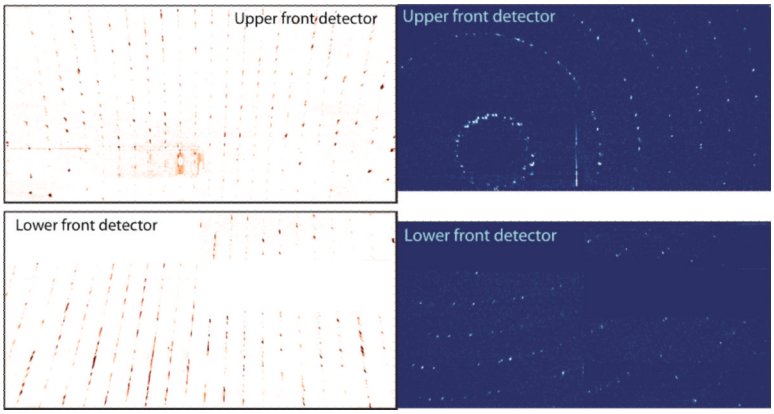


Fig. 9. Single-shot crystal diffraction patterns of PSI nanocrystals

(a, Left) A pattern from a crystal oriented with its *c* axis almost horizontal, perpendicular to the X-ray beam. (b, Right) Reciprocal lattice planes are seen as arcs, due to the intersection of the Ewald sphere with the lattice. Cuts through the shape transform, due to the finite crystal size, are evident when the intersection of the Ewald sphere with the reciprocal lattice planes is close to tangential, such as the almost continuous lines of intensity in the bottom left of (a) and the spots in the complete circle of reflection in the top left of (b). Figure and caption taken from [94]

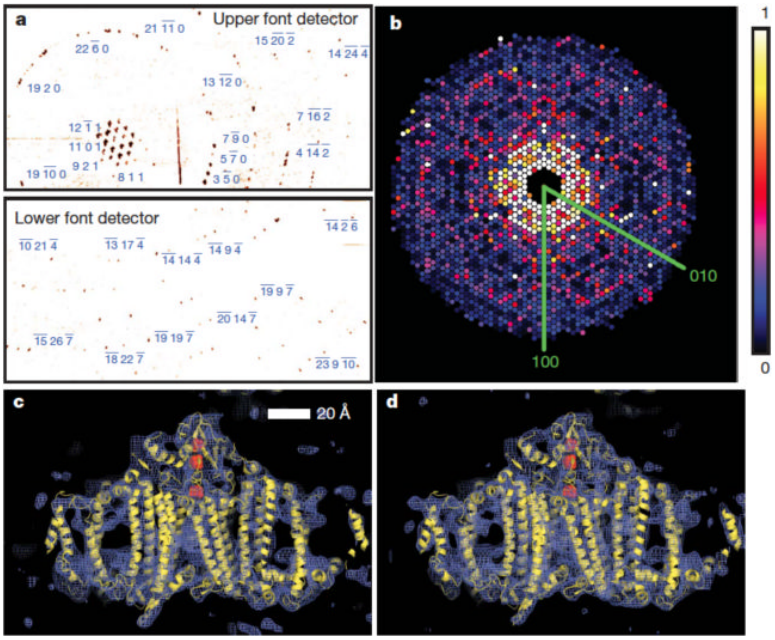


Fig. 10. Indexed diffraction patterns and electron density of PSI from the LCLS data

(a) Diffraction pattern recorded on the front pnCCDs with a single 70-fs pulse after background subtraction and correction of saturated pixels. Some peaks are labeled with their Miller indices. The resolution in the lower detector corner is 8.5 Å, (b) Precision-style pattern of the [001] zone for Photosystem I, obtained from merging femtosecond nanocrystal data from over 15,000 nanocrystal patterns, displayed on the linear color scale shown on the right. (c) Region of the the $2mF_0 2DF_c$ electron density map at 1.0σ (purple mesh), calculated from the 70-fs data and (d) from conventional synchrotron data truncated at a resolution of 8.5 Å and collected at a temperature of 100 K. The refined model of PSI is depicted in yellow. Figure and caption modified from [94].

