Serial millisecond crystallography for routine room-temperature structure determination at synchrotrons

Tobias Weinert1, Natacha Olieric1, Robert Cheng2, Steffen Brünle3, Daniel James1, Dmitry Ozerov4, Dardan Gashi1,5, Laura Vera6, May Marsh6, Kathrin Jaeger1, Florian Dworkowski6, Ezequiel Panepucci6, Shibom Basu6, Petr Skopintsev1, Andrew S. Doré7, Tian Geng7, Robert M. Cooke7, Mengning Liang8, Andrea E. Prota1, Valerie Panneels1, Przemyslaw Nogly1, Ulrich Ermler3, Gebhard Schertler1,9, Michael Hennig2, Michel O. Steinmetz1,10, Meitian Wang6 & Jörg Standfuss1

Historically, room-temperature structure determination was succeeded by cryo-crystallography to mitigate radiation damage. Here, we demonstrate that serial millisecond crystallography at a synchrotron beamline equipped with high-viscosity injector and high frame-rate detector allows typical crystallographic experiments to be performed at room-temperature. Using a crystal scanning approach, we determine the high-resolution structure of the radiation sensitive molybdenum storage protein, demonstrate soaking of the drug colchicine into tubulin and native sulfur phasing of the human G protein-coupled adenosine receptor. Serial crystallographic data for molecular replacement already converges in 1,000–10,000 diffraction patterns, which we collected in 3 to maximally 82 minutes. Compared with serial data we collected at a free-electron laser, the synchrotron data are of slightly lower resolution, however fewer diffraction patterns are needed for de novo phasing. Overall, the data we collected by room-temperature serial crystallography are of comparable quality to cryo-crystallographic data and can be routinely collected at synchrotrons.
Serial femtosecond crystallography (SFX) at X-ray free-electron lasers (XFELs) is a unique method for the determination of challenging structures without radiation damage at room-temperature. In particular, the development of high-viscosity injectors that have markedly reduced sample consumption has opened the door for the structure determination of novel membrane proteins and protein complexes, and de novo phasing of pharmaceutical targets. Room-temperature structure determination is of increasing importance in the elucidation of protein dynamics. As opposed to cryo-crystallography, room-temperature X-ray structures reveal realistic conformational flexibility that is crucial for protein function. In addition, room-temperature crystallography enables time-resolved studies that reveal protein dynamics and enzyme catalysis with atomic detail. However, since the beamtime available at XFELs will remain scarce for the foreseeable future, routine room-temperature structure determination at these sources is difficult to achieve. A promising solution is to adapt the methodology developed for macromolecular crystallography at XFELs to synchrotron sources, where radiation damage cannot be outrun but where the radiation dose per crystal can be reduced by using many crystals.

Here, the term serial crystallography is commonly used in a broader sense, covering loop- or sandwich-scanning techniques as well as the combination of a rotation series with the use of dozens to hundreds of individual crystals. The most direct approach is serial millisecond crystallography (SMX), which utilizes the same high-viscosity injectors successful at XFELs to distribute the radiation dose over thousands of crystals to determine room-temperature structures with minimum radiation-damage.

In pioneering SMX experiments, data were collected continuously (shutterless), exposing one crystal per image with a 10 × 30 μm² beam on a PILATUS 6 M detector with a low frame rate of 10 Hz. During the revision of this manuscript a study performed at advanced photon source was published, describing data collection at 10 Hz with a PILATUS3 6 M detector. In another approach, data were collected with a very small 2 × 3 μm² beam, using a shutter to allow short exposure times between 10 and 50 ms and a Rayonix detector with a maximal frame rate of 17 Hz. In these approaches data collection times were very long and took up to 5 days in case of the membrane protein bacteriorhodopsin. Here, we demonstrate how the combination of a high-viscosity injector with the high frame rates of an EIGER 16 M detector in a 'crystal scanning' approach allows SMX to be routinely applied to the three most common crystallographic techniques: high-resolution structure determination, ligand soaking and de novo phasing (Fig. 1).

To demonstrate the general applicability of SMX, we have chosen three real-life targets: the radiation sensitive molybdenum storage protein (MOSTO), the αβ-tubulin-darpin complex (TD1) soaked with the drug colchicine and the thermostabilised adenosine A2A G protein-coupled receptor (A2AR). All individual experiments including native single-wavelength

---

**Fig. 1** SMX setup, features and overview of results. The schematic of the SMX experimental setup at the Swiss Light Source shows the high-viscosity injector set to extrude a 50 μm diameter stream of crystal laden LCP (or other viscous carrier) vertically through the interaction region with the X-ray beam and into a catcher for safe and clean disposal of material. The micro-focused beam is centered on the extruded medium ~50 μm below the nozzle. The jet speed determines exposure time and is calculated from the nozzle diameter and sample flow rate (which is calculated from the pump flow rate). The extrusion is controlled with a helium sheath, which can be adjusted with the pressure regulator to maximize jet stability. A typical beam size in our experiments is 20 μm horizontal and 5 μm vertical with respect to the extruded medium. This size allows for a large interaction volume and maximum flux. Data are sent to the beamline data collection servers where real-time peak finding is fed back to the user.
anomalous dispersion (SAD) phasing can be performed in less than a single 8-h synchrotron shift, with data convergence of native data sets already being achieved after maximally 82 min.

Results

Optimizing data collection parameters and radiation damage.

Lipidic cubic phases (LCP) are an ideal delivery medium for serial crystallography because of the simple sample preparation, the compatibility with both soluble and membrane proteins and an easily controllable extrusion by high-visibility injectors. Similar to the search for suitable cryo-protectants in classical cryo-crystallography, different crystals may tolerate the embedding process differently. To solve this problem, a variety of different viscous media have been described. In our hands, all crystals tolerated LCP very well and hence we only used this well-established medium in the present study. In our SMX setup (Fig. 1) we use nozzles between 50 and 75 μm, which is sufficient to avoid clogging by crystals, as larger crystals were usually crushed during mixing with a standard syringe coupler, without indications for loss of diffraction quality in our cases.

