IUCrJ

Comparing serial X-ray crystallography and microcrystal electron diffraction (MicroED) as methods for routine structure determination from small macromolecular crystals

Alexander M. Wolff,[a,b] Iris D. Young,[b,c] Raymond G. Sierra,[d] Aaron S. Brewster,[c] Michael W. Martynowycz,[e,f] Eriko Nango,[g,h] Michihiro Sugahara,[g] Takanori Nakane,[i]‡ Kazutaka Ito,[b,j] Andrew Aquila,[d] Asmit Bhowmick,[c] Justin T. Biel,[a,b] Sergio Carbajo,[d] Aina E. Cohen,[k] Saul Cortez,[l] Ana Gonzalez,[k] Tomoya Hino,[m,n] Dohyun Im,[h] Jake D. Koralek,[d] Minoru Kubo,[g,o] Tomas S. Lazarou,[p] Takashi Nomura,[g] Shigeki Owada,[g,q] Avi J. Samelson,[r] Tomoyuki Tanaka,[g,h] Rie Tanaka,[g,h] Erin M. Thompson,[b,s] Henry van den Bedem,[t] Rahel A. Woldeyes,[b,s] Fumiaki Yumoto,[u] Wei Zhao,[v] Kensuke Tono,[g,q] Sebastien Boutet,[d] So Iwata,[g,h] Tamir Gonen,[e,f,w] Nicholas K. Sauter,[c] James S. Fraser[b] and Michael C. Thompson[b]*

[a]Graduate Program in Biophysics, University of California, San Francisco, San Francisco, California, USA, [b]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA, [c]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, [d]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, California, USA, [e]Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, California, USA, [f]Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, California, USA, [g]RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan, [h]Department of Cell Biology, Graduate School of Medicine, Kyoto University, Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan, [i]Department of Biological Science, Graduate School of Science, The University of Tokyo, Tokyo, Japan, [j]Laboratory for Drug Discovery, Pharmaceuticals Research Center, Asahi Kasei Pharma Corporation, 632-1 Mifuku, Izunokuni-shi, Shizuoka 410-2321, Japan, [k]SSRL, SLAC National Accelerator Laboratory, Menlo Park, California, USA, [l]Department of Biology, San Francisco State University, San Francisco, California, USA, [m]Department of Chemistry and Biotechnology, Graduate School of Engineering, Tottori University, 4-101 Koyama-cho, Minami, Tottori 680-8552, Japan, [n]Center for Research on Green Sustainable Chemistry, Tottori University, Tottori, Japan, [o]Graduate School of Life Science, University of Hyogo, Ako-gun, Hyogo 678-1297, Japan, [p]Department of Chemistry, New York University, New York, USA, [q]Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan, [r]Institute for Neurodegenerative Diseases, University of California, San Francisco, San Francisco, California, USA, [s]Graduate Program in Chemistry and Chemical Biology, University of California, San Francisco, San Francisco, California, USA, [t]Bioscience Department, SLAC National Accelerator Laboratory, Menlo Park, California, USA, [u]Structural Biology Research Center, Institute of Materials Structure Science, KEK/High Energy Accelerator Research Organization, Tsukuba, Ibaraki 305-0034, Japan, [v]Department of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, USA, and [w]Department of Physiology, University of California, Los Angeles, Los Angeles, California, USA. *Correspondence e-mail: mct.ucsf@gmail.com

Innovative new crystallographic methods are facilitating structural studies from ever smaller crystals of biological macromolecules. In particular, serial X-ray crystallography and microcrystal electron diffraction (MicroED) have emerged as useful methods for obtaining structural information from crystals on the nanometre to micrometre scale. Despite the utility of these methods, their implementation can often be difficult, as they present many challenges that are not encountered in traditional macromolecular crystallography experiments. Here, XFEL serial crystallography experiments and MicroED experiments using batch-grown microcrystals of the enzyme cyclophilin A are described. The results provide a roadmap for researchers hoping to design macromolecular microcrystallography experiments, and they highlight the strengths and weaknesses of the two methods. Specifically, we focus on how the different physical conditions imposed by the sample-preparation and delivery methods required for each type of experiment affect the crystal structure of the enzyme.

Unit-cell distribution

1. Introduction

In macromolecular crystallography, collecting full data sets from small crystals has been a challenge because their weaker
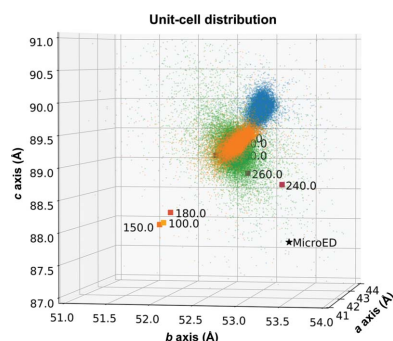
diffracting power limits the amount of signal that can be successfully obtained before the effects of X-ray radiation damage become significant. Thus, crystallographers have always been faced with a practical need to either optimize the growth of relatively large crystals or to make the most of smaller crystals by implementing clever data-collection and merging strategies (Cusack *et al.*, 1998; Zander *et al.*, 2015). Methodological advances that have facilitated the measurement of diffraction data from smaller and smaller crystals, such as the introduction of crystal cryocooling and the development of microfocus X-ray beams, have enabled the structure determination of increasingly challenging targets, for which large crystals could only be obtained with difficulty, if at all (Liu *et al.*, 2013; Zhou *et al.*, 2016). Additionally, small crystals have proven to be advantageous in a number of other contexts. For example, when crystals display long-range disorder, such as high mosaicity, a reduction in the total number of mosaic blocks reduces the spread of Bragg peaks and generally improves the overall data quality (Chernov, 1999). Small crystals are also advantageous when the diffraction experiment is preceded by a perturbation to the crystal. This includes common crystal treatments such as flash-cooling or ligand soaking, as well as more uncommon perturbations such as the stimulation of crystallized molecules for time-resolved experiments (Coquelle *et al.*, 2018; Olmos *et al.*, 2018). Because they have substantially less volume and a limited number of unit cells, perturbations can be applied more rapidly and uniformly to smaller crystals than to larger crystals, and smaller crystals accumulate less strain resulting from changes in crystal lattice dimensions. The development of protein 'microcrystallography' techniques, which are optimized for measuring crystals with physical dimensions of tens of micrometres or smaller, has offered access to these opportunities and benefits, and has led to a shift in what is considered to be a valuable specimen for experimental characterization.

The past decade has seen an explosion of new technologies for protein microcrystallography. The increased brightness available for crystallography at modern X-ray light sources, including synchrotrons and X-ray free-electron lasers (XFELs), led to the development of 'serial crystallography.' In a serial crystallography experiment the X-ray beam is typically very bright and tightly focused, so that extremely short exposure times produce measurable diffraction images, even for very small crystals (Chapman *et al.*, 2011). The intense X-ray beams used in these experiments destroy the samples rapidly, allowing only a single diffraction image to be collected per crystal. Therefore, the crystals must be rapidly (or 'serially') replenished at the X-ray interaction point in order for the experiment to be efficient. By measuring single diffraction snapshots of many randomly oriented crystals, it is possible to completely sample the reflections in reciprocal space and integrate the Bragg intensities. Importantly, serial crystallography experiments are generally conducted at room temperature, potentially providing more physiologically relevant insight into molecular structure by avoiding the artifacts associated with cryocooling.

Alongside the development of serial crystallography, major recent breakthroughs have been made in the field of microcrystal electron diffraction (MicroED). Specifically, advancements have resulted in the facile collection of continuous rotation data sets (Nannenga *et al.*, 2014; Shi *et al.*, 2013) from flash-cooled microcrystals using a transmission electron microscope that is equipped with a compustage and a cryoholder and operating in diffraction mode. Because the microscopes required for MicroED are now widespread as a result of the booming interest in electron cryomicroscopy (cryoEM), MicroED holds great potential for the determination of both protein and small-molecule structures (Nannenga & Gonen, 2019). Collectively, these new frontiers in macromolecular microcrystallography have created new opportunities for structural biology. Examples include structure determination from crystals as small as a few hundred nanometres in each of their dimensions (Chapman *et al.*, 2011; Nannenga & Gonen, 2019; de la Cruz *et al.*, 2017), and a new generation of challenging time-resolved measurements (Young *et al.*, 2016; Nango *et al.*, 2016) at high spatial and temporal resolution.

Despite the interesting possibilities that are now within reach, the optimization of sample-preparation and data-collection protocols for microcrystallography experiments remains challenging. Firstly, it is necessary to decide which measurement technique (*i.e.* X-ray or MicroED) is best suited to a given sample or research question. Then, if appropriately sized crystals are not obtained serendipitously, the experimenter must generate microcrystals with the correct size and density (crystals per microlitre), either by targeted growth or by the manipulation of larger crystal specimens. Next, it is essential to choose an appropriate method for delivering the microcrystals to the X-ray beam or to the column of the electron microscope. In the case of serial X-ray crystallography, multiple strategies have been explored for rapidly replenishing crystals at the X-ray interaction point. In addition to fixed-target approaches (Baxter *et al.*, 2016; Mueller *et al.*, 2015; Fuller *et al.*, 2017; Hunter *et al.*, 2014), in which crystals are mounted on a solid support and moved through the X-ray interaction region using automation, several methods have come into widespread use that exploit microfluidics to create free-standing streams, or 'jets', of microcrystal slurries. Various different types of microfluidic devices, collectively referred to as 'sample injectors', have been developed for this purpose (Sierra *et al.*, 2012; Weierstall *et al.*, 2014). Each type uses a different physical principle for carrying microcrystals to the X-ray beam by generating a stream of flowing liquid that is tens to hundreds of micrometres in diameter, and each method subjects the crystals to different conditions which could potentially affect the quality of the data acquired or the structure of the molecule itself. These conditions include exposure to strong electric fields (Sierra *et al.*, 2012), high pressure (Weierstall *et al.*, 2014) and additives (Sugahara *et al.*, 2017) that change the chemical properties of mother-liquor solutions.

Similarly, in MicroED several sample-preparation methods have been reported, such as direct pipetting of nanocrystals

onto EM grids (Rodriguez *et al.*, 2015), sonication, vortexing, vigorous pipetting or crushing to break large crystals into fragments and create a nanocrystal slurry (de la Cruz *et al.*, 2017). Crystals are then drop-cast onto EM grids, as traditionally used for cryoEM, and the grids are then blotted of excess solvent and flash-frozen in supercooled ethane. Once frozen, microcrystals prepared in this fashion can be used directly for data collection, or they can be subjected to a milling procedure that utilizes a scanning electron microscope with a focused ion beam (FIB-SEM) to create crystalline lamellae of the desired thickness (Duyvesteyn *et al.*, 2018; Martynowycz *et al.*, 2019*a*,*b*). Because electrons interact with matter more strongly than X-rays do, the ideal crystal thickness for MicroED measurements is only several hundred nanometres (Martynowycz *et al.*, 2019*b*), and the milling process is critical for samples that exceed this thickness. As for serial crystallography, sample preparation and delivery for MicroED involves subjecting crystals to unusual conditions that are not typically encountered when samples are prepared for traditional crystallographic experiments. These conditions include dehydration and exposure to shear forces that are produced by the flow of solvent during blotting (Martynowycz *et al.*, 2019*a*), as well as the potential damage induced by FIB milling. All of these considerations create a complex landscape, and designing the best experiment for a new system of interest is often a nontrivial process. This calculus is further complicated by the fact that the extent to which the unusual experimental conditions affect the quality of data acquired, or the structure of the molecule itself, has not been rigorously characterized.

Here, we discuss the planning, optimization and execution of protein microcrystallography experiments, using human cyclophilin A (CypA) as a model protein system. CypA is a proline isomerase enzyme that is highly abundant in human cells and plays important biological roles both as a protein-folding chaperone and a modulator of intracellular signaling pathways. Prior work has shown that CypA readily forms large crystals, which have been successfully used for traditional rotation crystallography at synchrotron X-ray sources and for fixed-target measurements at an XFEL source (Fraser *et al.*, 2009; Keedy *et al.*, 2015). Starting from crystallization conditions that produce large (hundreds of micrometres in each dimension) CypA crystals, we optimized the preparation of high-density microcrystal slurries. We then used these samples for a variety of microcrystallography experiments, including serial X-ray crystallography with three different microfluidic sample injectors, and MicroED. Because the data collected across the different types of experiments were derived from similarly prepared microcrystal samples and analyzed using comparable protocols, we were able to perform a rigorous comparison of the results. For each method, we evaluate the ease of sample preparation and delivery, the statistical quality of the measured data and the properties of the resulting atomic models. Our results illustrate the inherent strengths and weaknesses of these new and exciting techniques for macromolecular microcrystallography, and lay out a roadmap for optimization of this promising category of experiments.

2. Methods

2.1. Protein expression and purification

Wild-type human cyclophilin A (CypA) was expressed and purified as described previously (Fraser *et al.*, 2009). Briefly, following purification the protein was stored in a solution consisting of 20 m*M* HEPES pH 7.5, 20 m*M* NaCl, 0.5 m*M* TCEP at 4°C until use. Finally, samples were concentrated using Amicon centrifugal filters and then crystallized as described in Section 2.2.

2.2. Crystal formation and optimization

For exploration of the crystallization phase space of CypA, crystallization trays were set up as follows. Well solutions consisting of 100 m*M* HEPES pH 7.5, 5 m*M* TCEP and PEG 3350 (at a varying concentration) were distributed into Greiner 96-well Imp@ct microbatch crystallization plates. Each well contained 2 µl of the respective well solution mixed with 2 µl protein solution at the respective concentration. These drops were then vapor-sealed using 12 µl paraffin oil. For the large-scale production of crystals in batch, 600 µl of protein solution at 60 mg ml^{-1} was combined with 400 µl 50% PEG 3350 in a glass vial and stirred with a stir bar at a constant rate (varying rotational speed). Crystallization was robust over a temperature range spanning 20–25°C.

2.3. Crystal analysis

Raw images of microcrystal slurries were taken using a Nikon Ti microscope in differential interference contrast mode with a Nikon DS-Qi2 camera. The data were interpreted using the *Fiji* software (Rueden *et al.*, 2017). In addition to imaging crystalline slurries, particle densities (crystals per millitre) were analyzed using an INCYTO C-Chip. Diffraction tests were carried out on Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-2 using a 20 × 40 µm beam at 0.9795 Å wavelength. Crystals were diluted and loaded onto a MiTeGen MicroMesh 700/25 loop. Frames were collected for 1.0 s with a 1.0° oscillation. The angular extent of diffraction and the unit-cell parameters were assessed using *Adxv*.

2.4. Sample preparation for serial X-ray experiments

Crystals were formed in batch, as described in Section 2.2, at a constant stirring rate of 500 rev min^{-1}. Further preparation was determined by the delivery method. When using the MESH injector (Sierra *et al.*, 2012), the microcrystal slurry was delivered as is in a Hamilton syringe. When using the LCP injector (Weierstall *et al.*, 2014), the crystals were mixed with a viscogen: either polyethylene oxide (PEO), lipidic cubic phase (LCP) or cellulose. For PEO mixtures, the microcrystal slurry was combined with a viscogen consisting of 10% PEG and 10% PEO, and various ratios of crystal slurry to viscogen were tested. For LCP mixtures, the crystal slurries were centrifuged and the supernatant was removed, with a minimal volume (100 µl) added back to suspend the crystals. The crystals were then mixed with monoolein (9.9 MAG) in a 1:1.5 mass-to-mass ratio using coupled glass syringes (Ishchenko *et al.*, 2016). For

cellulose mixtures, crystal slurries were centrifuged, the supernatant was removed and the crystals were directly mixed with 20% hydroxyethyl cellulose in a 1:9 crystal-to-cellulose ratio as described previously (Sugahara *et al.*, 2017).