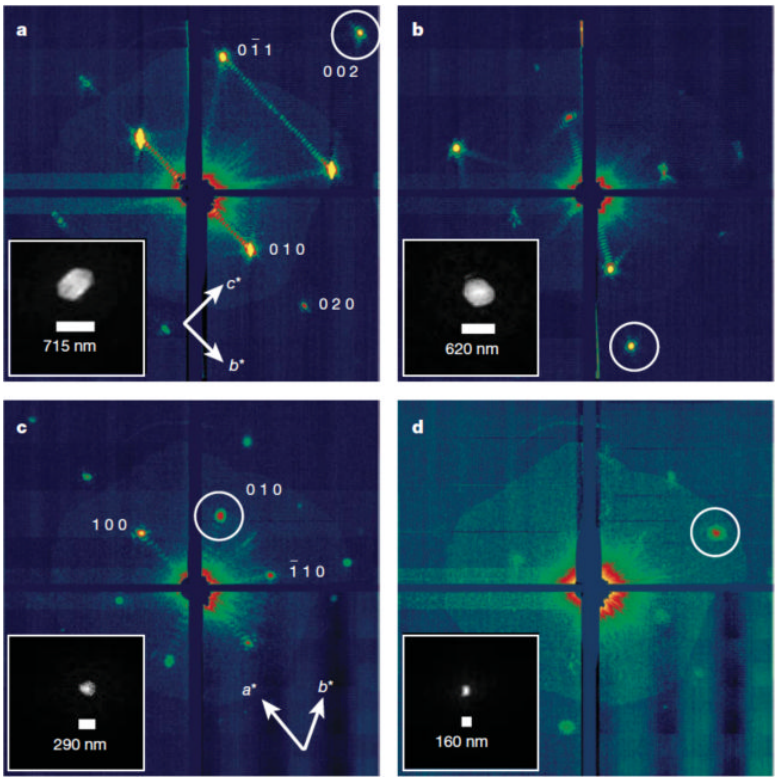


Fig. 11. Coherent crystal diffraction

Low-angle diffraction patterns recorded on the rear pnCCDs, revealing coherent diffraction from the structure of the PSI nanocrystals, shown using a logarithmic, false-color scale. The intensities away from the Bragg condition are an interference effect caused by the small number of unit cells in the crystal. The Miller indices of the peaks in (a) were determined from the corresponding high-resolution pattern. In (c) seven fringes are counted in the b* direction, corresponding to nine unit cells, or 250-nm PSI crystal. Insets, real-space images of the nanocrystal, determined by phase retrieval (using the Shrinkwrap algorithm [105] of the circled coherent Bragg shape transform). Figure and caption modified from [94].

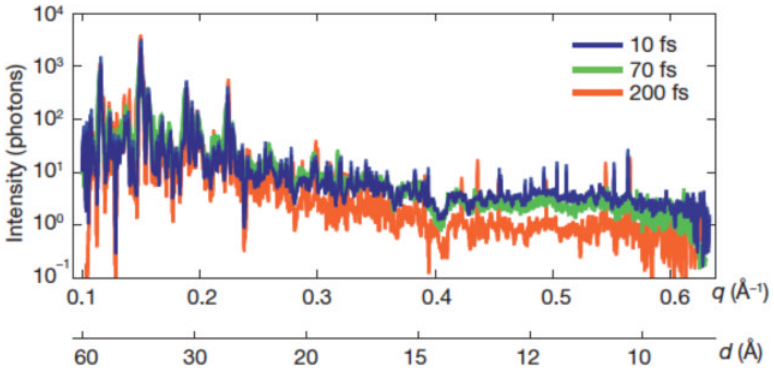


Fig. 12. Pulse-duration dependence of diffraction intensities

Plot of the integrated Bragg intensities of Photosystem-I-nanocrystal diffraction as a function of photon momentum transfer, $q = (4\pi/\lambda)\sin(\theta) = 2\pi/d$, (wavelength, λ; scattering angle, 2θ; resolution, *d*) for pulse durations of 10, 70 and 200 fs. Averages were obtained by isolating Bragg spots from 97,883, 805,311, and 66,063 patterns, respectively, normalized to pulse fluence. The error in each plot is indicated by the thickness of the line. The decrease in scattered intensity for 200-fs pulses and $d < 25$ Å indicates radiation damage for these long pulses, which is not apparent for 70-fs pulses and shorter.. Figure and caption modified from [94].