In optimization experiments with lysozyme crystals of an average size of ~10 × 10 × 10–20 μm³, we determined optimal beam parameters for data collection (see Supplementary Table 1). To test if a smaller beam is beneficial for data quality, we collected three data sets with three different foci of 20 × 5 μm², 10 × 5 μm² and 5 × 5 μm² with similar overall flux and flux densities of 0.4 × 10¹⁰ ph/ (s × μm²), 0.8 × 10¹⁰ ph/ (s × μm²) and 1.5 × 10¹⁰ ph/ (s × μm²), respectively. Using a data set collected in 20 min with the 5 × 5 μm² beam allowed us to determine the structure of lysozyme with data extending to 1.5 Å resolution (see Table 1 and Supplementary Fig. 1). The smaller 5 × 5 μm² beam resulted in the best data characteristics, with slightly higher overall signal-to-noise ratio and better CC* in the high-resolution shell (see Supplementary Table 1). However, the effect of the smaller beam was not very pronounced, resulting only in an 18% increase in overall I/σ and a 0.09 Å increase in resolution (based on a CC* >0.5 criterion) as compared with the largest beam tested, despite the flux density being reduced to a quarter in this experiment (see Supplementary Table 1). Based on these results we decided to use the 20 × 5 μm² beam for most of the study as crystals from our other three proteins were of comparable size and it allowed us to use the full intensity beam. In some runs, we varied the longer side of the beam to adjust hit rate and signal to noise depending on the specific characteristics of the sample. The wider beam perpendicular to the flow of the LCP extrusion allowed scanning a larger volume of the crystal-carrying medium. Generally, this led to increased hit rates and thus balanced sample consumption, data collection speed and signal-to-noise ratio.

The speed of extrusion, the average crystal size, the beam shape and the applied flux determine the average radiation dose per crystal (see online methods and Supplementary Data 1). Hence, we adjusted the extrusion speed to 250 μm/s for most cases while using the maximum available photon flux (see Table 1) to balance strong diffraction vs. radiation dose and shorter time for data collection.

Data were collected continuously using the EIGER 16 M detector at 50 Hz. Instead of recording only one image per crystal the low jet speed, a thin beam and a frame rate of 50 Hz allowed each crystal being traversed by one beam width of 5 μm in one frame to avoid recording of multiple frames from the same crystal region. Using eight runs from our A2A receptor 6 keV data with a hit rate of ~40%, we compared unbinned 50 Hz data with twofold binned 25 Hz data and fivefold binned 10 Hz data (see Supplementary Table 2). As in the original 50 Hz data, diffraction from the average sized crystal spans about six images (see Supplementary Fig. 2), the 10 Hz binned data nicely demonstrate the case where on average one diffraction pattern is collected per crystal. Comparison of data sets revealed that scanning a crystal with higher frame rate data collection for SMX is superior to lower frame rate data collection. This was supported not only by anomalous peak heights (see Fig. 2) but also by other data statistics such as signal-to-noise ratio in the higher resolution shells (see Supplementary Table 2). Using the anomalous signal to assess data quality, we show that the improved statistics are not owing to multiple patterns collected from a single crystal, which may bias statistics based on CC* (see Fig. 2). Thus, there are clear advantages of the latest detector generation with high frame rate and negligible dead time over slower detectors for SMX data acquisition.

Using the described data collection scheme, we observed 1–8 diffraction patterns per crystal (see Supplementary Fig. 2) corresponding to overall doses between 30 and 480 kGy per entire crystal, but maximally 60 kGy for each exposed crystal region (Table 1) as calculated with RADDOS3D. Such dose generally does not lead to a strong loss of diffraction. Consequently, we did not observe site specific damage features such as broken disulfide bonds or decarboxylation of residues in all four case studies.

Sample optimization.

To examine the time needed to adapt protein crystallization and crystal preparation for SMX, we selected the soluble MOSTO complex whose crystallization was previously optimized for classical crystallography. Room-temperature structure determination of this complex is challenging because of the unusually high metal content, which makes it susceptible to radiation damage. After a simple optimization step, we were able to produce numerous microcrystals suitable for SMX in hanging drops (Supplementary Fig. 3). After embedding these crystals into LCP, we collected 68,788 indexable diffraction patterns in ~6 h with much less time necessary to collect a fully converged data set (see Fig. 4). The data extended to 1.7 Å resolution, and the resulting electron density map is of excellent quality (see Fig. 3a and b). Clearly, high-resolution structure determination at room-temperature by SMX can be fast and straightforward.

Ligand soaking.

Next, we explored the suitability of SMX for the method that is inarguably responsible for most of the structures deposited in public or corporate databases: ligand soaking. We used microcrystals of the TDI complex, and soaked them with the microtubule-destabilizing agent colchicine—a common medication against gout. The colchicine-binding site is the target for a new class of chemotherapeutic agents. After a simple sample preparation (see materials and methods) and an incubation of 2 h, we collected 62,245 indexable diffraction patterns in ~9 h (see Table 1). The diffraction data extended to 2.1 Å (Table 1), and the ligand was clearly resolved and bound with full occupancy (see Fig. 3d). Colchicine showed the same binding mode as described previously. For comparison, we further collected a ligand free data set containing 66,271 images in 8 h. Useful data extended up to 2.1 Å resolution resulting in a high-quality electron density (see Fig. 3e). Despite using a much lower dose per crystal, the resolution of our TDIapo SMX structure with 2.1 Å is only slightly lower than the 1.8 Å structure of TDIapo obtained by cryo-crystallography from a larger crystal (see Table 1 and Supplementary Table 4).
Table 1 Data statistics