2.5. Serial X-ray data collection and analysis

For the MESH and LCP XFEL data sets, we collected data at LCLS-MFX (Sierra *et al.*, 2019) on an MX170-HS Rayonix detector in 2 × 2 binning mode. Crystals were delivered to the X-ray interaction point using either a MESH injector (Sierra *et al.*, 2012, 2016) or an LCP injector (Weierstall *et al.*, 2014). Data were collected using a 3 μm beam at 9.5 keV energy, pulsed at 10 Hz, with a pulse duration of 40 fs on average. Powder diffraction patterns of silver(I) behenate were used to estimate the detector distance. The *cctbx.xfel* GUI was used for real-time feedback on the hit rate and indexing rate, as well as to submit processing jobs onsite. Data were indexed and integrated using *dials.stills_process* (Winter *et al.*, 2018). Initial indexing results were used to refine the detector model, as well as crystal models (Brewster *et al.*, 2018). Refinement of the detector distance and panel geometry improved the agreement between measured and predicted spots. The data were then merged and post-refined using *cxi.merge* (Sauter, 2015). Error estimates were treated according to the Ev11 method (Brewster *et al.*, 2018; Evans, 2011), in which error estimates were increased using terms refined from the measurements until they could better explain the uncertainty observed in the merged reflection intensities. For the cellulose XFEL data set, we collected data at SACLA (Ishikawa *et al.*, 2012) using the Diverse Application Platform for Hard X-ray Diffraction in SACLA (DAPHNIS; Tono *et al.*, 2015) on BL2 (Tono *et al.*, 2019). Diffraction images were collected using a custom-built 4M pixel detector with multi-port CCD (mpCCD) sensors (Kameshima *et al.*, 2014). Data collection was supported by a real-time data-processing pipeline (Nakane *et al.*, 2016) developed in *Cheetah* (Barty *et al.*, 2014). Identified hit images were processed in *CrystFEL* version 0.6.3 (White *et al.*, 2016). Diffraction spots were indexed by *DirAx* (Duisenberg, 1992). Intensities were merged by Monte Carlo integration with the *process_hkl* command in the *CrystFEL* suite with linear scale factors and per-image resolution cutoff. We note that the data collected at SACLA could not be processed using *dials.stills_process* owing to spot-shape irregularities that are an artifact of the mpCCD detector. Specifically, when strong X-ray signals are recorded on this detector, large streaks are produced across horizontal rows of pixels. Serial crystallography data analysis at SACLA relies on the *CrystFEL* pipeline, which has been optimized to handle the idiosyncrasies of the hardware at this facility. The raw data are publicly available at CXIDB.

2.6. MicroED sample preparation

Samples for MicroED were prepared as described previously (Martynowycz *et al.*, 2019a). A 2 μl aliquot of crystals from the batch solution was applied onto a glow-discharged Quantifoil Cu 200 mesh R2/2 holey carbon grid.

The grid was gently blotted from the back (by hand) in an FEI Vitrobot for 10 s at 100% humidity and then vitrified in liquid ethane. Grids were stored in liquid nitrogen until further use, and all additional sample-manipulation and data-collection procedures were performed under cryogenic conditions. Prior to data collection, the grids were clipped and loaded into an FEI Versa FIB/SEM at liquid-nitrogen temperature and milled as described previously (Martynowycz *et al.*, 2019a). The grids were coated with a thin layer of amorphous platinum to increase the contrast during FIB/SEM imaging (Martynowycz *et al.*, 2019b). Large crystals (10–50 μm) near the center of the grid square were identified using a 2 kV SEM. The crystals were milled using a 30 kV gallium ion beam with a stepwise decreasing beam current as the sample slowly approached its final thickness of approximately 200 nm. The final 10 nm on either side of the crystalline lamellae were milled at 10 pA to polish the crystalline surface.

2.7. MicroED data collection and analysis

Grids containing milled crystals were transferred into an FEI Arctica TEM operating at an accelerating voltage of 200 kV under liquid nitrogen. Crystalline lamellae were initially identified in an all-grid atlas taken at 155× magnification, in which crystals were apparent as semitransparent areas suspended over a sharp, straight strip of empty area created by the milling process. Continuous-rotation MicroED data were collected in diffraction mode over an angular wedge between −60° and 0° from the untilted orientation at a rotation rate of 0.3° s⁻¹. The camera length was set to 2055 mm and frames were read out every 2 s. Data were recorded on a CetaD detector operating in rolling-shutter mode with 2 × 2 binning. The camera length was calibrated using a molybdenum foil. MicroED data were converted from the FEI SER format to SMV for data analysis using in-house software that is freely available (https://cryoem.ucla.edu/). The data were indexed and integrated with *XDS* and scaled in *XSCALE* (Kabsch, 2010). The raw data are publicly available at the SBGrid Data Bank.

2.8. Model refinement and analysis

The data were reduced as described in Sections 2.5 and 2.7. Initial phases were calculated by molecular replacement using *Phaser* (McCoy *et al.*, 2007), with PDB entry 4yum (Keedy *et al.*, 2015) as the search model. Prior to initial atomic refinement, R_{free} flags were carried over from PDB entry 4yum and random displacements ($\sigma = 0.5$ Å) were applied to the atomic coordinates to help remove model bias. Next, iterative cycles of model building and further refinement were performed until the models reached convergence. Individual atomic coordinates, atomic displacement parameters (*B* factors) and occupancies were refined using *phenix.refine* (Afonine *et al.*, 2012; Liebschner *et al.*, 2019). Automatic identification of ordered solvent was performed during the early cycles of model refinement. Models and maps were visualized and rebuilding steps were performed using *Coot* (Emsley *et al.*, 2010). The final structural models were visualized using *PyMOL* (Schrödinger) and were also used for ensemble

refinement using *phenix.ensemble_refine* (Burnley *et al.*, 2012). The input parameters for ensemble refinement (`pTLS`, `tx` and `wxray_coupled_tbath_oset`) were optimized for each data set.

3. Results

3.1. Optimization of batch crystallization

Large CypA crystals (Supplementary Fig. S1), of the order of hundreds of micrometres, or even millimetres, are readily obtained by vapor-diffusion methods (Fraser *et al.*, 2009; Keedy *et al.*, 2015); however, these crystals are too large for either microfluidic serial crystallography or MicroED. Therefore, we sought to optimize the production of microcrystals in batch format so that they could be easily delivered to the X-ray beam for serial crystallographic measurements using microfluidic sample injectors. As a first step towards this goal, we systematically explored the phase space of CypA crystallization in the vicinity of the conditions that yield large crystals [protein concentration in the range 80–100 mg ml^{-1} with 20–25%(*w/v*) PEG 3350 as a precipitant and HEPES buffer pH 7.5]. We adapted the established CypA crystallization protocol (Fraser *et al.*, 2009) to a microbatch format (rather than vapor diffusion) and tested an array of conditions by varying the protein and precipitant concentrations across the two axes of a 96-well crystallization plate. The lowest concentration of protein and precipitant led to the formation of large crystals that were ideal for data collection under traditional rotation conditions. In microbatch format, conditions that resulted in

large, single crystals contained substantially lower protein and precipitant concentrations relative to vapor-diffusion experiments that yield similarly sized crystals. Increasing the protein concentration led to the formation of a greater number of smaller crystals; however, they tended to cluster together and displayed a needle-like morphology (Fig. 1). High protein concentrations also led to a large variation in crystal size, which we sought to avoid since crystal monodispersity is desirable for serial crystallography experiments. In addition to modulating protein concentration, the precipitant concentration was also varied. Increasing the concentration of the precipitant led to increased crystal density while maintaining better monodispersity. At the highest precipitant concentrations that we tested the protein tended to aggregate, rather than crystallizing. Given these characteristics, we found that we could consistently obtain dense crystal slurries when the final PEG 3350 concentration was near 20% (Fig. 1). Increasing the protein concentration beyond 35 mg ml^{-1} did not lead to appreciable increases in crystal density, so we chose a final protein concentration of 35 mg ml^{-1} and a final PEG 3350 concentration of 20% for further optimization.

After identifying ideal protein and precipitant concentrations using the microbatch method described in Section 2.2, our next goal was to scale up the microbatch procedure to produce crystal slurries on the millilitre scale. We developed a batch crystallization protocol in which 0.9 ml CypA solution is stirred using a magnetic stir bar inside a glass vial and 0.6 ml PEG 3350 solution [50%(*w/v*)] is added dropwise to produce a solution with final prot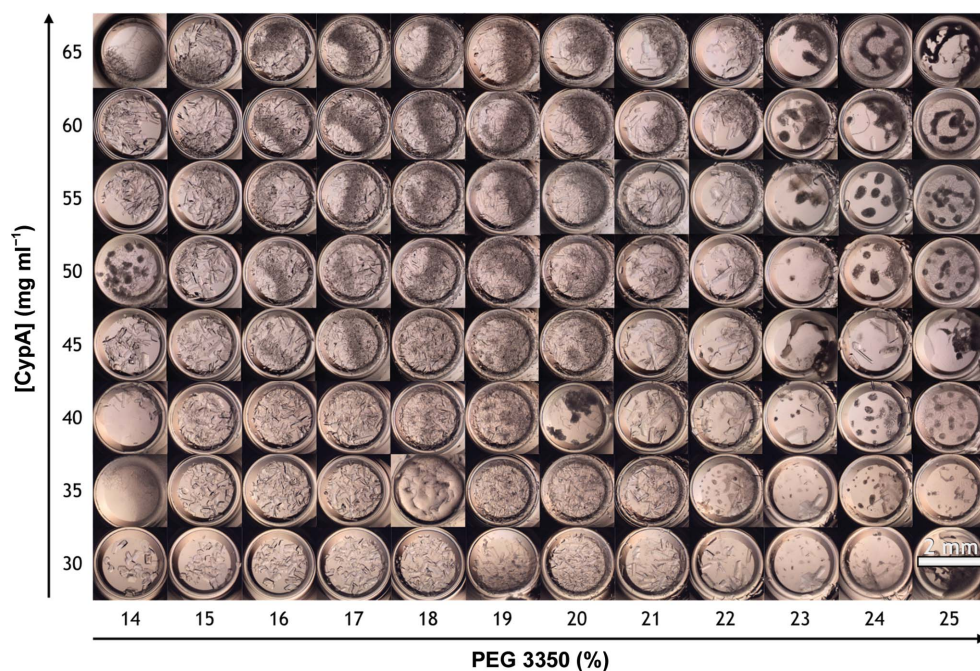ein and precipitant concentrations of 36 mg ml^{-1} and 20%(*w/v*), respectively. Initially, we used a rotating mixer to mix the slurry by inversion, but adding a stir bar yielded better results. Our batch-stirring protocol also improved the monodispersity, decreased the crystal size and increased the crystal density (crystals per millitre) relative to the micro-batch method. Increasing the final protein concentration above 36 mg ml^{-1} and increasing the final PEG 3350 concentration above 20% did not further improve the monodispersity, size or crystal density; however, we discovered that modulating the stirring speed of the crystallization solution allowed us to control the formation of differently sized crystals (Fig. 2). Within the range of stir rates that we tested (200–800 rev min^{-1}), we observed an increase in crystal density as the stir rate increased, which was coupled to a decrease in the average crystal size.

Figure 1
An array of images that illustrates the crystallization phase space of CypA. Concentrated solutions of CypA were mixed in a 1:1 volume ratio with solutions of PEG 3350 at varying concentrations. The labels on the axes indicate the final concentrations after mixing. CypA crystallizes readily in PEG 3350 solutions; however, the crystal size and morphology varies dramatically as a function of protein and PEG concentration. Specifically, at low CypA and low PEG 3350 concentrations (bottom left corner) the crystals that form are few and large, while at high PEG 3350 concentrations (right side) CypA aggregates and no crystals form. In the middle of the phase space, dense solutions of small crystals form.

Crystals of 50 × 50 × 50 μm or larger developed at lower stir rates, while crystals tended towards 15 × 15 × 15 μm at higher stir rates. In addition to modulating the density of the slurry, the stirring rate also affected the monodispersity (Ibrahim *et al.*, 2015). At 200 rev min^{-1} we observed greater variation in the size of the crystals, while at higher stir rates the crystals were more monodisperse but tended to clump together.

Next, optimized microcrystals were tested for their diffraction quality using a synchrotron beamline (SSRL 12-2) and exhibited Bragg peaks that were visible beyond 2.0 Å resolution (Supplementary Fig. S2). The indexed unit-cell parameters matched those of previously published CypA structures (Keedy *et al.*, 2015), indicating that the batch crystallization method does not have an adverse effect on the quality of the crystal lattice. The crystals used for this test were 50 × 50 × 50 μm in size, while typical CypA crystals used for single-crystal X-ray crystallographic structure determination are over 100 μm in each of their three dimensions.

3.2. Serial X-ray crystallography: microfluidic sample delivery and data collection

We used CypA microcrystal slurries to perform several different types of serial crystallography experiments, all conducted in an ambient atmosphere, in order to assess the performance of different methods of microfluidic sample delivery and to determine whether the conditions created in the injector alter the outcome of the structural measurements. Specifically, we tested two different types of microfluidic sample injectors: one which utilizes the principle of electro-spinning to form a microfluidic stream (Sierra *et al.*, 2012) and another that performs high-pressure microextrusion of crystals embedded in a viscous material (Weierstall *et al.*, 2014). Because the microextrusion method requires that samples are extremely viscous, we tested three different viscogens as additives to our CypA samples. Thus, a total of four unique experimental conditions were explored.

The first injector system that we implemented, in an experiment conducted at the MFX endstation of the XFEL at the Linac Coherent Light Source (LCLS), is referred to as the microfluidic electrokinetic sample holder (MESH). This device relies upon the principle of electrospinning to break the surface tension of the crystal slurry and drive it into a microjet as it exits the tip of a capillary (Sierra *et al.*, 2012). Within the MESH system, gentle pressure (less than 20 psi; 1 psi = 6.9 kPa) from a syringe pump drives crystals through a capillary (250 μm internal diameter) towards the X-ray interaction point, and the application of an electrostatic potential of approximately 3000 V (Table 1) across the sample stretches the liquid into a thin jet as it emerges from the capillary tip. Although our crystals were much smaller than 250 μm, they had a tendency to cluster together (Fig. 2), so we used a
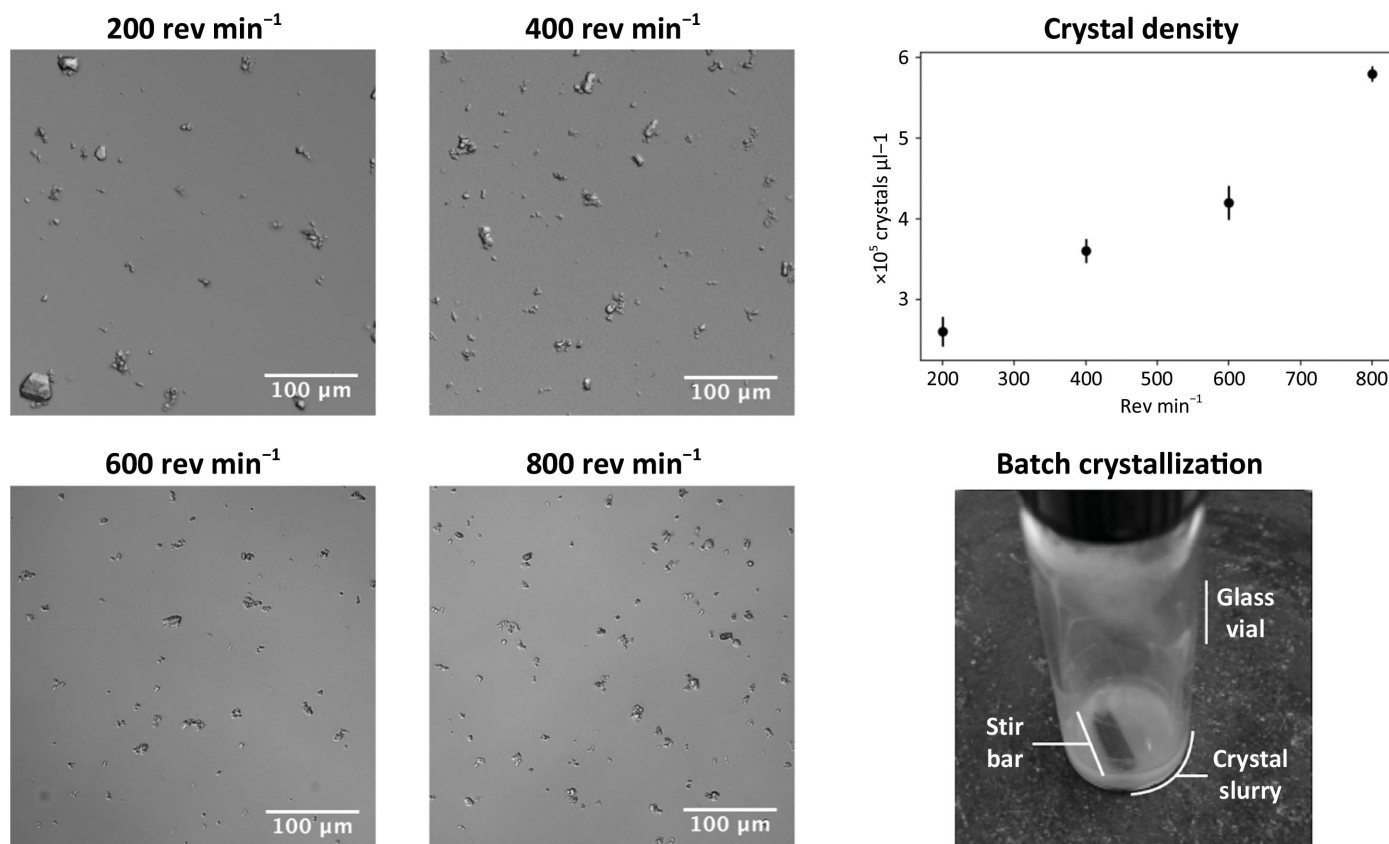

Figure 2
Images of microcrystals formed in batch with constant stirring. As the stir rate increased, the average size of the crystals decreased and the density of the slurry increased. This was confirmed by assessing the crystal density using a hemocytometer.