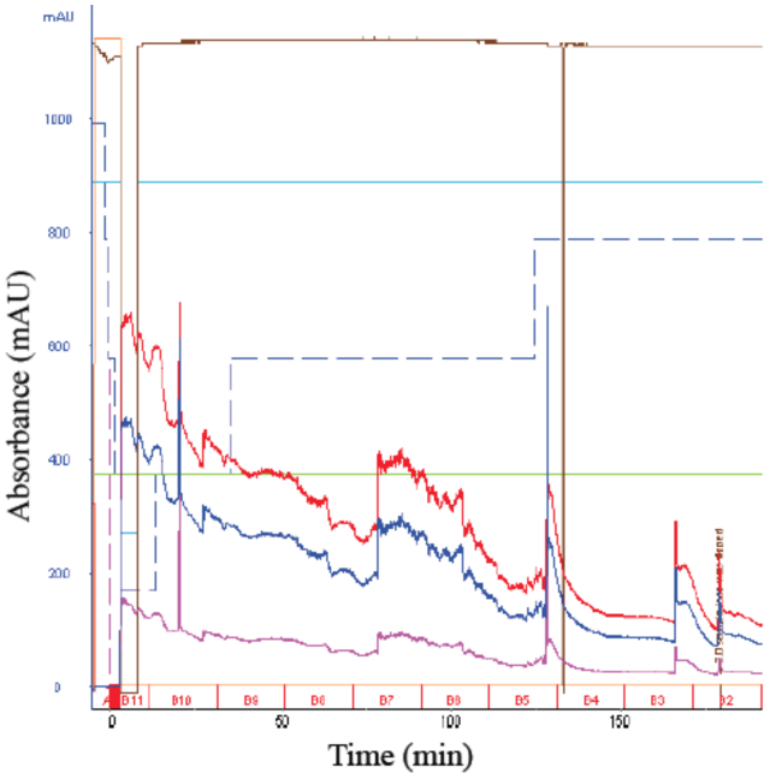


Fig. 13. Photosystem I crystallite settling with time

Plot of the absorbance at 680 nm (red), 600 nm (fushia), and 280 nm (blue) of suspension of 2-μm-filtered Photosystem I crystallites as monitored using an HPLC flowing at 10 mL/min. Sudden increases in absorbance at around 75 min and 130 min correspond to the sample loop being inverted.

Table 1

Statistics for the merged dataset of PSI at 70-fs pulse duration $R_{iso}(F)$ compares the LCLS structure factors to a conventional single-crystal dataset collected at the ALS beamline 8.2.2. The bottom center edge of the bottom detector occurs at a resolution length of 10.0 Å. Table and caption modified from [94]

Resolution shell (Å)	Number of unique reflections	Multiplicity	Completeness (%)	<Peak height/Noise>	Merged I/σ	$R_{iso}(F)$ (%)
30.0	416	1369	100	889	16.8	37.6
16.2	382	1278	100	29.3	15.9	15.1
13.6	362	1195	100	19.2	16.2	12.3
12.1	366	1009	100	20.3	14.7	12.9
11.1	357	678	100	19.5	12.7	14.6
10.4	355	492	100	15.4	10.6	28.7
9.85	349	283	100	14.0	9.20	22.2
9.39	352	111	100	11.8	7.12	28.1
9.01	352	28.6	99.4	9.50	4.99	24.8
8.68	88	3.92	25.1	6.70	3.29	30.8
Overall	3379	717	92.8	123	11.9	22.1