| Data set        | Lysozyme | A2AR SMX | A2AR SMX | A2AR 19 h 6 keV | A2AR SFX | TD1smx | TD1CAl | MOSTO |
|-----------------|----------|----------|----------|----------------|----------|--------|--------|--------|
| X-ray energy (keV) | 12.4    | 12.4    | 12.4    | 24.84          | 24.84    | 13.0   | 14.3   | 12.4   |
| Measurement time (h) | 90      | 200      | 90      | 31,065,416     | 12,000   | 2,000  | 12,000 | 0.7    |
| Nozzle size (μm) | 70       | 200      | 70      | 30,837         | 30,837   | 30,837 | 30,837 | 2.0    |
| Beam size (μm) | 4.0       | 200      | 4.0     | 30,759         | 30,759   | 30,759 | 30,759 | 2.0    |
| Phase (ph/°) | 3.9 x 10¹¹ | 1.5 x 10⁻¹² | 3.9 x 10¹¹ | 1.5 x 10⁻¹² | 4 x 10¹⁰ | 1.5 x 10⁻¹² | 3.9 x 10¹¹ | 1.5 x 10⁻¹² |
| Frame rate | 50 Hz    | 50 Hz    | 50 Hz    | 10 Hz          | 50 Hz    | 120 Hz | 50 Hz  | 50 Hz  |
| Crystal size (μm³) | 15 x 10 x 10 | 30 x 10 x 10 | 30 x 10 x 10 | 30 x 30 x 30 | 30 x 30 x 30 | 30 x 30 x 30 | 30 x 30 x 30 | 30 x 30 x 30 |
| Dose per pixel (kGy) | 30      | 30      | 30      | 6,000,000      | 6,000    | 6,000  | 6,000  | 6,000  |
| Oscillation range | -       | -       | 1.0     | 4.1            | 2.0      | -      | -      | -      |
| Jet speed (μm/s) | 250      | 250      | 250      | 250            | 250      | 250    | 250    | 250    |
| Space group | P2121    | P2121    | P2121    | P2121          | P2121    | P2121  | P2121  | P2121  |
| Unit cell (a, b, c in Å) | 78.6, 78.6, 38.9 | 78.6, 78.6, 38.9 | 78.6, 78.6, 38.9 | 78.6, 78.6, 38.9 | 78.6, 78.6, 38.9 | 78.6, 78.6, 38.9 | 78.6, 78.6, 38.9 | 78.6, 78.6, 38.9 |
| Beam size (μm) | 20 × 5  | 20 × 5  | 20 × 5  | 20 × 5  | 20 × 5  | 20 × 5  | 20 × 5  | 20 × 5  |
| Patterns indexed | 27,000 | 27,000 | 27,000 | 27,000 | 27,000 | 27,000 | 27,000 | 27,000 |
| Patterns indexed (%) | 46.5 | 46.5 | 46.5 | 46.5 | 46.5 | 46.5 | 46.5 | 46.5 |
| Crystal size (μm³) | 13,000 | 13,000 | 13,000 | 13,000 | 13,000 | 13,000 | 13,000 | 13,000 |
| Collected images | 35,000 | 35,000 | 35,000 | 35,000 | 35,000 | 35,000 | 35,000 | 35,000 |
| X-ray energy (keV) | 12.4 | 12.4 | 12.4 | 12.4 | 12.4 | 12.4 | 12.4 | 12.4 |
| Beam size (μm) | 4.0 | 4.0 | 4.0 | 4.0 | 4.0 | 4.0 | 4.0 | 4.0 |
| Phase (ph/°) | 3.9 x 10¹¹ | 1.5 x 10⁻¹² | 3.9 x 10¹¹ | 1.5 x 10⁻¹² | 4 x 10¹⁰ | 1.5 x 10⁻¹² | 3.9 x 10¹¹ | 1.5 x 10⁻¹² |
| Frame rate | 50 Hz | 50 Hz | 50 Hz | 10 Hz | 50 Hz | 120 Hz | 50 Hz | 50 Hz |
| Crystal size (μm³) | 15 x 10 x 10 | 30 x 10 x 10 | 30 x 10 x 10 | 30 x 30 x 30 | 30 x 30 x 30 | 30 x 30 x 30 | 30 x 30 x 30 | 30 x 30 x 30 |
| Dose per pixel (kGy) | 30 | 30 | 30 | 6,000,000 | 6,000 | 6,000 | 6,000 | 6,000 |

Data convergence. In order to assess how much data are minimally necessary to obtain a data set that is converged, we generated simulated annealing (SA) ligand omit maps and plotted their cross correlation to the ligand density calculated from the model vs. the number of images included in a data set (See Fig. 4). These plots illustrate that convergence does depend mostly on multiplicity, but also on other factors as TD1 (P2₁), A2AR SMX (C222₁) and MOSTO (P6₃22) converge at ~10,000, 1,000 and 2,000 indexed patterns corresponding to multiplicity of 21.3, 15.6 and 28.3, respectively. Based on the average indexing rate, data collection times for these converged data sets were 82, 3 and 10 min for TD1, A2AR SMX and MOSTO, respectively. However, ligands already reach 90% of their final cross correlation to the generated SA omit maps after 4,000, 900 and 900 indexed patterns (indicated by dotted lines in the plot) corresponding to 35, 3 and 5 min of data collection for TD1, A2AR SMX and MOSTO, respectively. Whether a ligand is bound can thus be already determined with relatively few images and after a short time of data collection. Interestingly, the SMX and SFX data for A2AR converge following a similar trajectory although we observe a slightly higher final correlation with the ligand for data collected at the XFEL likely because of the higher overall resolution of these data.

Native SAD phasing using SMX. To demonstrate the suitability of SMX for de novo phasing, we choose the human A2AR receptor, a G protein-coupled receptor. Although members of this class of intra-membrane receptors are common targets for pharmacological intervention, they are known to be difficult to crystallize. Furthermore, crystals of the A2AR receptor decay quickly in room-temperature experiments, making these crystals ideal to demonstrate the suitability of SMX for difficult targets.