Table 1
Sample-injection parameters for serial XFEL experiments.

	MESH injector	LCP	Cellulose
Sample flow rate (μl min^{-1})	0.3	0.088	0.625
Capillary diameter (μm)	250	50	75
Linear jet velocity (mm s^{-1})	Not determined	0.75	9.43
Electric field (V cm^{-1})	~3000	0	0
Pressure (psi)	<20	~8500	~3200
Delivery matrix	As crystallized	LCP-like lipid mixture	20% hydroxyethyl cellulose

Table 2
Crystallographic statistics for data collection.

Values in parentheses are for the highest resolution shell.

	XFEL MESH	XFEL LCP	XFEL cellulose
X-ray source	MFX, LCLS	MFX, LCLS	BL2, SACLA
Photon energy (keV)	9.4	9.5	10
X-ray pulse duration (fs)	~50	~45	~10
Photons per pulse	~1×10^{12}	~2.5×10^{12}	~1×10^{11}
Repetition rate (Hz)	10	10	30
Hit rate (%)	19	20	20
Indexing rate (%)	63	30	75
No. of images	18358	11821	23947
Resolution range	20.0–1.62 (1.65–1.62)	20.0–1.65 (1.68–1.65)	20.0–1.56 (1.58–1.56)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters			
a (Å)	43.32 ± 0.11	43.10 ± 0.18	43.00 ± 0.26
b (Å)	52.94 ± 0.09	52.65 ± 0.13	52.60 ± 0.23
c (Å)	89.87 ± 0.21	89.29 ± 0.26	89.20 ± 0.37
$\alpha = \beta = \gamma$ (°)	90	90	90
Total reflections	5928801 (37371)	2477563 (11714)	7874934 (59803)
Multiplicity	224.08 (28.81)	98.90 (9.63)	261 (40.5)
Completeness (%)	99.95 (100)	99.89 (99.92)	100 (100)
Mean $I/\sigma(I)$	5.312 (0.747)	4.062 (1.076)	7.91 (1.46)
Wilson B factor (Å^2)	20.21	19.84	24.82
R_{split} (%)	8.3 (28.4)	14.9 (54.3)	7.94 (66.0)
CC_{int} (%)	99.3 (86.5)	94.5 (56.7)	99.3 (58.6)
Software	*cctbx + cxi.merge*	*cctbx + cxi.merge*	*CrystFEL*

capillary with a relatively large internal diameter to avoid clogging. A full description of the optimal operating parameters for our experiment using the MESH is provided in Table 1. During this experiment, the position and physical dimensions of the Taylor cone and microfluidic jet formed by electrospinning (Supplementary Fig. S3) had a tendency to fluctuate. However, by positioning the injector so that the X-ray beam was at the approximate position where the liquid within the Taylor cone accelerated and became the jet, we were able to collect data at a hit rate of approximately 19% and an indexing rate of 63% (Table 2). Moving the injector so that the beam pointed at the jet itself resulted in an unacceptably low hit rate (<1%), and moving the injector so that the beam was positioned at a more stable, but thicker, region of the Taylor cone resulted in extremely high background scattering owing to the excessive volume of solvent in the beam path.

The second type of microfluidic sample injector that we implemented for our experiments was a viscous extrusion system. Several variations on this device have been created, all of which operate using high pressure to extrude a viscous crystal slurry through a capillary, which is stabilized by a sheath gas to form a relatively slow-moving column of material as it exits the capillary at the X-ray interaction point

(Weierstall *et al.*, 2014). Within the injector, crystals were exposed to pressures as high as 8500 psi (Table 1) with no electrostatic potential. Given the high operating pressures of these injectors, we experienced no issues with clogging, so we used a 50 μm capillary to minimize sample consumption. Because these injector systems require that the sample be much more viscous than our CypA crystal slurries, their use required that we add viscogens, or 'carrier media', to our samples after crystallization but prior to injection. We experimented with three different types of carrier media: lipidic cubic phase (LCP) formed from monoolein, hydroxyethyl cellulose and polyethylene oxide (PEO). Carrier-media solutions were prepared independently and were mixed with crystal slurries to embed the microcrystals in the viscous material immediately prior to loading samples into the injector reservoirs. For LCP and PEO, the carrier media and crystal slurries were prepared in separate glass syringes and mixed using a coupling device, while samples containing cellulose were prepared by mixing crystal slurry and carrier medium on a glass surface using a spatula (Sugahara *et al.*, 2015). Visual inspections using a microscope equipped with a cross-polarizer confirmed that while it seems harsh, the process of mixing CypA microcrystals into viscous material does not visibly damage them. Crystal slurries prepared with LCP and PEO carrier media were delivered to the XFEL interaction point of the MFX endstation at the LCLS using an injector device developed by Weierstall *et al.* (2014). Crystal slurries prepared with hydroxyethyl cellulose were delivered to the XFEL interaction point of the SPring-8 Angstrom Compact Linear Accelerator (SACLA) using an injector setup similar to that used in studies on photosystem II (Suga *et al.*, 2017; Nango *et al.*, 2016; Kubo *et al.*, 2017). We observed that samples prepared with both LCP and hydroxyethyl cellulose formed microfluidic jets that were highly stable and maintained consistent physical dimensions for long periods of time, allowing efficient data collection. We obtained an average hit rate of 20% and an indexing rate of 30% for our experiment with LCP as the carrier medium, and we obtained an average hit rate of 20% and an indexing rate of 75% for our experiment with cellulose as the carrier medium (Table 2). On the other hand, and in contrast to reports by Martin-Garcia *et al.* (2017), we were unable to obtain useful data when PEO was used as the carrier medium. Samples prepared using PEO did not form stable jets, but instead formed droplets at the tip of the injector nozzle, which grew to a critical mass and then

Table 3
Statistics for X-ray model refinement.

Values in parentheses are for the highest resolution shell.

	XFEL MESH	XFEL LCP	XFEL cellulose
Resolution range (Å)	19.91–1.62	17.43–1.65	19.9–1.56
	(1.72–1.62)	(1.75–1.65)	(1.64–1.56)
Unique reflections	26445 (4324)	25034 (4076)	29528 (4159)
Reflections used in refinement	25613 (4189)	24247 (3950)	28598 (4027)
Reflections used for R_{free}	832 (135)	787 (126)	929 (132)
R_{work}	0.1362 (0.2398)	0.1434 (0.2316)	0.1348 (0.2376)
R_{free}	0.1569 (0.2455)	0.1671 (0.2367)	0.1509 (0.2804)
No. of non-H atoms			
Total	1522	1558	1551
Macromolecules	1383	1399	1401
Protein residues	163	163	163
R.m.s.d., bonds (Å)	0.005	0.004	0.015
R.m.s.d., angles (°)	0.844	0.656	1.298
Ramachandran favored (%)	96.89	96.89	96.89
Ramachandran allowed (%)	3.11	3.11	3.11
Ramachandran outliers (%)	0.00	0.00	0.00
Rotamer outliers (%)	0.68	2.68	0.00
Clashscore	3.24	1.43	1.78
Average B factor (Å^2)			
Overall	26.29	26.78	29.62
Macromolecules	24.78	25.29	28.02
PDB code	6u5c	6u5d	6u5e
Ensemble refinement			
R_{work}	0.1241	0.1304	0.1296
R_{free}	0.1477	0.1517	0.1512

Table 4
All-atom r.m.s.d. values for comparison of the three serial crystallography structures.

Pair	R.m.s.d. (Å)
MESH/LCP	0.048
MESH/cellulose	0.063
LCP/cellulose	0.069

slowly dripped towards the X-ray interaction point and became unstable in the sheath gas. This problem persisted despite considerable effort to optimize the sample and sheath-gas flow parameters and attempts to use both helium and nitrogen as the sheath gas.

In addition to assessing the data quality under different sample-delivery conditions, we also wanted to determine whether the different types of sample-delivery methods bias the orientation of the crystals as they are delivered to the X-ray beam (Supplementary Fig. S4). Because serial crystallography methods assume that crystals are delivered to the beam in random orientations in order to sample all of reciprocal space, the extent to which this is not true limits the efficiency of the experiment. We found that for the MESH data set the crystals do not appear to have an orientation bias as they are delivered to the X-ray beam, while the data set collected using LCP as a carrier medium for a viscous extrusion injector showed significant orientation bias.

3.3. Serial X-ray crystallography: data quality and atomic structure are robust across sample-delivery strategies

Following data collection, the individual data sets were processed and the reduced data were compared, revealing that the high-quality diffraction typical of CypA crystals is consistent across the different sample-delivery methods that we implemented in our serial crystallography experiments. Raw diffraction images collected at LCLS (MESH and LCP conditions) were indexed and integrated using *dials.stills_process* (Winter *et al.*, 2018), and the individual measurements were merged (with post-refinement) using *cxi.merge* (Sauter,

2015). Raw diffraction images collected at SACLA (cellulose condition) were indexed, integrated and merged using *CrystFEL*, following hitfinding with *Cheetah*. We used different software to process data obtained at different XFEL light sources because we observed that the optimal software performance generally depends on idiosyncratic features of the experimental endstations, such as detector behavior and spectral characteristics of the X-ray pulses. The data sets comprise 18 358, 11 821 and 23 947 indexed diffraction images for the MESH, LCP and cellulose samples, respectively. The large number of indexed patterns used to construct each data set resulted in very high completeness and multiplicity, and the CypA crystals diffracted to high resolution under all three delivery strategies (Table 2). The diffraction resolutions of the data sets reported here fall within the range of resolutions reported for CypA structures solved using large crystals and rotation geometry, and the modest differences in maximum resolution (1.65–1.56 Å) between the data sets are likely to be owing to differences in the numbers of indexed patterns contributing to each data set, rather than to significant differences in the quality of diffraction under the three distinct sample-delivery conditions. Statistical metrics reflecting the measurement precision and strength of the diffraction signal were favorable for all data sets, indicating that none of the delivery methods compromised the integrity of the crystal lattice (Table 2).

Using the reduced data sets, we performed molecular replacement to calculate initial phases, followed by iterative cycles of model building and atomic refinement to determine the structure of CypA under each of the different sample-delivery conditions. After molecular replacement and before the initial cycle of manual model building, we applied random perturbations ($\sigma = 0.5$ Å) to the atomic coordinates and refined them against the X-ray data in order to eliminate any effect of model bias that might arise from using the same molecular-replacement search model for the three independent structures. We performed iterative rounds of model building and atomic refinement until the procedure reached convergence, and found that the models that we obtained from each of the three experiments were of comparable statistical quality in terms of their fit to the experimental data and their overall geometry (Table 3). We discovered that the method of sample delivery in each of the three serial crystallography experiments has a minimal impact on the average structure of CypA; however, the three individual structures are not identical.

Pairwise alignment of the three structures and comparison of atomic coordinates revealed that for each pair of structures the root-mean-squared deviation (r.m.s.d.) of atomic positions

is less than 0.1 Å (Table 4). The conformational heterogeneity of a key network of residues in CypA that extends from the core of the protein to the active site has been studied previously using ambient-temperature crystallographic experiments. Rotameric exchange of residues in this network, which include Arg55 (the catalytic residue), Met61, Ser99 and Phe113, is required for enzymatic turnover (Eisenmesser *et al.*, 2005; Fraser *et al.*, 2009). Notably, all three of our serial data sets revealed evidence for multiple conformations of these

residues (Fig. 3). Small differences between the structures existed as differences in rotamers (or mixtures of rotamers) for side chains with generally weak electron density, such as Met61.

Furthermore, the structure of CypA did not appear to be significantly perturbed by either the electric field within the MESH jet or by the high pressures within the LCP injector used for the LCP and cellulose data sets (Fig. 3). Local features within the models matched the maps well, with only
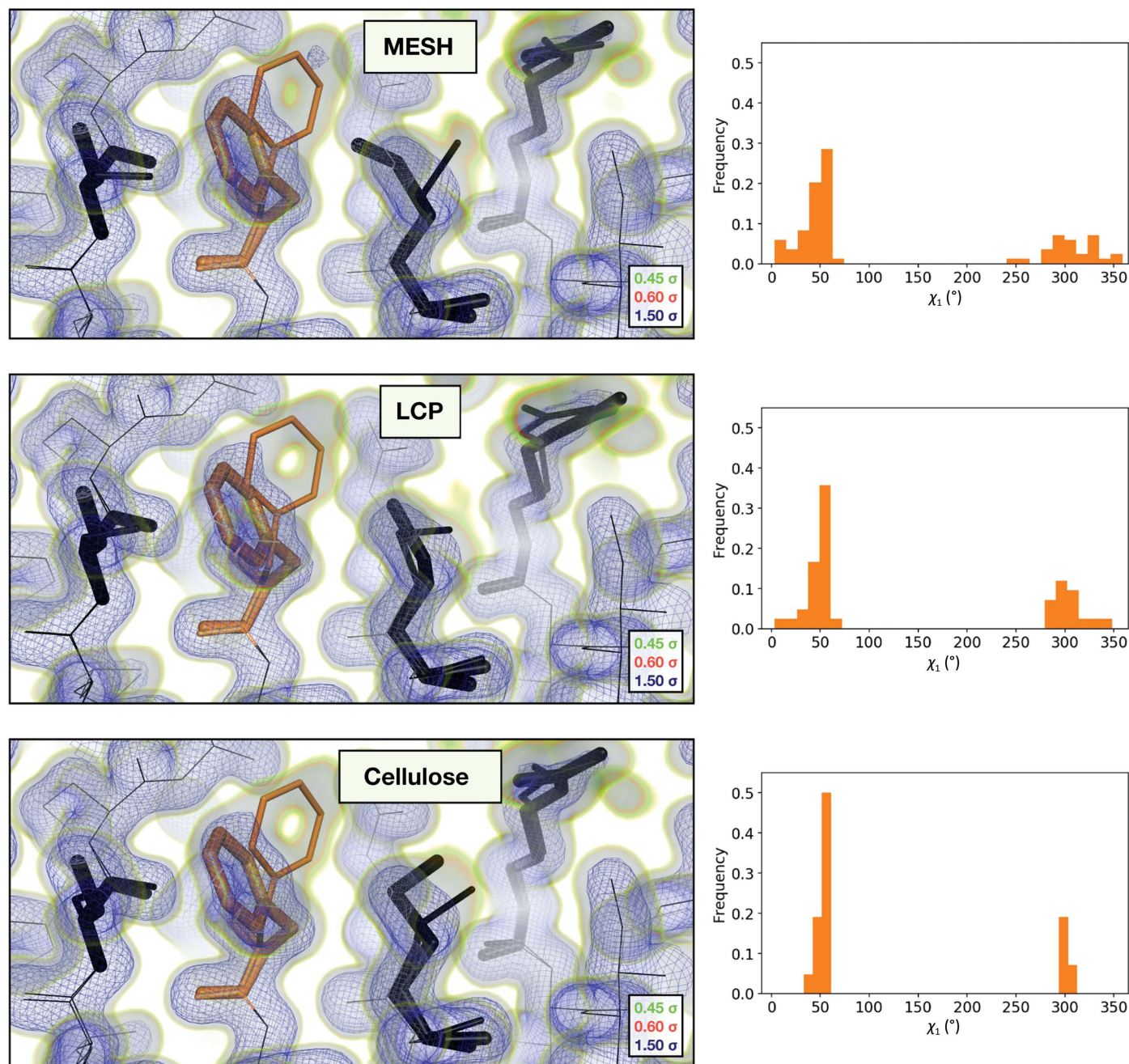
Figure 3
Comparison of the $2mF_o - F_c$ maps and the refined multi-conformer models produced from each serial XFEL experiment. Maps were visualized at multiple contour levels to show evidence of alternative conformations. Following multi-conformer refinement, ensembles were generated from each model using *phenix.ensemble_refine*. In the right panel, a histogram of the χ_1 angles for residue 113 is plotted for the ensemble. Multi-conformer models plus maps, and the distribution of χ_1 angles across the ensemble models, are similar for all three XFEL data sets.

subtle differences noticeable in the maps. The model statistics were similar across all three data sets; only the average B factor differed appreciably between the data sets (Table 3). Comparing the normalized atomic B factors of atoms within each structure (Supplementary Fig. S4) revealed that the increase in the average was not owing to any localized change in conformational heterogeneity, but instead resulted from a global increase in the refined B factors. These global differences in atomic B factors across structures could be owing to varied perturbation of the crystal lattice (but not the molecular structure) that results from exposure to different sample-delivery conditions or owing to small differences in the data-processing parameters. When the structure was expanded to an ensemble, the resulting multiconformer models settled into nearly equivalent minima, confirming the similarity of the three data sets (Supplementary Figs. S5 and S6).

3.4. MicroED: grid preparation and data collection

In order to obtain CypA microcrystals on copper grids that were suitable for MicroED data collection, we tested several sample-preparation strategies. The ideal crystal thickness for MicroED samples is approximately 300–500 nm (Martyno-

wycz *et al.*, 2019*b*), which is smaller than any crystal that is visible using light microscopy. Firstly, we prepared grids using a CypA microcrystal slurry containing visible crystals of the order of 10 µm in size, with the hope that this sample would also contain much smaller crystal fragments that would be acceptable for data collection. We examined this sample in the microscope and observed only large (several micrometres or larger) microcrystals on the grid. As a next step, we attempted to reduce the size of the CypA microcrystals using several physical agitation methods, including vortexing and crushing the crystals using either a pipette tip or glass beads. Samples exposed to physical agitation were used for grid preparation and were examined under the microscope, again revealing an absence of suitably sized crystals for data collection. Attempts to improve the grid preparation by changing the glow-discharge and blotting methods also did not result in suitable samples. We hypothesize that the difficulties in preparing grids with submicrometre-sized CypA crystals result from a combination of the surface properties of the crystals and the strong lateral forces that are introduced by the blotting process, which could pull small crystals off of the grid. We note that we used only grids with amorphous holey carbon supports, and did not attempt to prepare samples using grids
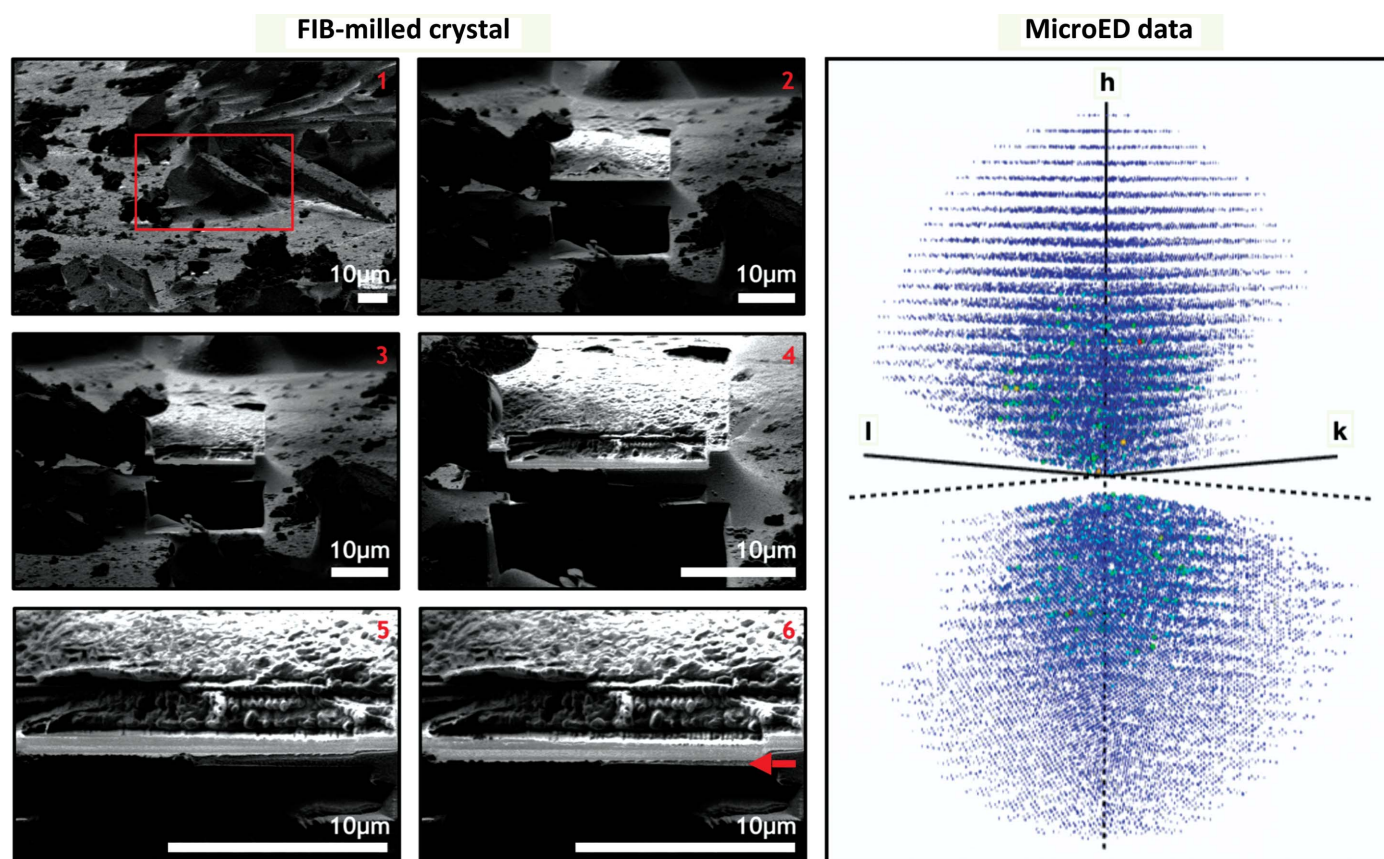

Figure 4
MicroED data collection. CypA crystals were deposited on a copper grid with an amorphous carbon support material and frozen in vitreous ice (left). A single crystal (highlighted in a red box in image 1) is shown at various stages (1–6) of the FIB-milling process. The edge of the final crystalline lamella is denoted with a red arrow in image 6. Also shown is the intensity-weighted reciprocal lattice (right) representing the MicroED data that were collected from the single crystal shown in the left panel.

Table 5
Crystallographic statistics for MicroED data.

Values in parentheses are for the highest resolution shell.

No. of images	135
Resolution range (Å)	30.5–2.50 (2.57–2.50)
Space group	$P2_12_12_1$
Unit-cell parameters	
a (Å)	42.40
b (Å)	53.40
c (Å)	87.76
$\alpha = \beta = \gamma$ (°)	90
Total reflections	22370 (1668)
Multiplicity	1.95 (1.94)
Completeness (%)	86.0 (87.3)
Mean $I/\sigma(I)$	3.23 (1.01)
Wilson B factor (Å^2)	35.53
R_{meas} (%)	24.9 (87.7)
$CC_{1/2}$ (%)	95.2 (44.8)
Unique reflections	6236 (608)
Reflections used in refinement	6236 (608)
Reflections used for R_{free}	213 (22)
R_{work}	0.1854
R_{free}	0.2237
No. of non-H atoms	
Total	1280
Macromolecule	1248
Protein residues	163
R.m.s.d., bonds (Å)	0.011
R.m.s.d., angles (°)	0.754
Ramachandran favored (%)	96.27
Ramachandran allowed (%)	3.73
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.76
Clashscore	0.41
Average B factor (Å^2)	
Overall	30.60
Macromolecule	30.97
PDB entry	6u5g
Ensemble refinement	
R_{work}	0.2351
R_{free}	0.2587

with more exotic support materials such as gold or graphene oxide.

Because we were unsuccessful in preparing samples for MicroED using traditional methods of applying crystals to grids, we turned to a method that utilizes a focused ion beam (FIB) to mill larger crystals down to an appropriate thickness. We observed that crystals larger than several micrometres were able to stick to the holey carbon grids, and we prepared a frozen grid with crystals that were approximately $5 \times 5 \times 5$ to $20 \times 20 \times 20$ μm in size. Prior to MicroED data collection in the transmission electron microscope (TEM), the frozen samples were loaded into a scanning electron microscope (SEM) equipped with a gallium ion FIB. A single crystal of approximately $20 \times 20 \times 20$ μm in size was identified in the SEM and was subsequently FIB-milled to form a crystalline lamella that was approximately 200 nm thick (Fig. 4). The grid containing the lamella was then transferred to the TEM for MicroED data collection. Three separate data sets were collected from unique regions of the lamella, two of which were merged to produce the final reduced data set that was used for structure determination. Inclusion of the third data set degraded the quality of the merged data. Also, because the rotation range of the microscope stage is restricted, and the crystal orientation was the same for all three data sets owing to

the fact that they were collected from a single lamella, the final merged data set suffered from a missing wedge of reciprocal space (Fig. 4 and Supplementary Fig. S7) and had an overall completeness of 86%. Bragg peaks were measured out to 1.9 Å resolution, but integrating data beyond the 2.5 Å cutoff that we implemented resulted in poorer data quality. Additional information about the quality of the merged MicroED data is provided in Table 5.

3.5. CypA structure determination from MicroED data

After merging the integrated MicroED data to obtain a high-quality data set, we implemented exactly the same structure-determination procedure as was used for analysis of the serial X-ray diffraction data sets. Specifically, we performed molecular replacement followed by the application of random coordinate perturbations, and then iterative model building and refinement until the R_{work} and R_{free} values converged and no additional improvements to the model could be made. Analysis of the CypA crystal structure determined by MicroED revealed two notable features.

Firstly, during the indexing stage of the data-reduction procedure, we observed that the unit cell had unusual dimensions (Fig. 5). Specifically, while the crystallographic a and c axes match well to those of other CypA structures determined at cryogenic temperatures, the b axis is approximately 1% longer than the corresponding axes in typical CypA structures determined at 'physiological' temperatures (>260 K). Because of indexing challenges resulting from the systematic incompleteness of the data (14% of the predicted observations are missing), such as the inability to observe reciprocal-lattice points along the principal k and l axes within the 60° wedge that we measured (Fig. 4), we took several additional steps to ensure that the unit cell was indexed accurately and that the space-group symmetry was correctly assigned. To address the possibility that an optical distortion in the microscope or challenges related to the flatness of the Ewald sphere (Clabbers & Abrahams, 2018) could lead to the incorrect measurement of unit-cell dimensions, we performed a structure-refinement procedure that also simultaneously refined the coordinates and the lengths of the unit-cell axes (Clabbers et al., 2018) using REFMAC5 (Winn et al., 2011; Murshudov et al., 2011). This procedure resulted in refined unit-cell dimensions that were similar to the input (Supplementary Table S1), without improvement in R_{work} and R_{free}, indicating that the refined unit cell does not result in a model that is more consistent with the measured intensities. To confirm that the elongation of the b axis does not also break the crystallographic $P2_12_12_1$ space-group symmetry, we reduced the raw data three separate times in space group $P2_1$. In each of these three data sets, the twofold symmetry operation was preserved along a different crystallographic axis (i.e. $P2_111$, $P12_11$ and $P112_1$ relative to the parent $P2_12_12_1$). Refinement of the CypA structure against the data with lower symmetry produced models with worse overall quality than when the data were reduced in space group $P2_12_12_1$, confirming the validity of the space-group assignment.

The second notable observation that we made about the crystal structure determined by MicroED is that while the unit cell is distorted relative to other CypA structures, the structure of the molecule within the unit cell is essentially the same as in X-ray structures that were determined using cryocooled crystals (PDB entry 3k0m; Fraser *et al.*, 2009), with an r.m.s.d. of 0.22 Å. Structures of CypA determined from cryocooled crystals, using both X-ray and MicroED, lack key conformations that are visible in their ambient temperature counterparts. In particular, ambient temperature structures of CypA reveal alternative conformations of an important network of amino-acid side chains (the catalytic residue Arg55, as well as Met61, Ser99 and Phe113), while structures determined using cryocooled samples, including the MicroED structure

presented here, reveal only a single conformation of these side chains (Fig. 6). Truncating the resolution of room-temperature X-ray data sets to match the MicroED resolution limit and generating electron-density maps revealed that the loss of the alternate conformation in the maps calculated from MicroED data could not be attributed to the differences in resolution (Supplementary Fig. S8).

3.6. Effect of experimental conditions on unit-cell dimensions

Comparing the structures that we determined using different microcrystallography techniques revealed that the unit-cell dimensions of the CypA crystals were noticeably affected by the conditions required for each of the different
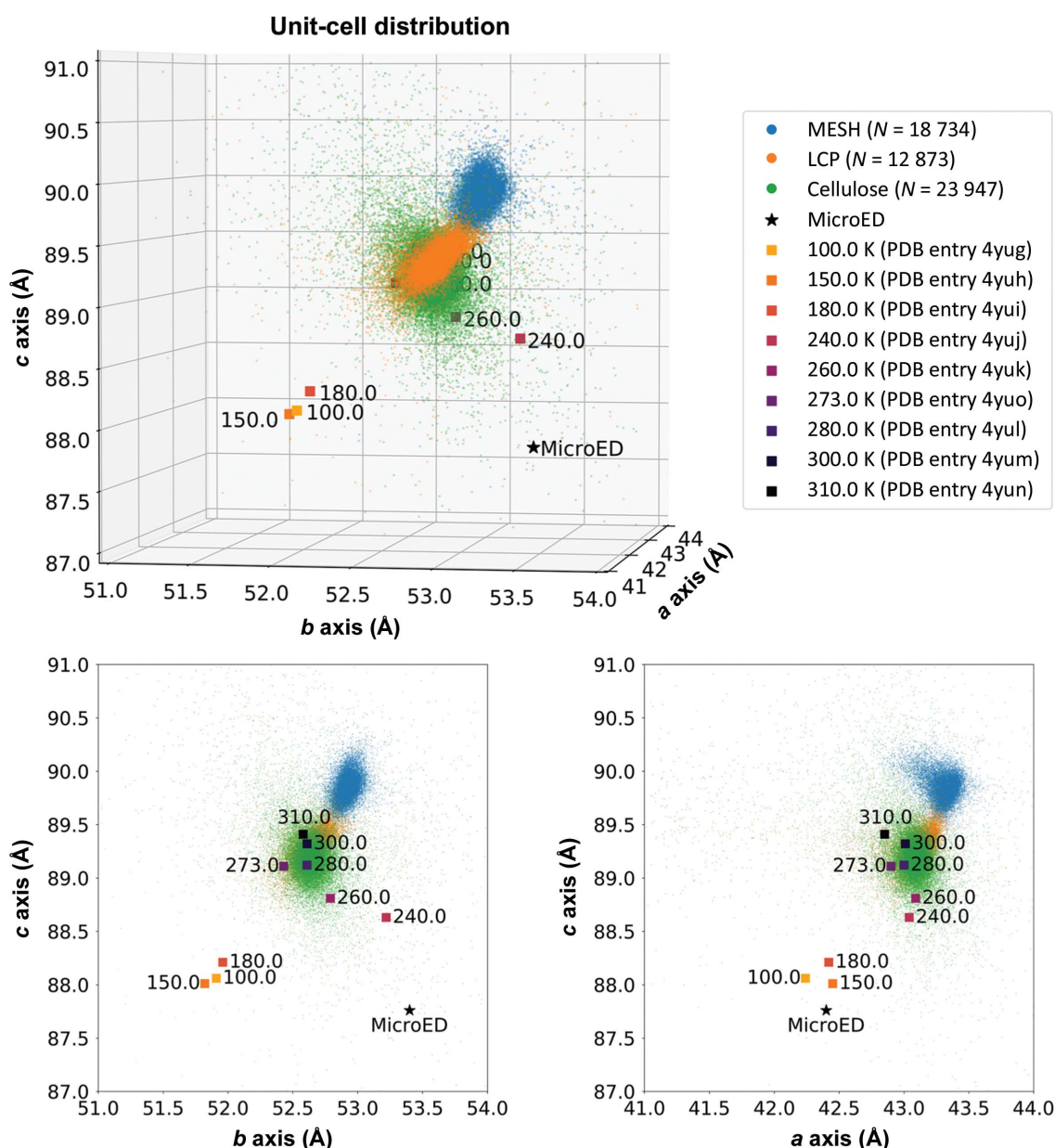