We collected a large data set of the A2AR receptor at a photon energy of 6 keV where the anomalous signal is prominent, and X-ray absorption effects can still be neglected. Structure solution was attempted in parallel to data collection and already succeeded during 19 h of beamtime (Fig. 6). SHELXD identified 18 sites with CCₐₚₙ of 25.5 and CCₜₐₚₚₚ of 11.9, which were clearly separated from non-solutions (Fig. 5b). SHELXE phasing, density modification and chain tracing resulted in a good separation of hands after 10 cycles (Fig. 5c). Automatic model building of the A2AR receptor on the 6 keV data alone improved initial phases, resulting in a map of sufficient quality for manual improvement with a model map CC of 60% to the final model.

Criteria for the identification of a diffraction hit were chosen conservatively, and by using different hit finding criteria as well as peaks identified by CHEETAH for indexing we could retrospectively solve the structure with data collected in 5 h (Supplementary Table 3, Fig. 6d and methods).

SHELXD identified 14 sites with CCₜₐₚₚₚ of 25.5 and CCₜₙₐₜₚₚ of 9.3 and clearly separated solutions from non-solutions (Fig. 5e). It was necessary to perform 100 cycles of SHELXE phasing, density modification and chain tracing to obtain a good separation of hands (Fig. 5f). Automatic model building of the A2AR receptor on the 2.7 Å data collected at 6 keV alone improved initial phases, resulting in a map of sufficient quality for manual improvement of the model with a model map CC of 55% to the final model. Overall, native-SAD phasing was possible using only ~10 μL LCP containing 12 mg ml⁻¹ A2AR receptor. The high efficiency of high-visibility injectors together with our high frame rate serial-scanning approach, thus allows to target proteins for which only such submilligram quantities are readily available, even in the case of a challenging membrane protein.
We demonstrate this approach using microcrystals of the human A$_{2A}$ receptor. Distributing much larger doses over thousands of crystals is especially true for small crystals from pharmaceutically interesting membrane proteins such as G protein-coupled receptors. This is always limited by the maximum attainable resolution, and this is especially true for small crystals from pharmaceutically interesting membrane proteins such as G protein-coupled receptors. To gain detailed understanding of protein functionality, we achieve time resolutions in the millisecond range and beyond. Precisely timed and short-pulsed lasers are able to increase the shot noise ratio to up to 1.7 Å resolution for the A$_{2A}$R even with only a very short data collection at LCLS.

Our current SMX method combines the strengths of our previous SMX approach that was demonstrated on Bacteriorhodopsin, namely short exposure times to mitigate crystal motions and a small beam to improve the signal-to-noise ratio, with the benefits of continuous shutterless data collection and a broader beam to overcome the problem of slow data acquisition. As opposed to another shutterless approach demonstrated on the PILATUS detector with lysozyme as a model system, we collected multiple high frame rate-patterns instead of a single diffraction pattern per crystal, used a much smaller beam and the opposite beam geometry (20 × 5 μm$^2$ instead of 10 × 30 μm$^2$). Using 50 Hz frame rates as made possible by high frame rate detectors such as EIGER and PILATUS3, we were able to counteract the negative impact of crystal movement during exposure of each individual frame as realized by using the shutter in our previous work. High frame rate data collection also avoids empty exposure of indexable frames after crystals already passed the beam. This leads to an increased signal-to-noise ratio by reducing the recorded background (see Supplementary Fig. 4). In our experiments, the jet speed and detector readout were ideally configured such that although one detector frame was recorded, the extruded medium (and with it the crystals) moved one beam width. This allowed for collection of multiple diffraction patterns per crystal, without repeatedly exposing the same portion of the crystal. In case of the A$_{2A}$R, we used a 5 μm-thin beam and crystals traversed 5 μm during collection of one pattern (see Table 1 and Supplementary Data 1). This setup improved data collection statistics significantly. The improvement of data statistics (see Supplementary Table 2) by collecting multiple diffraction patterns per crystal is supported by dramatically increased anomalous signals (see Fig. 2), excluding bias as a source of statistics improvement.

At the same time, keeping the broadest dimension of the beam perpendicular to the direction of crystal travel minimized sample consumption and improved hit rates in our experiments. In the case of MOSTO, where we initially suffered from low hit rates, we used a very wide but thin 40 × 5 μm$^2$ beam. In combination with a larger nozzle size, a larger volume could be “scanned”, allowing to collect a large data set within 6 h of beamtime while still obtaining data to high resolution.

We call this approach “crystal scanning”, because it balances the hit rate and the signal-to-noise ratio by scanning across the extruded viscous carrier medium with a wide but thin beam (see Supplementary Fig. 4). In addition, it maximizes the amount of unique diffraction patterns by scanning crystals with a high frame rate that is matched to the crystal movement speed and the beam dimension in the direction of crystal movement. The fast readout of the EIGER detector and its duty cycle better than 99% are crucial for this strategy. At the frame rates we used, also PILATUS3 detectors with a duty cycle of ~95% at 50 Hz would appear to be suitable for the strategy presented here, although the higher deadtime, which is not evenly spread-out over the individual frames, may lead to a slight decay in data quality. Our data collection scheme is very well suited for future time-resolved experiments, because it can combine the ultra-high frame rates of the EIGER detectors (up to 3,000 Hz) with laser excitation to achieve time resolutions in the millisecond range and beyond. Other approaches like choppers, single bunch operation or precisely timed and short-pulsed lasers are able to increase achievable time resolutions even further, but the time resolution obtained simply by detector readout speeds will allow to collect whole-time series in a single experiment. To allow processing of high frame-rate data, several images can be merged together, resulting in lower frame rates but stronger diffraction patterns to...
allow indexing, integration and background subtraction. Leaving out individual images from the merged images and using the Hadamard transform\(^3\) to deconvolute them, may allow time resolved analysis far beyond indexable individual diffraction patterns at these frame rates. With the advent of single shot serial crystallography, data processing progressed significantly by improving detector geometries\(^3\), as well as including prediction- and post-refinement\(^3\). We expect these developments to continue at their fast pace, allowing challenging time resolved experiments to be even more feasible.