Figure 5
Comparison of unit-cell dimensions across data-collection strategies. Published structures are provided as a reference for the effect of temperature upon unit-cell dimensions. The unit cells measured using serial XFEL experiments resemble data from published room-temperature structures. An FIB-milled crystal used for MicroED revealed dimensions that were unique from the unit-cell compression normally seen in cryogenic X-ray data.
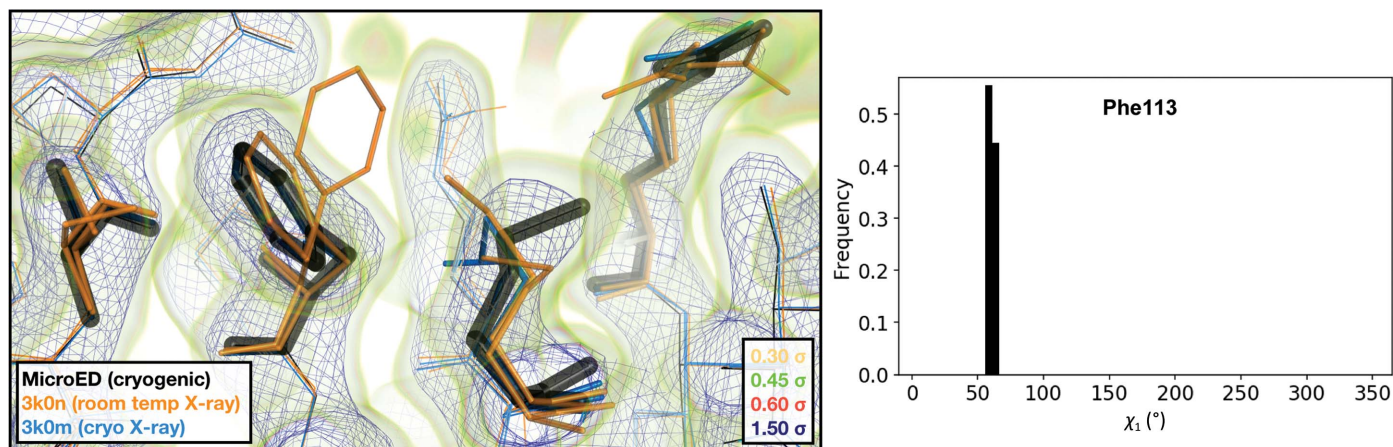
Figure 6
Visualization of the $2mF_o - F_c$ map and the refined model measured from an FIB-milled crystal using MicroED. The conformation of residues coupled to the catalytic site resembles structures previously solved under cryogenic conditions using X-ray crystallography (PDB entry 3k0m). For some regions of the structure, the cryogenic X-ray and MicroED structures are indistinguishable. A previously published multi-conformer model produced from data acquired at room temperature is provided for comparison (PDB entry 3k0n). Following refinement, ensembles were generated using *phenix.ensemble_refine*. In the right panel, a histogram of the χ_1 angles for residue 113 is plotted for the ensemble. All members of the ensemble adopted the same rotameric position as previous cryogenic structures.

experiments (Fig. 5). The three serial X-ray data sets closely resemble previous room-temperature data collected under traditional rotation conditions. Additionally, we observed that for sample-delivery methods that involve embedding the crystals in a viscous carrier medium (LCP or cellulose) the unit cell tended to be slightly smaller than for experiments that do not require such additives (MESH). Average unit-cell dimensions for the LCP and cellulose data sets were approximately 0.5–0.7% smaller than for the MESH data set, and the small differences between these average values were found to be statistically significant using one-way ANOVA and *post hoc* Tukey tests (Supplementary Table S2). We hypothesize that the observed shrinkage of the unit cell could be owing to crystal dehydration. The MicroED data revealed a previously unreported unit cell with an expanded *b* axis relative to any other CypA structure that has been reported (Fig. 5). The unit cell was different from both ambient temperature (PDB entry 3k0n; Fraser *et al.*, 2009) and cryogenic (PDB entry 3k0m) X-ray structures, and matches most closely to a structure determined at 240 K (PDB entry 4yuj). The unusual unit cell observed in the MicroED experiment is not the result of measurement error (Supplementary Table S1), and we hypothesize that it could be the result of cooling the crystals in ethane rather than nitrogen, or could be caused by the grid-blotting procedure or FIB milling, both of which are unique to MicroED. Despite the small variations in crystal packing that cause changes in unit-cell parameters for structures determined using different methods, the refined coordinates of the CypA molecules themselves are consistent.

4. Discussion

The ability to measure diffraction signals from ever-smaller crystal samples has enabled a variety of new and innovative experiments in macromolecular crystallography; however, there is still a relative absence in the literature of practical guidelines for optimizing microcrystallography experiments. The work that we present here attempts to address this knowledge gap by providing a detailed description of how we optimized the growth of cyclophilin A (CypA) microcrystals and measured their diffraction using two emerging microcrystallography techniques: serial XFEL crystallography and microcrystal electron diffraction (MicroED). Our results compare and contrast serial X-ray and MicroED methodologies, and highlight some important considerations and pitfalls that might be encountered during the preparation of microcrystalline samples for the respective experiments. Because we measured MicroED data from only a single crystal, we focus our comparison on technical issues related to sample preparation and ease of data collection. Consequently, this case study provides a roadmap for experimenters who are interested in performing structural measurements using crystalline samples with dimensions on the scale of nanometres to micrometres.

For decades, macromolecular crystallographers have strived to grow large (hundreds of micrometres) single crystals that can be used for crystallographic measurements using rotation X-ray methods, but new data-collection methods such as serial X-ray crystallography and MicroED require the reliable formation of crystals that are much smaller: typically hundreds of nanometres to tens of micrometres. Precise control of crystal size over this range is challenging, and others have developed methods that employ specialized equipment for *in situ* light-scattering measurements to evaluate crystal size in real time (Baitan *et al.*, 2018; Schubert *et al.*, 2017) and halt crystallization as it progresses. Instead, our work with CypA demonstrates a simple, alternative method for controlling the size of crystals during batch growth. Starting from crystallization conditions identified by microbatch screening in 96-well plates, we scaled up the crystallization volume and introduced agitation (by stirring) to control the crystal size. We observed that at higher stir rates (*i.e.* greater agitation)

crystals tend to be smaller and more concentrated in the resulting slurry. We speculate that stirring fractures the crystals when they reach a critical size, which exerts control over the crystal dimensions and also actively introduces seeds into the slurry. For CypA, we found that batch crystallization with stirring could be used to generate relatively monodisperse crystal slurries, in millilitre volumes, with crystal sizes in the range of micrometres to hundreds of micrometres. We believe that the batch-stirring protocol is likely to be useful for a variety of crystal systems beyond CypA; however, it may have limited utility for crystal systems that are more susceptible to physical damage.

With the ability to create large batches of CypA crystals, we could perform serial X-ray crystallography experiments at XFEL light sources, which generally consume a large amount of sample. We utilized CypA microcrystal slurries, prepared in an identical fashion, to evaluate several commonly used, injector-based sample-delivery strategies, including both electrospinning and viscous extrusion using two types of crystal carrier media. These delivery strategies exposed the crystals to extreme experimental conditions, including strong electric fields, high pressures and unusual carrier media. Crystal structures determined using each method revealed how the conditions imposed by the different sample-delivery systems perturbed either the crystal lattice or the protein structure.

We first observed that the different sample-delivery methods produce measurable differences in the distributions of unit-cell axis lengths for CypA microcrystals. Crystals measured using the MESH device (electrospinning) tended to have longer a, b and c axes than crystals measured using viscous carrier media. This could result from dehydration of the crystals by the viscogens (LCP or cellulose), which reduce the relative humidity of the crystallization mother liquor, or from the high pressures required to extrude the viscous carrier media through the injectors. The magnitude of the change in unit-cell parameters across different sample-delivery methods is similar to that which is observed for crystal cryocooling (Fig. 5). Additionally, we observed that crystals delivered using the MESH (electrospinning) device tended to be oriented more randomly than crystals delivered in a viscous carrier such as LCP (Supplementary Fig. S4). In the case of our CypA crystals, the dihedral space-group symmetry prevents the crystals from having a dipole moment that could cause them to assume a preferred orientation in the electric field introduced by the MESH injector. On the other hand, the slightly elongated crystal morphology led to orientation bias in the high-viscosity injector system, likely owing to shear forces resulting from the flow of the highly viscous liquid and the narrower inner diameter of the capillaries used in that device relative to the MESH. We expect that crystals with different properties, such as polar space-group symmetry or more isotropic morphologies, would have different behaviors with respect to preferred crystal orientation in the various injector systems.

Despite differences in unit-cell parameters and preferred crystal orientations, the overall quality of the reduced data sets resulting from each of the serial X-ray experiments was generally equivalent (Table 3). A notable inconsistency is that the $CC_{1/2}$ and R_{split} values for the LCP data set are slightly worse than for the other two data sets. This could be owing to the fact that the LCP data set contains the least images, or it could result from the high background created by X-ray scattering from the LCP matrix. The low-resolution signal is not as strong relative to the high-resolution signal for the LCP data set relative to the others, which we believe supports the latter hypothesis.

The atomic models generated from the three data sets were also strikingly similar (Table 4). Isomorphous difference maps ($F_o - F_o$) are a very sensitive method for detecting subtle changes in molecular structure, and we found that calculating such maps using pairs of serial X-ray data sets did produce strong ($>3.0\sigma$) peaks (Supplementary Fig. S10). The refined coordinates showed that these difference peaks corresponded to changes in atomic positions of the order of 0.1 Å or less, which are not significant with respect to the interpretation of structure–function relationships at this resolution. We also utilized a multi-conformer ensemble refinement approach as a way to assess the level of heterogeneity (model variance) that was present in each of the data sets. Our analysis focused on a network of catalytically important residues which are known to be dynamic (Eisenmesser et al., 2005; Fraser et al., 2009). We observed that within the catalytic network the refined ensembles reflect a similar level of heterogeneity across the different structures (Supplementary Fig. S6), which is generally supported by the correlations in the B factors derived from standard refinements (Supplementary Fig. S5). We did, however, observe that the ensemble derived from the MESH data shows enhanced heterogeneity relative to the other two data sets for a loop region including residues 69–74 (Supplementary Fig. S9). The conformation of this loop is stabilized by a key charged residue (Arg69; Caines et al., 2012), which may be perturbed by the electric field. While it has been shown that electric fields can be used to perturb conformational dynamics in proteins (Hekstra et al., 2016), we expect that the effect should be minimal in our MESH experiment because the crystals are randomly oriented relative to the applied electric field and the field is more than two orders of magnitude less than those that are intentionally used for perturbing conformational dynamics (Hekstra et al., 2016).

Our results show that the choice of microfluidic sample-delivery method has a minimal effect on the static crystal structure of CypA. Consequently, the choice of sample-delivery method for a serial X-ray crystallography experiment should be selected based on practical considerations related to the experiment, such as the requirement for laser perturbation or mixing in a time-resolved experiment. We note that an important aspect of microfluidic sample delivery that has not been rigorously explored is whether the effects of pressure or electric fields might have a time-dependent component as microcrystals exit the injector device and pass through the X-ray interaction region.

In addition to serial X-ray crystallography, MicroED also offers the ability to determine macromolecular structures at

high resolution using very small crystals with moderate solvent content. In stark contrast to serial X-ray crystallography experiments, which require hundreds of milligrams of protein and millilitres of microcrystal slurries, MicroED lies at the other extreme, allowing crystal structure determination using as little as a single microcrystal. Additionally, MicroED experiments have a significant advantage in that they are much more accessible than experiments performed at XFELs and require a substantially lower investment of time and resources. Using CypA crystals derived from our batch protocol, we encountered several challenges in preparing appropriately sized microcrystal samples on grids for measurement in the TEM. MicroED requires extremely small crystals, ideally less than 500 nm thick (Martynowycz *et al.*, 2019*b*), and we encountered difficulties in getting such small CypA crystals to remain on the grids after blotting away excess solvent. This may be owing to the specific surface chemistry of CypA crystals or may be a more general trend of high solvent-content crystals. As a result, we turned to a recent development in sample preparation that is widening the scope of MicroED by enabling measurements from crystals that are initially tens of micrometres thick by utilizing an FIB-milling process to machine large crystals into thin lamellae that are optimal for MicroED measurements (Martynowycz *et al.*,

2019*a*). The FIB-milling procedure allowed us to determine the structure of CypA by MicroED using a single crystal that was initially (before milling) similar in size to those which we used for serial X-ray experiments. We observed that the MicroED crystal structure of CypA had a slightly distorted unit cell relative to other reported CypA structures and we hypothesize that this could be owing to the damage from either blotting or FIB milling; however, more rigorous studies will be required to evaluate the specific effects of these sample-preparation procedures on MicroED structures. The sensitivity of the CypA crystals during preparation for MicroED could be related to their high solvent content. Our structure of CypA has the highest solvent content of any non-membrane-protein MicroED crystal structure determined to date using three-dimensional crystals (Fig. 7), demonstrating how improved sample preparation is expanding the technique to include more challenging crystal systems.

Our evaluation of modern protein microcrystallography techniques reveals that MicroED and serial X-ray crystallography are complementary methods for structural biology (Zatsepin *et al.*, 2019). The optimal experimental method for a microcrystallography experiment will depend upon various aspects of the macromolecular system that is being studied. For the determination of static, low-energy macromolecular structures, MicroED has substantial advantages over serial X-ray crystallography in terms of sample-preparation requirements, the quantity of material required and the ease of data collection. However, while serial X-ray crystallography experiments require large amounts of sample, specialized equipment that is only available at select X-ray light sources and substantial optimization of sample-delivery parameters, they also have their advantages. Importantly, serial X-ray measurements are performed at ambient temperature and can reveal physiological conformational ensembles of the crystallized molecules. On the other hand, we compared our MicroED structure with a cryogenic X-ray structure of CypA (PDB entry 3k0m) and observed that it was nearly identical and suffered from the same temperature-dependent reduction in conformational heterogeneity (Fig. 6). Because cooling rate is related to crystal size (Halle, 2004), it remains to be seen whether MicroED experiments using very small crystals (hundreds of nanometres) might capture a more physiological conformational ensemble. Our data do not shed light on this question, since the crystals used in our experiments were approximately $20 \times 20 \times 20$ μm at the time of freezing, before they were FIB-milled to an appropriate thickness. Finally, we note that for CypA, as well as other examples from the literature including lysozyme (Shi *et al.*, 2013; Nannenga *et al.*, 2014) and proteinase K (Hattne *et al.*, 2016), refinement R factors are much higher and resolutions are generally lower for MicroED structures than for X-ray structures. In our case, some of this might be improved by collecting more complete data from multiple FIB-milled crystals. However, more generally, we expect that this discrepancy will only improve as we gain a better understanding of how electrons interact with macromolecular crystals and develop data-analysis software that handles the processing of MicroED data and the
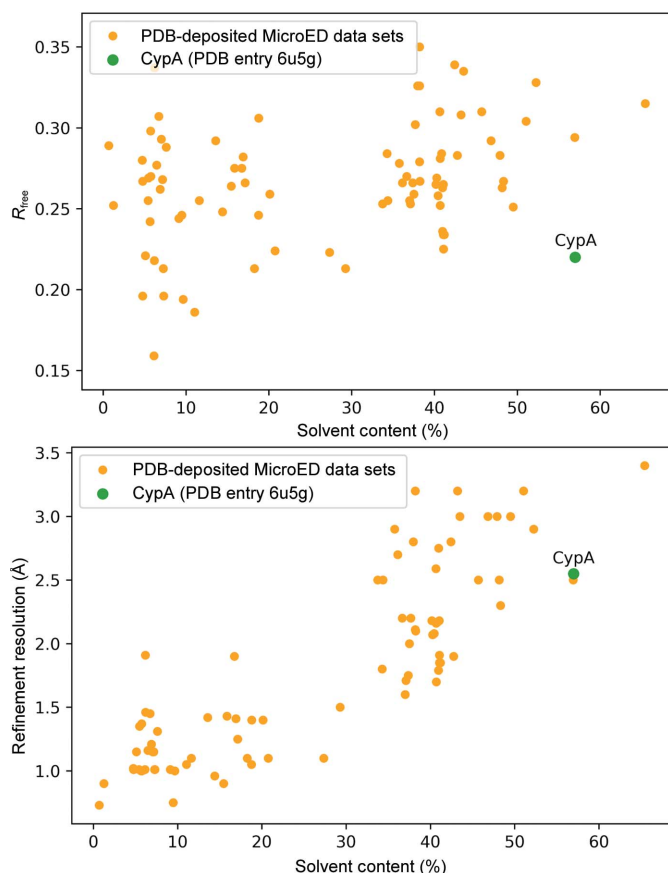

Figure 7
Survey of MicroED data sets deposited in the PDB. Structures determined from 3D crystals are shown as orange points and CypA is shown as a green point. The highest solvent-content point is PDB entry 3j7t, which is a membrane protein.

research papers

refinement of structural models based on electron scattering more appropriately.

References

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.

Baitan, D., Schubert, R., Meyer, A., Dierks, K., Perbandt, M. & Betzel, C. (2018). *J. Vis. Exp.*, e57070.

Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A. & Chapman, H. (2014). *J. Appl. Cryst.* **47**, 1118–1131.

Baxter, E. L., Aguila, L., Alonso-Mori, R., Barnes, C. O., Bonagura, C. A., Brehmer, W., Brunger, A. T., Calero, G., Caradoc-Davies, T. T., Chatterjee, R., Degrado, W. F., Fraser, J. M., Ibrahim, M., Kern, J., Kobilka, B. K., Kruse, A. C., Larsson, K. M., Lemke, H. T., Lyubimov, A. Y., Manglik, A., McPhillips, S. E., Norgren, E., Pang, S. S., Soltis, S. M., Song, J., Thomaston, J., Tsai, Y., Weis, W. I., Woldeyes, R. A., Yachandra, V., Yano, J., Zouni, A. & Cohen, A. E. (2016). *Acta Cryst.* D**72**, 2–11.