Using this strategy, we were able to solve the structure of a human membrane protein pharmaceutical target, the \(A_{2A}\) receptor, by native-SAD phasing, which demonstrates that even high quality data necessary for native-SAD phasing are obtainable in 5 h of beamtime with injector-based SMX. This was possible on very small crystals (30 \(\times\) 30 \(\times\) 5 \(\mu\)m\(^3\)) and at room-temperature. Based on reports of gadolinium based\(^4\), \(^5\) and native sulfur/chloride SAD phasing at XFEL sources\(^6\), phasing exploiting other anomalous scatterers than sulfur should be even easier to implement. At the synchrotron, both the data collection time and the number of indexed images were lower than those required for phasing of the \(A_{2A}\) receptor at the LCLS\(^7\). Reasons for this unexpected finding may include (i) the use of larger crystals in our experiment at the Swiss Light Source (SLS), (ii) lower partiality of individual reflections due to slight rotation of crystals during exposure at the synchrotron, (iii) shot to shot
variation in the spectrum and intensity of XFEL pulses, as well as (iv) the higher dynamic range of the single photon counting EIGER 16 M detector vs. the integrating CSPAD.

Interestingly, we observe unexcepted differences between cryo- and room-temperature structures of tubulin. Visualizing protein backbone mobility by ensemble models (see Methods), we observed overall similar patterns of mobility in different parts of the structures, with overall increased mobility of loop regions for room-temperature structures, as expected. However, for the colchicine unbound structure, the T7 loop of β-tubulin that shields ligands bound to the colchicine site from the solvent and the α-tubulin T5 loop, which is also involved in ligand binding to the colchicine site, show an opposing mobility pattern in both structures. The T7 loop appears flexible and the T5 loop more rigid in the room-temperature structure, whereas the cryo-structure seems to exhibit a more mobile T5 loop and a more rigid T7 loop (Supplementary Fig. 5 and Supplementary Table 4). This observation indicates that apart from allowing time resolved experiments, room-temperature structures may provide additional information that could help to understand ligand binding and protein dynamics.

When comparing cryo-crystallography, SMX and SFX, we observed an immediate resolution advantage of XFEL over synchrotron crystallography (Table 1 and Fig. 6). Clearly, one advantage of XFELs is the ability to deliver and diffract photon amounts that will destroy any crystal in a synchrotron experiment. This advantage is most relevant when small, weakly diffracting crystals are studied or in cases where short pulses are needed to collect time-resolved data. Nevertheless, XFEL beamtime will not be available for routine experiments in the foreseeable future and synchrotron based SMX taps into the large pool of available synchrotron beamlines. The upcoming diffraction-limited storage rings and specialized beamlines will improve on flux densities and instrumentation. They will further reduce the time needed for successful SMX experiments and close the already small resolution gap between cryo-crystallography and practically damage free room-temperature structure determination by SMX.

In conclusion, we show that SMX at synchrotrons allows fast, straightforward structure determination at room-temperature for large soluble macromolecular complexes as well as membrane proteins. Using SMX, a much larger dose can be distributed over many crystals, resulting in higher resolution structures with less-radiation damage compared with classical room-temperature methods. Collecting data using modern fast frame rate detectors produces results of excellent quality, making even native SAD phasing possible in less than a single 8-h synchrotron shift. Using ultra high frame rate detectors and next-generation diffraction-limited sources, time-resolved measurements in the micro- and perhaps nanosecond range may become possible at synchrotrons. Given the comparatively simple sample preparation, data collection, data processing and its great potential for automation, we believe that SMX is the method of choice for room-temperature structure determination and fragment screening approaches.

**Methods**

**Ligand preparation.** Chicken egg white lysozyme from Alfa Aesar (60701) was dissolved at a concentration of 50 mg ml⁻¹ in 100 mM Na Acetate pH 3.0. Of this lysozyme solution, 500 µl were mixed with 500 µl of precipitant, 19.04% NaCl, 5.44% PEG 8,000, 68 mM Na Acetate pH 3.0 in an Eppendorf tube and incubated overnight at 20 °C. Crystals grown in batch were spun down at room-temperature with 2,000 g for 1 min. Almost all supernatant was removed and crystals were re-suspended using a 200 µl pipette. Crystals were then pipetted into the back of a Hampton syringe and mixed with ~70% monoolein through a 400 µm LCP coupler at room temperature until the cubic phase was formed.

**TD1 sample preparation.** Synthetic DNA encoding for the DARPin D1 was cloned into a pET-based expression vector (see Supplementary Table 3) and expressed in *E. coli* BL21 DE3 (purchased from Stratagen). The DARPin protein was purified by Ni-affinity chromatography and size exclusion chromatography. Bovine brain tubulin was purchased from the Centro de Investigaciones Biológicas (Microtubule Stabilizing Agents Group), CSIC, Madrid, Spain. The TD1 was formed by mixing tubulin and Darpin 1 in a 1:1.1 ratio and concentrated to 20 mg ml⁻¹ before crystallization. TD1 was crystallized by hanging drop vapor diffusion in 20 °C in 20–24% PEG 3350, 0.2 M Ammonium Sulfate, 0.1 M Bis-Tris Methane pH 5.5, mixing 2 µl of the complex with 2 µl of precipitant solution. On the next day, a few TD1 crystals were observed in each drop. To reach a larger number of crystals 1 µl water was then added and the drop mixed a few times. After another night incubation at 20 °C, each crystallization drop contained numerous TD1 microcrystals.