Brewster, A. S., Waterman, D. G., Parkhurst, J. M., Gildea, R. J., Young, I. D., O'Riordan, L. J., Yano, J., Winter, G., Evans, G. & Sauter, N. K. (2018). *Acta Cryst.* D**74**, 877–894.

Burnley, B. T., Afonine, P. V., Adams, P. D. & Gros, P. (2012). *eLife*, **1**, e00311.

Caines, M. E. C., Bichel, K., Price, A. J., McEwan, W. A., Towers, G. J., Willett, B. J., Freund, S. M. V. & James, L. C. (2012). *Nat. Struct. Mol. Biol.* **19**, 411–416.

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.-U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R. M., 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, J. C. H. (2011). *Nature*, **470**, 73–77.

Chernov, A. A. (1999). *J. Cryst. Growth*, **196**, 524–534.

Clabbers, M. T. B. & Abrahams, J. P. (2018). *Crystallogr. Rev.* **24**, 176–204.

Clabbers, M. T. B., Gruene, T., Parkhurst, J. M., Abrahams, J. P. & Waterman, D. G. (2018). *Acta Cryst.* D**74**, 506–518.

Coquelle, N., Sliwa, M., Woodhouse, J., Schirò, G., Adam, V., Aquila, A., Barends, T. R. M., Boutet, S., Byrdin, M., Carbajo, S., De la Mora, E., Doak, R. B., Feliks, M., Fieschi, F., Foucar, L., Guillon, V., Hilpert, M., Hunter, M. S., Jakobs, S., Koglin, J. E., Kovacsova, G., Lane, T. J., Lévy, B., Liang, M., Nass, K., Ridard, J., Robinson, J. S., Roome, C. M., Ruckebusch, C., Seaberg, M., Thepaut, M., Cammarata, M., Demachy, I., Field, M., Shoeman, R. L., Bourgeois, D., Colletier, J.-P., Schlichting, I. & Weik, M. (2018). *Nat. Chem.* **10**, 31–37.

Cruz, M. J. de la, Hattne, J., Shi, D., Seidler, P., Rodriguez, J., Reyes, F. E., Sawaya, M. R., Cascio, D., Weiss, S. C., Kim, S. K., Hinck, C. S., Hinck, A. P., Calero, G., Eisenberg, D. & Gonen, T. (2017). *Nat. Methods*, **14**, 399–402.

Cusack, S., Belrhali, H., Bram, A., Burghammer, M., Perrakis, A. & Riekel, C. (1998). *Nat. Struct. Mol. Biol.* **5**, 634–637.

Duisenberg, A. J. M. (1992). *J. Appl. Cryst.* **25**, 92–96.

Duyvesteyn, H. M. E., Kotecha, A., Ginn, H. M., Hecksel, C. W., Beale, E. V., de Haas, F., Evans, G., Zhang, P., Chiu, W. & Stuart, D. I. (2018). *Proc. Natl Acad. Sci. USA*, **115**, 9569–9573.

Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E. & Kern, D. (2005). *Nature*, **438**, 117–121.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.

Fraser, J. S., Clarkson, M. W., Degnan, S. C., Erion, R., Kern, D. & Alber, T. (2009). *Nature*, **462**, 669–673.

Fuller, F. D., Gul, S., Chatterjee, R., Burgie, E. S., Young, I. D., Lebrette, H., Srinivas, V., Brewster, A. S., Michels-Clark, T., Clinger, J. A., Andi, B., Ibrahim, M., Pastor, E., de Lichtenberg, C., Hussein, R., Pollock, C. J., Zhang, M., Stan, C. A., Kroll, T., Fransson, T., Weninger, C., Kubin, M., Aller, P., Lassalle, L., Bräuer, P., Miller, M. D., Amin, M., Koroidov, S., Roessler, C. G., Allaire, M., Sierra, R. G., Docker, P. T., Glownia, J. M., Nelson, S., Koglin, J. E., Zhu, D., Chollet, M., Song, S., Lemke, H., Liang, M., Sokaras, D., Alonso-Mori, R., Zouni, A., Messinger, J., Bergmann, U., Boal, A. K., Bollinger, J. M. Jr, Krebs, C., Högbom, M., Phillips, G. N. Jr, Vierstra, R. D., Sauter, N. K., Orville, A. M., Kern, J., Yachandra, V. K. & Yano, J. (2017). *Nat. Methods*, **14**, 443–449.

Halle, B. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 4793–4798.

Hattne, J., Shi, D., de la Cruz, M. J., Reyes, F. E. & Gonen, T. (2016). *J. Appl. Cryst.* **49**, 1029–1034.

Hekstra, D. R., White, K. I., Socolich, M. A., Henning, R. W., Šrajer, V. & Ranganathan, R. (2016). *Nature*, **540**, 400–405.

Hunter, M. S., Segelke, B., Messerschmidt, M., Williams, G. J., Zatsepin, N. A., Barty, A., Benner, W. H., Carlson, D. B., Coleman, M., Graf, A., Hau-Riege, S. P., Pardini, T., Seibert, M. M., Evans, J., Boutet, S. & Frank, M. (2014). *Sci. Rep.* **4**, 6026.

Ibrahim, M., Chatterjee, R., Hellmich, J., Tran, R., Bommer, M., Yachandra, V. K., Yano, J., Kern, J. & Zouni, A. (2015). *Struct. Dyn.* **2**, 041705.

Ishchenko, A., Cherezov, V. & Liu, W. (2016). *J. Vis. Exp.*, e54463.

Ishikawa, T., Aoyagi, H., Asaka, T., Asano, Y., Azumi, N., Bizen, T., Ego, H., Fukami, K., Fukui, T., Furukawa, Y., Goto, S., Hanaki, H., Hara, T., Hasegawa, T., Hatsui, T., Higashiya, A., Hirono, T., Hosoda, N., Ishii, M., Inagaki, T., Inubushi, Y., Itoga, T., Joti, Y., Kago, M., Kameshima, T., Kimura, H., Kirihara, Y., Kiyomichi, A., Kobayashi, T., Kondo, C., Kudo, T., Maesaka, H., Maréchal, X. M., Masuda, T., Matsubara, S., Matsumoto, T., Matsushita, T., Matsui, S., Nagasono, M., Nariyama, N., Ohashi, H., Ohata, T., Ohshima, T., Ono, S., Otake, Y., Saji, C., Sakurai, T., Sato, T., Sawada, K., Seike, T., Shirasawa, K., Sugimoto, T., Suzuki, S., Takahashi, S., Takebe, H., Takeshita, K., Tamasaku, K., Tanaka, H., Tanaka, R., Tanaka, T., Togashi, T., Togawa, K., Tokuhisa, A., Tomizawa, H., Tono, K., Wu, S., Yabashi, M., Yamaga, M., Yamashita, A., Yanagida, K., Zhang, C., Shintake, T., Kitamura, H. & Kumagai, N. (2012). *Nat. Photon.* **6**, 540–544.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kameshima, T., Ono, S., Kudo, T., Ozaki, K., Kirihara, Y., Kobayashi, K., Inubushi, Y., Yabashi, M., Horigome, T., Holland, A., Holland, K., Burt, D., Murao, H. & Hatsui, T. (2014). *Rev. Sci. Instrum.* **85**, 033110.

Keedy, D. A., Kenner, L. R., Warkentin, M., Woldeyes, R. A., Hopkins, J. B., Thompson, M. C., Brewster, A. S., Van Benschoten, A. H., Baxter, E. L., Uervirojnangkoorn, M., McPhillips, S. E., Song, J., Alonso-Mori, R., Holton, J. M., Weis, W. I., Brunger, A. T., Soltis, S. M., Lemke, H., Gonzalez, A., Sauter, N. K., Cohen, A. E., van den Bedem, H., Thorne, R. E. & Fraser, J. S. (2015). *eLife*, **4**, e07574.

Kubo, M., Nango, E., Tono, K., Kimura, T., Owada, S., Song, C., Mafuné, F., Miyajima, K., Takeda, Y., Kohno, J., Miyauchi, N., Nakane, T., Tanaka, T., Nomura, T., Davidsson, J., Tanaka, R., Murata, M., Kameshima, T., Hatsui, T., Joti, Y., Neutze, R., Yabashi, M. & Iwata, S. (2017). *J. Synchrotron Rad.* **24**, 1086–1091.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.

Liu, W., Wacker, D., Gati, C., Han, G. W., James, D., Wang, D., Nelson, G., Weierstall, U., Katritch, V., Barty, A., Zatsepin, N. A., Li, D., Messerschmidt, M., Boutet, S., Williams, G. J., Koglin, J. E., Seibert, M. M., Wang, C., Shah, S. T. A., Basu, S., Fromme, R., Kupitz, C., Rendek, K. N., Grotjohann, I., Fromme, P., Kirian, R. A., Beyerlein, K. R., White, T. A., Chapman, H. N., Caffrey, M., Spence, J. C. H., Stevens, R. C. & Cherezov, V. (2013). *Science*, **342**, 1521–1524.

Martin-Garcia, J. M., Conrad, C. E., Nelson, G., Stander, N., Zatsepin, N. A., Zook, J., Zhu, L., Geiger, J., Chun, E., Kissick, D., Hilgart, M. C., Ogata, C., Ishchenko, A., Nagaratnam, N., Roy-Chowdhury, S., Coe, J., Subramanian, G., Schaffer, A., James, D., Ketwala, G., Venugopalan, N., Xu, S., Corcoran, S., Ferguson, D., Weierstall, U., Spence, J. C. H., Cherezov, V., Fromme, P., Fischetti, R. F. & Liu, W. (2017). *IUCrJ*, **4**, 439–454.

Martynowycz, M. W., Zhao, W., Hattne, J., Jensen, G. J. & Gonen, T. (2019a). *Structure*, **27**, 545–548.

Martynowycz, M. W., Zhao, W., Hattne, J., Jensen, G. J. & Gonen, T. (2019b). *Structure*, **27**, 1594–1600.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Mueller, C., Marx, A., Epp, S. W., Zhong, Y., Kuo, A., Balo, A. R., Soman, J., Schotte, F., Lemke, H. T., Owen, R. L., Pai, E. F., Pearson, A. R., Olson, J. S., Anfinrud, P. A., Ernst, O. P. & Miller, R. J. D. (2015). *Struct. Dyn.* **2**, 054302.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nakane, T., Joti, Y., Tono, K., Yabashi, M., Nango, E., Iwata, S., Ishitani, R. & Nureki, O. (2016). *J. Appl. Cryst.* **49**, 1035–1041.

Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A.-N., Standfuss, J., Neutze, R. & Iwata, S. (2016). *Science*, **354**, 1552–1557.

Nannenga, B. L. & Gonen, T. (2019). *Nat. Methods*, **16**, 369–379.

Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. (2014). *Nat. Methods*, **11**, 927–930.

Olmos, J. L., Pandey, S., Martin-Garcia, J. M., Calvey, G., Katz, A., Knoska, J., Kupitz, C., Hunter, M. S., Liang, M., Oberthuer, D., Yefanov, O., Wiedorn, M., Heyman, M., Holl, M., Pande, K., Barty, A., Miller, M. D., Stern, S., Roy-Chowdhury, S., Coe, J., Nagaratnam, N., Zook, J., Verburgt, J., Norwood, T., Poudyal, I., Xu, D., Koglin, J., Seaberg, M. H., Zhao, Y., Bajt, S., Grant, T., Mariani, V., Nelson, G., Subramanian, G., Bae, E., Fromme, R., Fung, R., Schwander, P., Frank, M., White, T. A., Weierstall, U., Zatsepin, N., Spence, J., Fromme, P., Chapman, H. N., Pollack, L., Tremblay, L., Ourmazd, A., Phillips, G. N. & Schmidt, M. (2018). *BMC Biol.* **16**, 59.

Rodriguez, J. A., Ivanova, M. I., Sawaya, M. R., Cascio, D., Reyes, F. E., Shi, D., Sangwan, S., Guenther, E. L., Johnson, L. M., Zhang, M., Jiang, L., Arbing, M. A., Nannenga, B. L., Hattne, J., Whitelegge, J., Brewster, A. S., Messerschmidt, M., Boutet, S., Sauter, N. K., Gonen, T. & Eisenberg, D. S. (2015). *Nature*, **525**, 486–490.

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T. & Eliceiri, K. W. (2017). *BMC Bioinformatics*, **18**, 529.

Sauter, N. K. (2015). *J. Synchrotron Rad.* **22**, 239–248.

Schubert, R., Meyer, A., Baitan, D., Dierks, K., Perbandt, M. & Betzel, C. (2017). *Cryst. Growth Des.* **17**, 954–958.

Shi, D., Nannenga, B. L., Iadanza, M. G. & Gonen, T. (2013). *eLife*, **2**, e01345.

Sierra, R. G., Batyuk, A., Sun, Z., Aquila, A., Hunter, M. S., Lane, T. J., Liang, M., Yoon, C. H., Alonso-Mori, R., Armenta, R., Castagna, J.-C., Hollenbeck, M., Osier, T. O., Hayes, M., Aldrich, J., Curtis, R., Koglin, J. E., Rendahl, T., Rodriguez, E., Carbajo, S., Guillet, S., Paul, R., Hart, P., Nakahara, K., Carini, G., DeMirci, H., Dao, E. H., Hayes, B. M., Rao, Y. P., Chollet, M., Feng, Y., Fuller, F. D., Kupitz, C., Sato, T., Seaberg, M. H., Song, S., van Driel, T. B., Yavas, H., Zhu, D., Cohen, A. E., Wakatsuki, S. & Boutet, S. (2019). *J. Synchrotron Rad.* **26**, 346–357.

Sierra, R. G., Gati, C., Laksmono, H., Dao, E. H., Gul, S., Fuller, F., Kern, J., Chatterjee, R., Ibrahim, M., Brewster, A. S., Young, I. D., Michels-Clark, T., Aquila, A., Liang, M., Hunter, M. S., Koglin, J. E., Boutet, S., Junco, E. A., Hayes, B., Bogan, M. J., Hampton, C. Y., Puglisi, E. V., Sauter, N. K., Stan, C. A., Zouni, A., Yano, J., Yachandra, V. K., Soltis, S. M., Puglisi, J. D. & DeMirci, H. (2016). *Nat. Methods*, **13**, 59–62.

Sierra, R. G., Laksmono, H., Kern, J., Tran, R., Hattne, J., Alonso-Mori, R., Lassalle-Kaiser, B., Glöckner, C., Hellmich, J., Schafer,

D. W., Echols, N., Gildea, R. J., Grosse-Kunstleve, R. W., Sellberg, J., McQueen, T. A., Fry, A. R., Messerschmidt, M. M., Miahnahri, A., Seibert, M. M., Hampton, C. Y., Starodub, D., Loh, N. D., Sokaras, D., Weng, T.-C., Zwart, P. H., Glatzel, P., Milathianaki, D., White, W. E., Adams, P. D., Williams, G. J., Boutet, S., Zouni, A., Messinger, J., Sauter, N. K., Bergmann, U., Yano, J., Yachandra, V. K. & Bogan, M. J. (2012). *Acta Cryst.* D**68**, 1584–1587.

Suga, M., Akita, F., Sugahara, M., Kubo, M., Nakajima, Y., Nakane, T., Yamashita, K., Umena, Y., Nakabayashi, M., Yamane, T., Nakano, T., Suzuki, M., Masuda, T., Inoue, S., Kimura, T., Nomura, T., Yonekura, S., Yu, L.-J., Sakamoto, T., Motomura, T., Chen, J.-H., Kato, Y., Noguchi, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Nango, E., Tanaka, R., Naitow, H., Matsuura, Y., Yamashita, A., Yamamoto, M., Nureki, O., Yabashi, M., Ishikawa, T., Iwata, S. & Shen, J.-R. (2017). *Nature*, **543**, 131–135.

Sugahara, M., Mizohata, E., Nango, E., Suzuki, M., Tanaka, T., Masuda, T., Tanaka, R., Shimamura, T., Tanaka, Y., Suno, C., Ihara, K., Pan, D., Kakinouchi, K., Sugiyama, S., Murata, M., Inoue, T., Tono, K., Song, C., Park, J., Kameshima, T., Hatsui, T., Joti, Y., Yabashi, M. & Iwata, S. (2015). *Nat. Methods*, **12**, 61–63.

Sugahara, M., Nakane, T., Masuda, T., Suzuki, M., Inoue, S., Song, C., Tanaka, R., Nakatsu, T., Mizohata, E., Yumoto, F., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Numata, K., Nango, E. & Iwata, S. (2017). *Sci. Rep.* **7**, 703.

Tono, K., Hara, T., Yabashi, M. & Tanaka, H. (2019). *J. Synchrotron Rad.* **26**, 595–602.

Tono, K., Nango, E., Sugahara, M., Song, C., Park, J., Tanaka, T., Tanaka, R., Joti, Y., Kameshima, T., Ono, S., Hatsui, T., Mizohata, E., Suzuki, M., Shimamura, T., Tanaka, Y., Iwata, S. & Yabashi, M. (2015). *J. Synchrotron Rad.* **22**, 532–537.

Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C. H., Doak, R. B., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Han, G. W., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G. J., Koglin, J. E., Seibert, M. M., Klinker, M., Gati, C., Shoeman, R. L., Barty, A., Chapman, H. N., Kirian, R. A., Beyerlein, K. R., Stevens, R. C., Li, D., Shah, S. T. A., Howe, N., Caffrey, M. & Cherezov, V. (2014). *Nat. Commun.* **5**, 3309.

White, T. A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K. R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C. H., Diederichs, K. & Chapman, H. N. (2016). *J. Appl. Cryst.* **49**, 680–689.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Cryst.* D**74**, 85–97.

Young, I. D., Ibrahim, M., Chatterjee, R., Gul, S., Fuller, F. D., Koroidov, S., Brewster, A. S., Tran, R., Alonso-Mori, R., Kroll, T., Michels-Clark, T., Laksmono, H., Sierra, R. G., Stan, C. A., Hussein, R., Zhang, M., Douthit, L., Kubin, M., de Lichtenberg, C., Vo Pham, L., Nilsson, H., Cheah, M. H., Shevela, D., Saracini, C., Bean, M. A., Seuffert, I., Sokaras, D., Weng, T.-C., Pastor, E., Weninger, C., Fransson, T., Lassalle, L., Bräuer, P., Aller, P., Docker, P. T., Andi, B., Orville, A. M., Glownia, J. M., Nelson, S., Sikorski, M., Zhu, D., Hunter, M. S., Lane, T. J., Aquila, A., Koglin, J. E., Robinson, J., Liang, M., Boutet, S., Lyubimov, A. Y., Uervirojnangkoorn, M., Moriarty, N. W., Liebschner, D., Afonine, P. V., Waterman, D. G., Evans, G., Wernet, P., Dobbek, H., Weis, W. I., Brunger, A. T., Zwart, P. H., Adams, P. D., Zouni, A., Messinger, J., Bergmann, U., Sauter, N. K., Kern, J., Yachandra, V. K. & Yano, J. (2016). *Nature*, **540**, 453–457.

Zander, U., Bourenkov, G., Popov, A. N., de Sanctis, D., Svensson, O., McCarthy, A. A., Round, E., Gordeliy, V., Mueller-Dieckmann, C. & Leonard, G. A. (2015). *Acta Cryst.* D**71**, 2328–2343.

Zatsepin, N. A., Li, C., Colasurd, P. & Nannenga, B. L. (2019). *Curr. Opin. Struct. Biol.* **58**, 286–293.

Zhou, X. E., Gao, X., Barty, A., Kang, Y., He, Y., Liu, W., Ishchenko, A., White, T. A., Yefanov, O., Han, G. W., Xu, Q., de Waal, P. W., Suino-Powell, K. M., Boutet, S., Williams, G. J., Wang, M., Li, D., Caffrey, M., Chapman, H. N., Spence, J. C. H., Fromme, P., Weierstall, U., Stevens, R. C., Cherezov, V., Melcher, K. & Xu, H. E. (2016). *Sci Data*, **3**, 160021.

IUCrJ

Volume 7 (2020)

Supporting information for article:

Comparing serial X-ray crystallography and microcrystal electron diffraction (MicroED) as methods for routine structure determination from small macromolecular crystals

Alexander M. Wolff, Iris D. Young, Raymond G. Sierra, Aaron S. Brewster, Michael W. Martynowycz, Eriko Nango, Michihiro Sugahara, Takanori Nakane, Kazutaka Ito, Andrew Aquila, Asmit Bhowmick, Justin T. Biel, Sergio Carbajo, Aina E. Cohen, Saul Cortez, Ana Gonzalez, Tomoya Hino, Dohyun Im, Jake D. Koralek, Minoru Kubo, Tomas S. Lazarou, Takashi Nomura, Shigeki Owada, Avi J. Samelson, Tomoyuki Tanaka, Rie Tanaka, Erin M. Thompson, Henry van den Bedem, Rahel A. Woldeyes, Fumiaki Yumoto, Wei Zhao, Kensuke Tono, Sebastien Boutet, So Iwata, Tamir Gonen, Nicholas K. Sauter, James S. Fraser and Michael C. Thompson
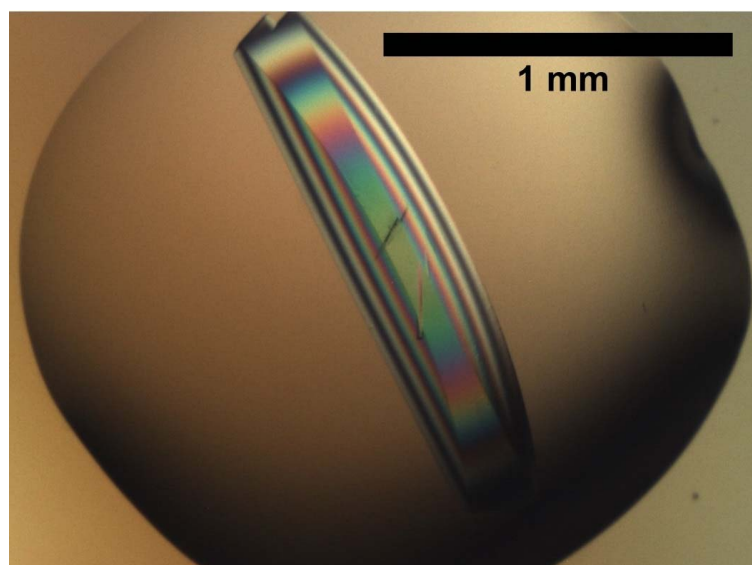
Figure S1 Previously, CypA crystallization conditions were optimized to produce very large single crystals. The pictured crystal (viewed using a cross-polarizer) was over 1 mm long and was visible to the naked eye.
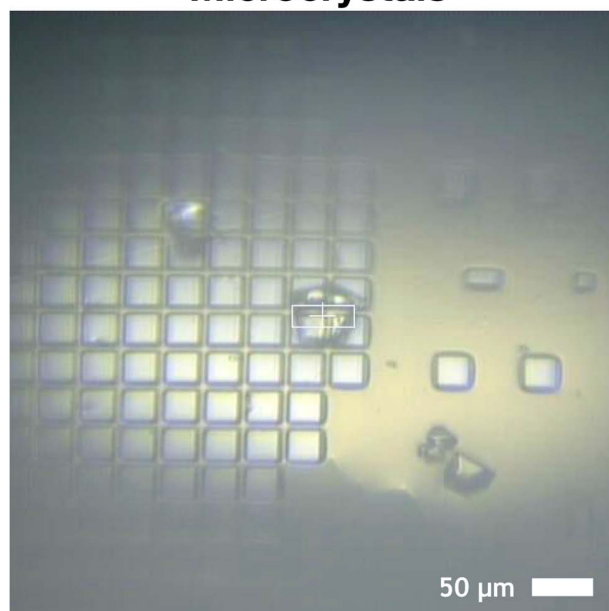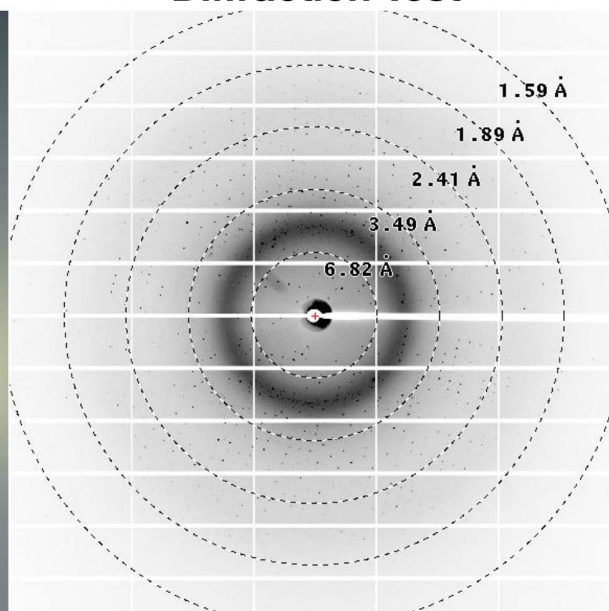

Figure S2 Image of microcrystals on a Mitegen micromesh grid (left) and the subsequent diffraction from one of these crystals (right). Crystals were 20-50 μm in all dimensions, and diffracted to the edge of the detector at 1.59 Å.
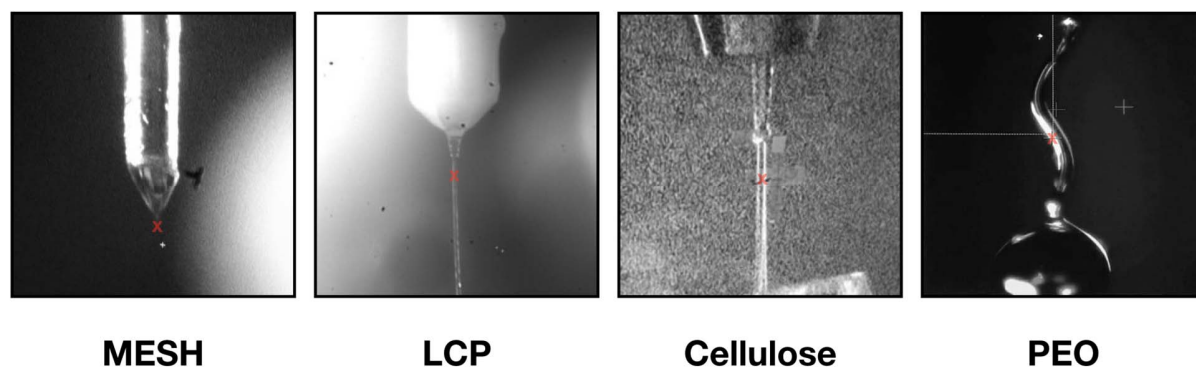
MESH   **LCP**   **Cellulose**   **PEO**