SMX sample preparation. Individual hanging drops were pooled by pipetting into an Eppendorf tube, the glass covers were carefully washed with mother liquor. Crystals were then spun down at room-temperature with 2,000 g for 1 min. Almost all supernatant was removed and crystals were re-suspended. Crystals were then pipetted into the back of a Hampton syringe and mixed with ~70% monoolein through a 400 µm coupler at room temperature until the cubic phase formed.
Ligand soaking. A mixture of 6 µl 100 mM colchicine in 100% DMSO, 6 µl of TD1 crystallization buffer (22% PEG 3350, 0.2 M ammonium sulfate, 0.1 M bis-tris methane pH 5.3) and 18 µl Monolein was prepared to a final concentration of 20 mM colchicine and 20% DMSO. The protein crystals were prepared as described above. A total of 5 µl ligand mixture was then mixed 100 times with 22 µl crystals dispersed in LCP using two Hamilton syringes and a standard 400 µm coupler, resulting in final concentrations of 3.7 mM colchicine and 3.7% DMSO. Crystals were incubated for 2 h before data collection to allow colchicine binding.

MOSTO sample preparation. MOSTO was heterologously produced in E. coli BL21 (DE3) cells with a combination of vectors harboring different molecular

A2AR receptor sample preparation. The A2AR receptor construct used in this study contains eight thermostabilizing mutations, A54L, T88A, R107A, K122A, L202A, L235A, V239A and S277A, and another mutation N154A to remove a glycosylation site. A FLAG tag has been added to the N-terminus and after K315, residues 25 to 34 containing the catalytic coupler at room temperature until the cubic phase was formed. The A2A receptor construct used in this study was expressed in the baculovirus system and purified using Ni-NTA super resin (Qiagen) in 40 mM Tris–HCl pH 7.6, 1 mM ethylenediaminetetraacetic acid supplemented with ethylenediaminetetraacetic acid-free protease inhibitor cocktail (Roche) and stored at −80 °C until use. To purify A2AR, membranes were thawed and solubilized in 1.5% (w/v) n-decyl-β-D-maltopyranoside (DM) for 1 h at 4 °C in the same buffer supplemented with 5 mM ZM241385 and 2% (v/v) 1,6-hexanediol. The applied gas pressures typically ranged between two and four bars. Pressures up to 40 bars were used to clean the nozzle from sample that was not extruded properly. Extruded sample was caught in a sample catcher connected to a membrane pump.

Data collection. Data were collected with a continuous beam using an EIGER 16 M detector at frame rates indicated in Supplementary Table 1. The widest dimension of the beam was horizontal to the ground level (see Supplementary Fig. 1). Data were collected in separate runs containing maximally 100,000 images each.

Data processing. Most data were processed using CrystFEL version 0.6.230 directly on the acquired images. Images were pre-selected by our live monitoring software that implemented Cheetah31 peak finding algorithm 8 for peak detection. Parameters used during the beamtime were: signal to noise ratio = 20, minimum number of pixels per peak = 3 and minimum number of spots 10. For processing optimization, data with minimum SNR = 5 and minimum number of peaks = 20 were used. The final processing with optimized parameters was performed on all images. In this approach, all peaks for indexing were identified using the Zaeferer algorithm as implemented in CrystFEL. We choose this approach for simplicity and to avoid unnecessary data duplication.

For 10 runs of the A2AR receptor 6 keV data (corresponding to 5 h of data collection time) we used the EIGER specific implementation of Cheetah31 developed by Takani Nakane (https://github.com/biochem-fan/cheetah/tree/eiger) and subsequently used the Cheetah found peaks for indexing.
were collected using A2AR crystals prepared in the same way to facilitate a direct comparison of the techniques. The SMX density is well defined, but the overall resolution (2.1 Å) is lower than in SFX (1.7 Å) and cryo-crystallography (1.95 Å). The substructure was identified using SHELXD, using data up to 2.6 Å resolution and $E$-values to 1.3 in a search for 12 sites with expanding two sites into disulfides. Phasing was possible entirely on anomalous data, including data up to 2.6 Å resolution by performing 10 cycles of SHELXE with auto tracing and search for alpha helices enabled.

Phasing with the minimal amount of collected data was achieved by processing the images as described above. Data preparation, sub structure identification and experimental phasing were performed using the SHELX pipe line with the graphical user interface hkl2map to data preparation using SHELX. The substructure was identified using SHELXD, using data up to 2.73 Å resolution and $E$-values to 1.2 in a search for 11 sites without expanding into disulfides. Phasing was possible entirely on anomalous data collected up to 2.67 Å resolution by performing 100 cycles of SHELXE with auto tracing and search for alpha helices enabled (see Supplementary Fig. 5).

Model building of the A2A receptor. For the minimal data set (5 h data collection), automatic model building of the A2AR receptor was performed using phases of SHELXE and the 6 keV A2AR data directly resulting in 305 (68%) residues built and 209 residues sequenced after BUCCANEER, resulting in a map of sufficient quality for manual improvement of the model with a model map CC of 55% to the final model. For the high-quality data set (19 h data collection), automatic model building of the A2AR receptor in Buccaneer was performed using phases of SHELXE and the 6 keV A2AR data directly resulting in 328 (73%) residues built and 142 residues sequenced after BUCCANEER, resulting in a map of sufficient quality for manual improvement of the model with a model map CC of 60% to the final model.