Figure S3 Images of jets delivering microcrystal slurries to the XFEL interaction point (red "x" in each image). Minimal viscogens were added to the crystalline slurry for the MESH injector system. When using a viscous extrusion type of injector, a variety of carrier media were tested, including LCP, Cellulose, and PEO.
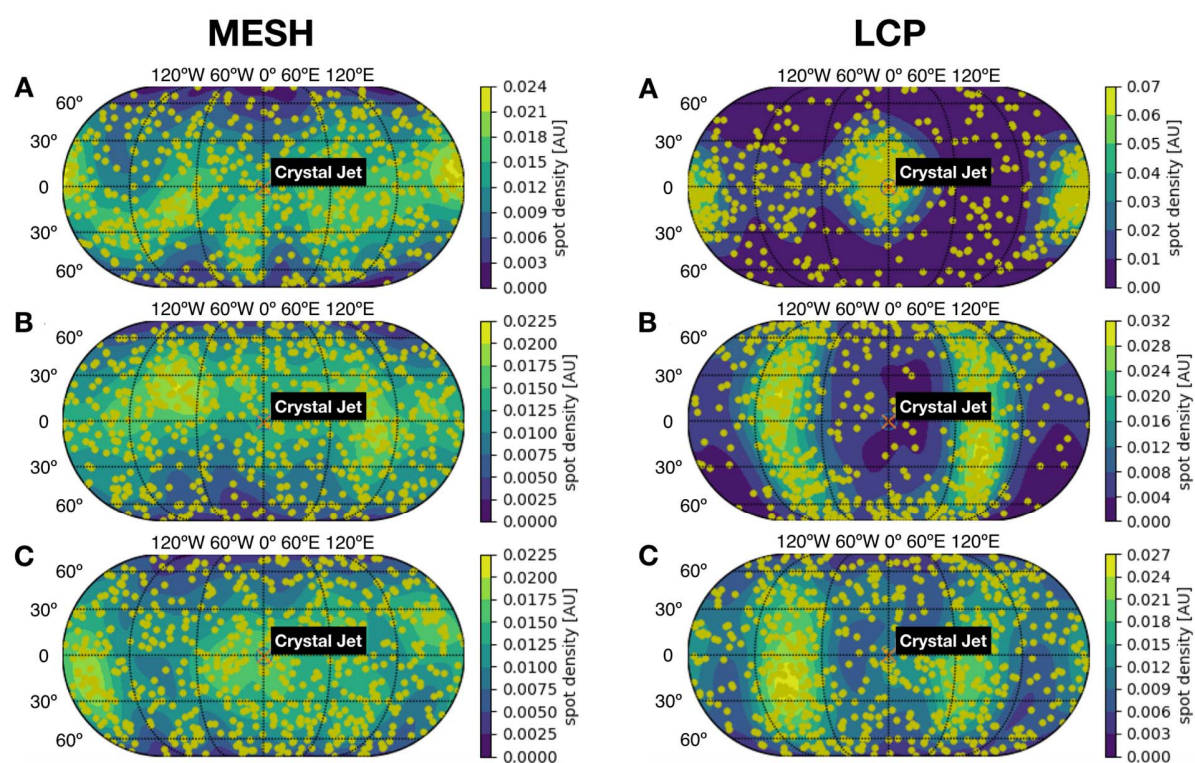
Figure S4   Maps of crystal axis (A, B, C) orientations from the MESH data collection (left panel) and from the LCP data collection (right panel). Axis orientations (in polar coordinates) for individual crystals are depicted as yellow spots, and the background color reflects the frequency with which a particular orientation was observed. A subset of each dataset is shown for visual clarity. The orientations appear evenly distributed for the MESH data, with no major bias introduced by the electric field created by the injection system. The viscosity of the LCP carrier media appears to have induced an orientation bias, but did not prohibit collection of a complete dataset with high redundancy.
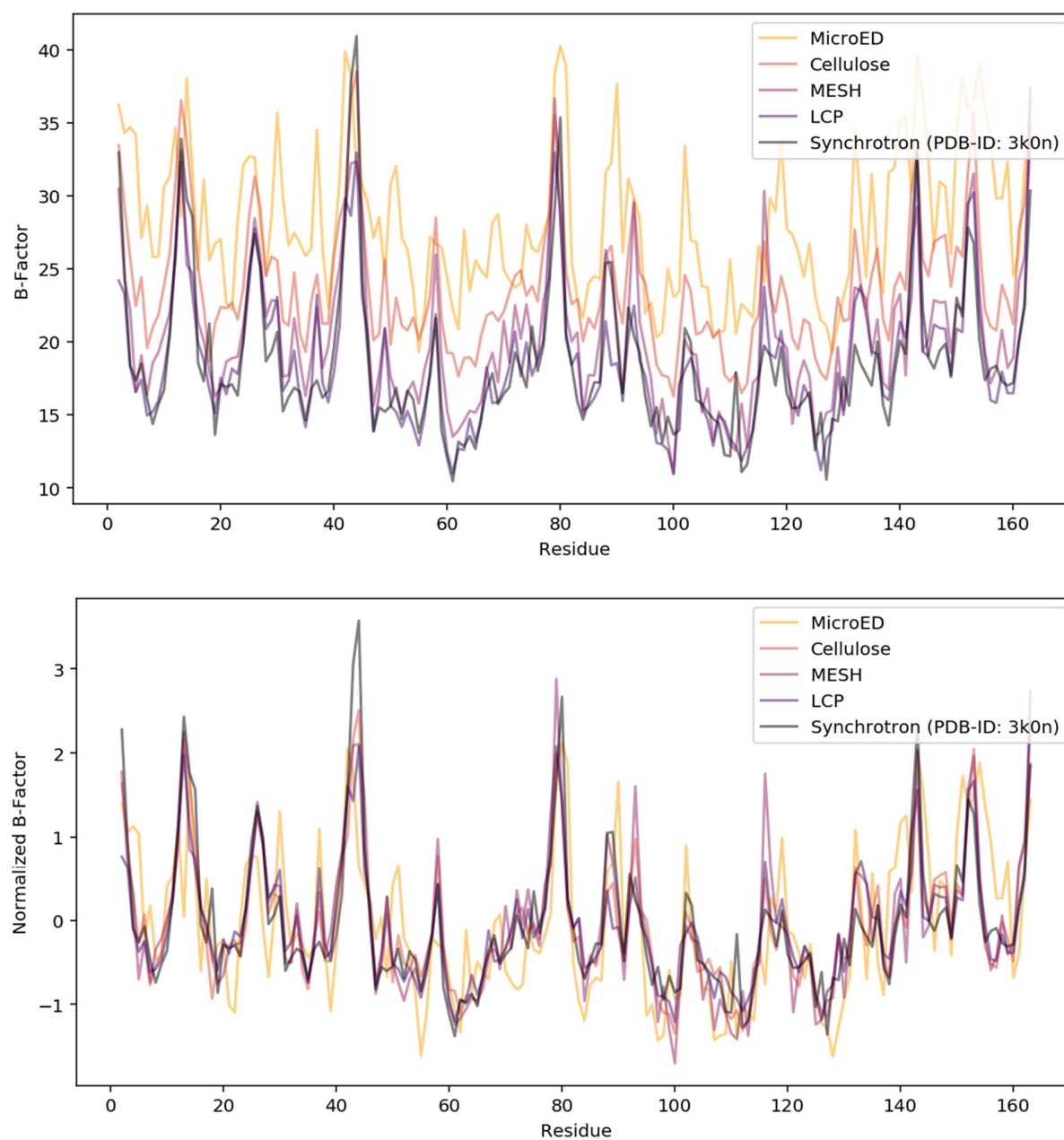
Figure S5 Raw (top) and normalized (bottom) B-factor per residue across data sets. A previously published structure, solved at room temperature using rotational collection from a single crystal (PDB ID *3K0N*) is provided for reference. Most variation is systematic, and thus is removed by normalization.
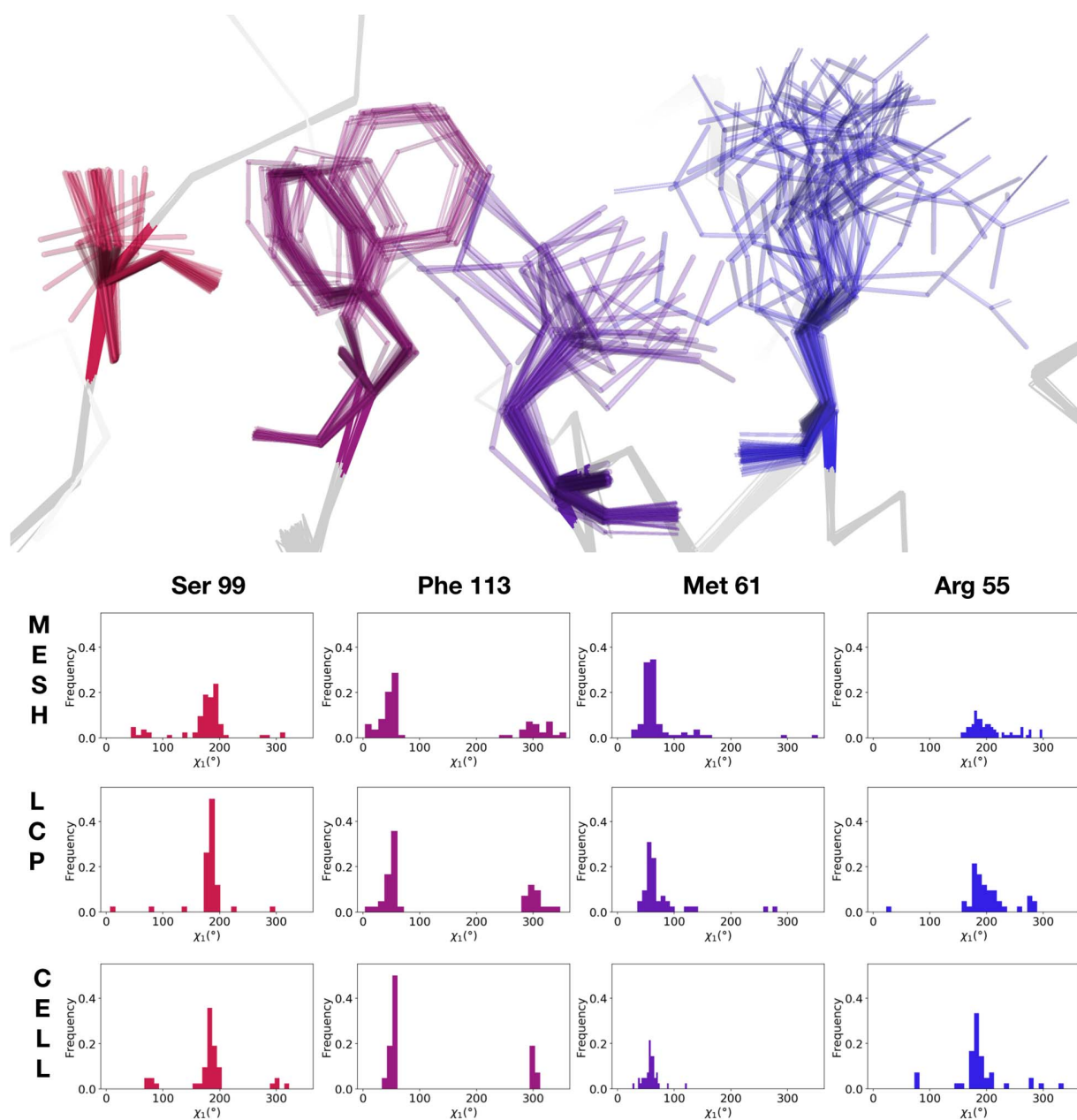
Figure S6 Visualization of the ensemble of conformers generated via phenix.ensemble_refine for the three serial XFEL datasets. The analysis focuses on a network of amino acid side chains that are known to be dynamic and important for function. In the top panel, the ensemble model for the cellulose dataset is displayed. Sticks are shown for the residues of interest (R55, M61, S99, F113), while the backbone is displayed as a ribbon for the rest of the structure. In the bottom panel, histograms of the distribution of chi1 angles are plotted for each of the four residues from the respective ensemble. Pairwise chi-square analysis revealed no significant differences between the rotamer distributions.
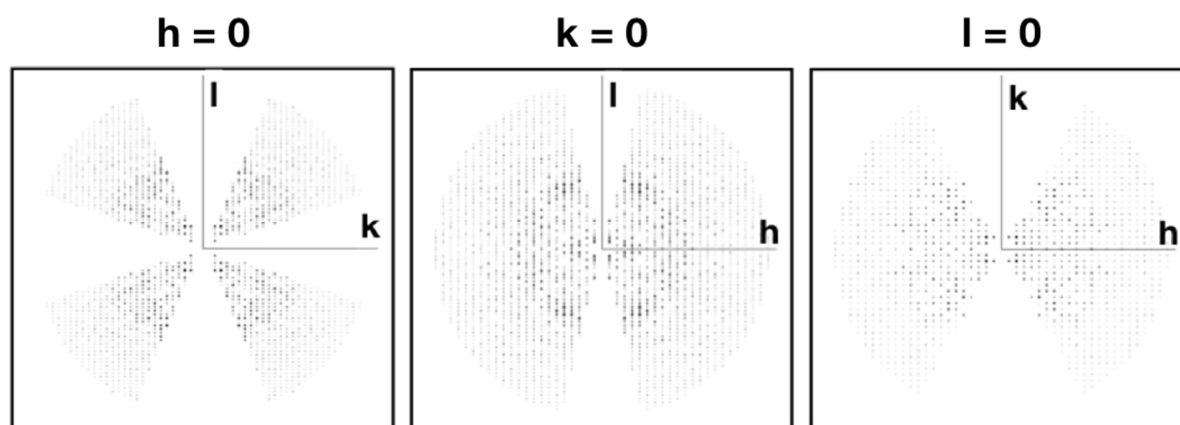
Figure S7 MicroED data visualized on two-dimensional slices of the reciprocal lattice. Measured reflections are visualized in black. Missing measurements along the kl plane may contribute to challenges in assignment of unit cell dimensions.
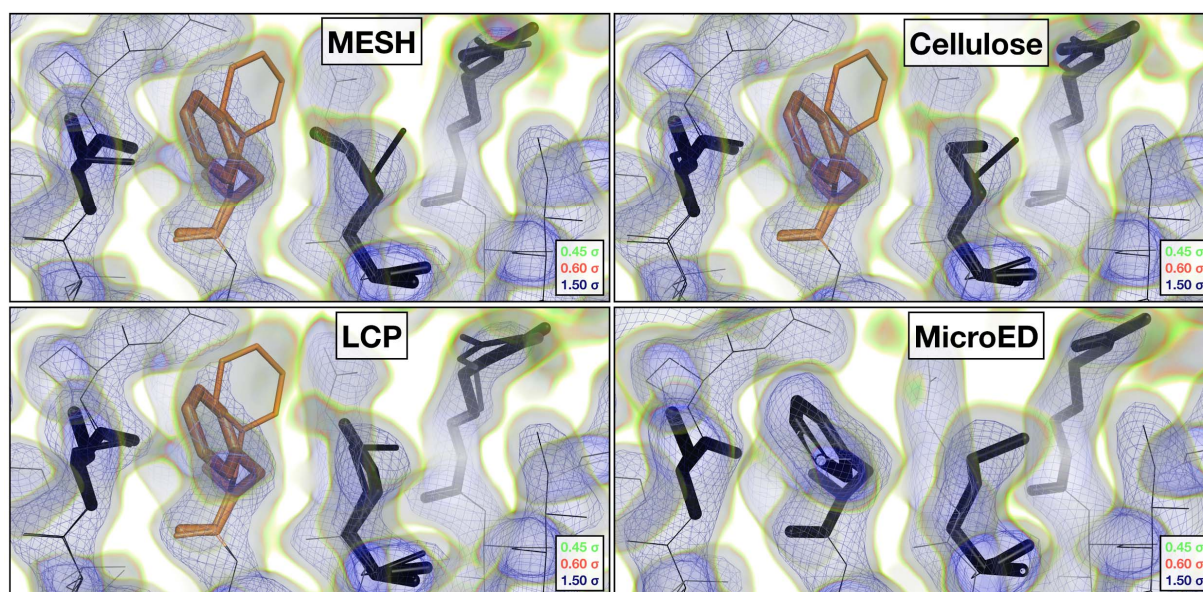

Figure S8 Visualization of all four datasets truncated at 2.5 Å. Comparison of the 2mFoFc maps and refined multi-conformer models reveals evidence for alternative conformations in the room temperature data (MESH, LCP, Cellulose), while cryogenic data (MicroED) supports a single conformer model.
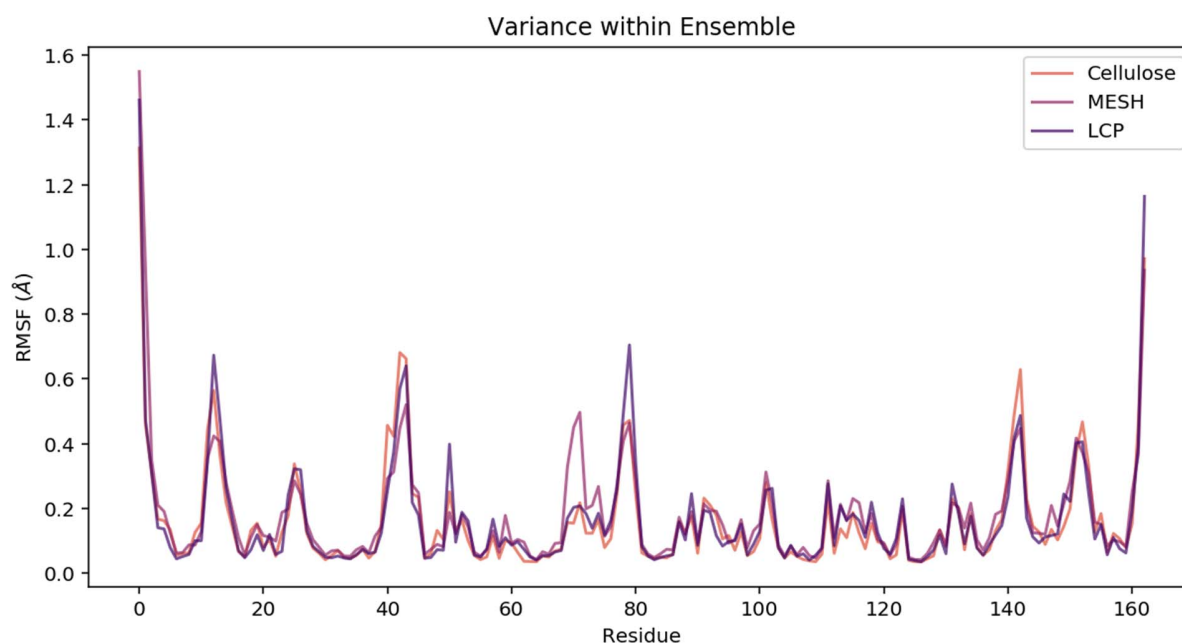
Figure S9 Average RMSF per residue for each ensemble generated for the serial XFEL datasets. A loop containing residues 64-74 samples more conformations in the MESH ensemble than in the other two.
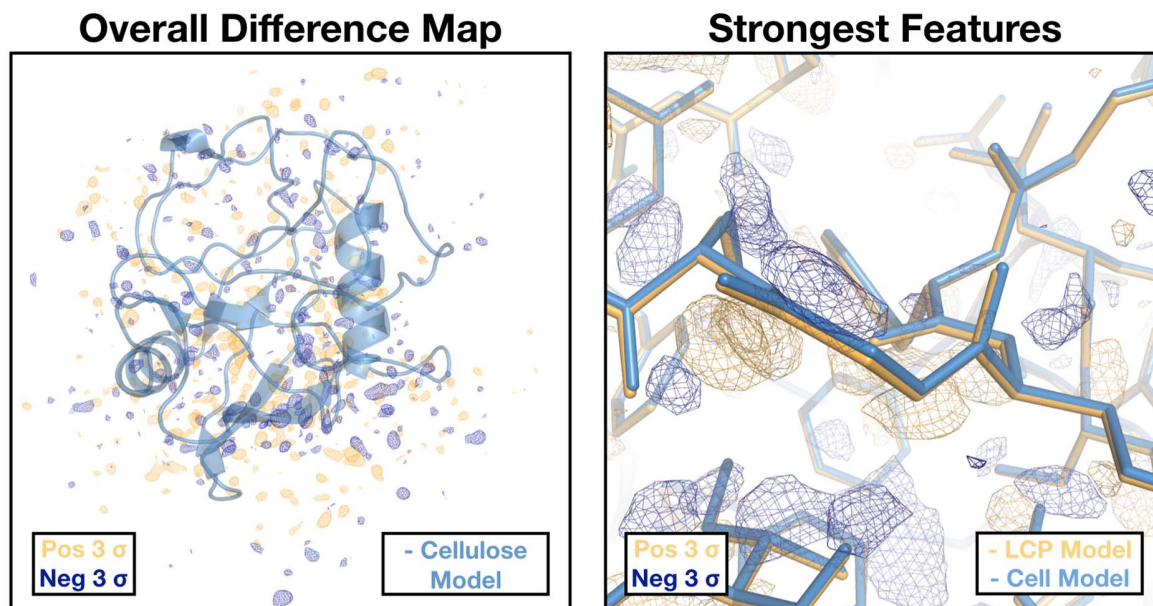

Figure S10 Visualization of an Fo-Fo difference map, displayed at 3 sigma, for the LCP and Cellulose datasets. In the left panel the overall difference map is shown, with the Cellulose model visualized for structural context. In the right panel the strongest features within the map are visualized, with both the LCP model and the Cellulose (Cell) model visualized for comparison.

Table S1   Results of data indexing with XDS and unit cell refinement with REFMAC5.

Unit cell axis	a	b	c
XDS indexed (merged)	42.40	53.40	87.76
XDS indexed (std. dev.)	0.05	0.03	0.06
REFMAC5 refined	42.56	53.56	88.07

Table S2   Quantitative comparison of unit cell distributions across XFEL datasets

One-way ANOVA with post-hoc Tukey HSD tests revealed that all three datasets were significantly different across all three edges of the crystalline unit cell, as denoted by the respective p-values for each test

	A axis	B axis	C axis
MESH Unit Cell (N = 18358)	43.32±0.11 Å	52.94±0.09 Å	89.87±0.21 Å
LCP Unit Cell (N = 11821)	43.10±0.18 Å	52.65±0.13 Å	89.29±0.26 Å
Cellulose Unit Cell (N = 23947)	43.00±0.26 Å	52.60±0.23 Å	89.20±0.37 Å
ANOVA	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
MESH vs. LCP Tukey HSD Post-hoc Test	Difference = -0.22 95 % CI = -0.2256 to -0.2144 $p < 0.0001$	Difference = -0.29 95 % CI = -0.2948 to -0.2852 $p < 0.0001$	Difference = -0.58 95 % CI = -0.5883 to -0.5717 $p < 0.0001$
MESH vs. Cellulose Tukey HSD Post-hoc Test	Difference = -0.32 95 % CI = -0.3247 to -0.3153 $p < 0.0001$	Difference = -0.34 95 % CI = -0.3440 to -0.3360 $p < 0.0001$	Difference = -0.67 95 % CI = -0.6769 to -0.6631 $p < 0.0001$
LCP vs. Cellulose Tukey HSD Post-hoc Test	Difference = -0.10 95 % CI = -0.1053 to -0.0947 $p < 0.0001$	Difference = -0.05 95 % CI = -0.0545 to -0.0455 $p < 0.0001$	Difference = -0.09 95 % CI = -0.0979 to -0.0821 $p < 0.0001$

.