Refinement and model building. Lysozyme: the structure was solved by molecular replacement with Molrep using PDB entry 3LZT as model. The structure was then refined in refmac5 and finally in PHENIX, sidechains and water molecules were adjusted using Coot. The structure was refined iteratively by refmac5 directly using PDB entry 4PET and manual adjustment in Coot. Final refinement cycles were carried out using autoBUSTER (Global Phasing Ltd.). TD1 (Apo and Col): The structure was solved using PDB entry 4DRX and refined iteratively using refmac5 and iterative model building in Coot. The last cycles of refinement were carried out using autoBUSTER (Global Phasing Ltd.), in case of TDCol the colchicine ligand was refined using quantum mechanical description of the ligand. A2AR (SMX): We extended the phases obtained from the high-quality 6 keV data set on the 12.4 keV room temperature SMX data and used automatic model building in PHENIX followed by iterative refinement in refmac5 and manual model building in Coot. In the final stages, refinement was carried out using PHENIX. A2AR (SFX): The refined model from A2AR (SFX) was directly refined against the SFX data using PHENIX. A2AR (cryo): The structure was solved using PDB entry 5IU4 and subsequently refined using PHENIX.

Simulated annealing omit maps/model map cross correlations. The final processed streams of serial data were scaled and merged using a partiality cutoff of 0.1 (see Supplementary Table 1). For the ensemble refinement of the cryo-structure without colchicine we used data collected on a single TD1 crystal harvested before addition of water to produce micro-crystals that was cryo-protected using 25% PEG 400, 8% photons per pulse. The collected data were preprocessed with Cheetah in order to select frames containing diffraction patterns. Selected images were indexed and integrated and merged using CrystFEL version 0.6.2.

Calculation of radiation doses. The radiation dose (diffraction weighted dose) was calculated using RADDose32 with the parameters given in Supplementary Table 1, assuming no rotation of the crystals and choosing the smallest crystal dimension collinear to the direction of the beam, hence the values represent maximum doses. Crystal movement was taken into account by matching the exposure time to the amount of time the crystals traverse one full height of the beam (0.02 s in case of 250 µm/s jet speed and 5 µm beam height, see Supplementary Data 1). The dose applied owing to the Gaussian beam shape to the region above the beam, that was fully exposed in each consecutive frame was neglected in the dose calculations, since it is small compared with the dose received in the full exposure.

Experimental phasing of the A2AR receptor. Initial phasing was achieved in parallel to data acquisition. Data processed with CrystFEL was converted into XSCALE format and passed through XSCALE. Data preparation, sub structure identification and experimental phasing were performed using the SHELX pipe line with the graphical user interface hkl2map to data preparation using SHELX. The substructure was identified using SHELXD, using data up to 2.65 Å resolution and $E$-values to 1.2 in a search for 11 sites without expanding into disulfides. Phasing was possible entirely on anomalous data collected up to 2.67 Å resolution by performing 100 cycles of SHELXE with auto tracing and search for alpha helices enabled (see Supplementary Fig. 5).

Model building of the A2AR receptor. For the minimal data set (5 h data collection), automatic model building of the A2AR receptor was performed using phases of SHELXE and the 6 keV A2AR data directly resulting in 305 (68%) residues built and 209 residues sequenced after BUCCANEER, resulting in a map of sufficient quality for manual improvement of the model with a model map CC of 55% to the final model. For the high-quality data set (19 h data collection), automatic model building of the A2AR receptor in Buccaneer was performed using phases of SHELXE and the 6 keV A2AR data directly resulting in 328 (73%) residues built and 142 residues sequenced after BUCCANEER, resulting in a map of sufficient quality for manual improvement of the model with a model map CC of 60% to the final model.

Refinement and model building. Lysozyme: the structure was solved by molecular replacement with Molrep using PDB entry 3LZT as model. The structure was then refined in refmac5 and finally in PHENIX, sidechains and water molecules were adjusted using Coot. The structure was refined iteratively by refmac5 directly using PDB entry 4PET and manual adjustment in Coot. Final refinement cycles were carried out using autoBUSTER (Global Phasing Ltd.). TD1 (Apo and Col): The structure was solved using PDB entry 4DRX and refined iteratively using refmac5 and iterative model building in Coot. The last cycles of refinement were carried out using autoBUSTER (Global Phasing Ltd.), in case of TDCol the colchicine ligand was refined using quantum mechanical description of the ligand. A2AR (SMX): We extended the phases obtained from the high-quality 6 keV data set on the 12.4 keV room temperature SMX data and used automatic model building in PHENIX followed by iterative refinement in refmac5 and manual model building in Coot. In the final stages, refinement was carried out using PHENIX. A2AR (SFX): The refined model from A2AR (SFX) was directly refined against the SFX data using PHENIX. A2AR (cryo): The structure was solved using PDB entry 5IU4 and subsequently refined using PHENIX.

Simulated annealing omit maps/model map cross correlations. The final processed streams of serial data were scaled and merged using a partiality cutoff of 0.1 (see Supplementary Table 1). For the ensemble refinement of the cryo-structure without colchicine we used data collected on a single TD1 crystal harvested before addition of water to produce micro-crystals that was cryo-protected using 25% PEG 400, 8%
PEG 3350, 0.2 M Ammonium Sulfate and 0.1 M Bis-Tris Methane pH 5.5. Data were collected at Beamline X06SA using a 20 × 5 μm² beam using the Eiger 16 M detector. The structure of the refined room temperature structure TDI<sub>apo</sub> was directly refined in PHENIX against the data, with manual adjustments in Coot (see Supplementary Table 4). For the ensemble refinement of the cryo-structure with colchicine we used PDB entry SEYP, replaced the α- and β- chains with the ones from our final refined room temperature model of TDI<sub>apo</sub> but kept the DARPin, which had a slightly different sequence. In all cases crystal packing was almost identical, despite slight variations of unit cell constants.

All structures for the ensemble refinement were prepared in the same way. First, a model with no gaps regardless of observed density was generated, sidechains for which no density was observed were removed. Then a refinement over 10 cycles was carried out in PHENIX, optimizing stereochemistry and atomic displacement parameter (B-Factor) weights as well as updating waters. Sidechains were adjusted to fit the density in Coot. Afterwards TLS groups were determined automatically. Then, another refinement round with the automatically determined TLS groups was carried out, with the option ‘Automatically add hydrogens to model’ enabled, but solvent update disabled. The models generated in this way were then refined with a fraction of 0.8 of all atoms included in TLS fitting with ordered solvent update using PHENIX Ensemble refinement against data cut to 2.35 Å resolution. The resolution of the data was reduced to ensure the same amount of data for the refinement as well as data with C<sub>Gl</sub> above 30% for all refined structures.

**Data availability.** Coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 5N1M, 5N1X, 5N2M, 5N4M, 5NQT, 5N5M, 505 W and SNQU. Other data are available from the corresponding author upon reasonable request.

Received: 30 March 2017 Accepted: 14 July 2017
Published online: 14 September 2017

**References**

1. Chapman, H. N. et al. Femtosecond X-ray protein nanocrystallography. Nature 470, 73–77 (2011).
2. Suga, M. et al. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. Nature 517, 99–103 (2014).
3. Weierstall, U. et al. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. Nat. Commun. 5, 3309 (2014).
4. Zhang, H. et al. Structure of the angiotensin receptor revealed by serial femtosecond crystallography. Cell 161, 833–844 (2015).
5. Kang, Y. et al. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. Nature 523, 561–57 (2015).
6. Batyuk, A. et al. Native phasing of x-ray free-electron laser data for a G protein-coupled receptor. Sci. Adv. 2, e1600292 (2016).
7. Fraser, J. S. et al. Accessing protein conformational ensembles using room-temperature X-ray crystallography. Proc. Natl. Acad. Sci. USA 108, 16247–52 (2011).
8. Fenwick, R. B., van den Bedem, H., Fraser, J. S. & Wright, P. E. Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. Proc. Natl. Acad. Sci. USA 111, E445–54 (2014).
9. Nobl, P. et al. Lipidic cubic phase injector is a viable crystal delivery system for time-resolved serial crystallography. Nat. Commun. 7, 12314 (2016).
10. Nango, E. et al. A three-dimensional movie of structural changes in bacteriorhodopsin. Science 354, 1552–1557 (2016).
11. Suga, M. et al. Light-induced structural changes and the site of O = O bond formation in PSII caught by XFEL. Nature 543, 131–135 (2017).
12. Gati, C. et al. Serial crystallography on in vivo grown microcrystals using synchrotron radiation over 10 cycles. IUCrJ 1, 87–94 (2014).
13. Coquelle, N. et al. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. Acta Crystallogr. Sect. D Biol. Crystallogr. 71, 1184–1196 (2015).
14. Schubert, R. et al. A multicrystal diffraction data-collection approach for studying structural dynamics with millisecond temporal resolution. IUCrJ 3, 393–406 (2015).
15. Huang, C. Y. et al. In meso in situ serial X-ray crystallography of soluble and membrane proteins. Acta Crystallogr. Sect. D Biol. Crystallogr. 71, 1238–1256 (2015).
16. Huang, C. Y. et al. In meso in situ serial X-ray crystallography of soluble and membrane proteins at cryogenic temperatures. Acta Crystallogr. Sect. D, Struct. Biol. 72, 93–112 (2016).
17. Botha, S. et al. Room-temperature serial crystallography at synchrotron X-ray sources using slowly flowing free-standing high-viscosity microstreams. Acta Crystallogr. Sect. D Biol. Crystallogr. 71, 387–397 (2015).
49. Segala, E. et al. Controlling the dissociation of ligands from the adenosine A2A receptor through modulation of salt bridge strength. J. Med. Chem. 59, 6470–6479 (2016).
50. White, T. A. et al. CrystFEL: a software suite for snapshot serial crystallography. J. Appl. Crystallogr. 45, 335–341 (2012).
51. Barty, A. et al. CheetaAH: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. J. Appl. Crystallogr. 47, 1118–1131 (2014).
52. Kabash, W. XDS. Acta. Crystallogr. D. Biol. Crystallogr. 66, 125–132 (2010).
53. Liang, M. et al. The Coherent X-ray Imaging (CIX) instrument at the Linac Coherent Light Source (LCLS). J. Synchrotron. Radiat. 22, 514–519 (2015).
54. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta. Crystallogr. D. Biol. Crystallogr. 66, 479–485 (2010).
55. Pape, T. & Schneider, T. R. HKL2MAP: a graphical user interface for macromolecular phasing with SHEXL programs. J. Appl. Crystallogr. 37, 843–844 (2004).
56. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta. Crystallogr. D. Biol. Crystallogr. 62, 1002–1011 (2006).
57. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 22–25 (2010).
58. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta. Crystallogr. D. Biol. Crystallogr. 67, 355–367 (2011).
59. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. 66, 213–221 (2010).
60. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta. Crystallogr. D. Biol. Crystallogr. 60, 2126–2132 (2004).
61. Collaborative Computational Project, N. 4. The CCP4 suite: programs for protein crystallography. Acta. Crystallogr. D. Biol. Crystallogr. 50, 760–3 (1994).

Acknowledgements

We thank SwissFEL for their constant logistic and financial support. We are grateful to Professor Uwe Weierstall for advice on the high-viscosity injector. Dr. Takanori Nakane for his implementation of CHEETAH for EIGER detectors as well as for setting up a tutorial including scripts how to directly process EIGER data. We thank the staff of the CXI beamline at the SLAC National Accelerator Laboratory. LCLS is an Office of Science User Facility operated for the United States Department of Energy Office of Science by Stanford University. We thank the staff of the CXI beamline at the SLAC National Accelerator Laboratory. LCLS is an Office of Science User Facility operated for the United States Department of Energy Office of Science by Stanford University. Emsley et al. (Acta Crystallogr. D. Biol. Crystallogr. 50, 760–763, 1994).