Fast and accurate data collection for macromolecular crystallography using the JUNGFRAU detector

Filip Leonarski1, Sophie Redford1, Aldo Mozzanica1, Carlos Lopez-Cuenca1, Ezequiel Panepucci1, Karol Nass1, Dmitry Ozerov1, Laura Vera1, Vincent Olieric1, Dominik Buntschu1, Roman Schneider1, Gemma Tinti1, Erik Frojdh1, Kay Diedrichs2, Oliver Bunk1, Bernd Schmitt1* and Meitian Wang1*

The accuracy of X-ray diffraction data is directly related to how the X-ray detector records photons. Here we describe the application of a direct-detection charge-integrating pixel-array detector (JUNGFRAU) in macromolecular crystallography (MX). JUNGFRAU features a uniform response on the subpixel level, linear behavior toward high photon rates, and low-noise performance across the whole dynamic range. We demonstrate that these features allow accurate MX data to be recorded at unprecedented speed. We also demonstrate improvements over previous-generation detectors in terms of data quality, using native single-wavelength anomalous diffraction (SAD) phasing, for thaumatin, lysozyme, and aminopeptidase N. Our results suggest that the JUNGFRAU detector will substantially improve the performance of synchrotron MX beamlines and equip them for future synchrotron light sources.

M X reveals 3D structures and elucidates functions of biomolecules with atomic resolution, thereby enabling researchers to make fundamental contributions to molecular biology and structure-based drug discovery. Synchrotron radiation, together with large-format 2D detectors, has been essential to the success of modern MX. In parallel with the evolution of synchrotron sources, several generations of X-ray detectors have been developed, namely, image plates1, multiwire proportional counters2, X-ray television detectors3, charge-coupled device (CCD) detectors4, and hybrid (pixel-array) photon-counting (HPC) detectors5. Currently, most MX beamlines are equipped with HPC detectors, or are scheduled to be.

Each new generation of X-ray detector has transformed MX data-collection strategies. The traditional high-dose and coarse-phi slicing data-collection strategy adapted for CCD detectors6–10 has been replaced by a continuous, low-dose and fine-phi slicing strategy that takes full advantage of HPC detectors11–13. Very recently, the EIGER detector14–16 enabled new data-collection protocols that incorporate fast raster scanning11 and serial crystallography14.

Some of the key features of HPC detectors are very low-noise detection and a point-spread response of a single pixel, achieved by counting of an incoming photon only in the pixel where it deposits at least 50% of its energy. Thus, photon counters have negligible readout noise, which means that the accuracy of their measurements is limited by their calibration, systematic effects, and Poisson statistics.

However, there are two intrinsic effects that may lead to photons going undetected by photon-counting devices, namely, charge-sharing and pileup. Charge-sharing results in spreading of photon-induced charges into adjacent pixels when photons hit the sensor near the pixel border (‘corner effect’). In such situations, the detection (counting) of the photons strongly depends on the threshold settings. The calibration of the threshold becomes less accurate at low photon energies (≤8 keV), and a 50% threshold might not be achievable. Therefore, the effects could be detrimental in low-energy applications such as native-SAD phasing, where highly accurate measurement of intensity is needed. The effect could be mitigated to a certain degree by an increase in the pixel size (for example, 170 μm pitch in the PILATUS17) or by a charge-summing and allocation method as implemented in MEDIPIX317, but such measures reduce spatial resolution and count-rate capability.

Pileup effects occur at high photon rates as a result of the dead time in the readout electronic circuit, which needs some time to reset before the next photon can be detected. This count-rate dependence of HPC detectors leads to a nonlinear response to photon flux, necessitating a count-rate correction18. A recent development in retriggering technology19 extends the count-rate capacity of HPC detectors but does not eliminate the problem. As an added complication, the count-rate correction in its simplest form is valid only with a constant flux of photons. However, in practice no count-rate correction is applied for a changing photon rate when a sharp Bragg peak moves through the diffraction condition during a single exposure. These count-rate-related issues are usually avoided in MX measurements, which are carried out with an attenuated beam at a low rotation speed. However, the count-rate capability will become acute for the next-generation synchrotrons with higher brilliance20–22.

A challenge in detector development is finding a way to overcome the aforementioned charge-sharing and pileup effects while maintaining low-noise performance at the single-photon level and with a high dynamic range. New charge-integrating hybrid pixel detectors could meet such challenges23–25; JUNGFRAU is one example26. Initially developed for X-ray free-electron laser (XFEL) applications27, JUNGFRAU features direct detection and dynamic gain-switching technology28. Instead of counting individual photons by using a threshold, JUNGFRAU measures the total amount

1Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland. 2Fachbereich Biologie, Universität Konstanz, Konstanz, Germany. 3e-mail: bernd.schmitt@psi.ch; meitian.wang@psi.ch
of charge accumulated during the integration time, by which it entirely eliminates both the charge-sharing problem and the count-rate limitation. As shown in Fig. 1, the JUNGFRAU has three separate gains per pixel, which means that low-intensity regions of the frame benefit from high gain and single-photon sensitivity, whereas strongly diffracting Bragg peaks are accurately measured thanks to the ability to switch to low gain and extend the dynamic range to 12 million counts per second (Mcps) per pixel at 12.0 keV, limited by the current 1.1-kHz frame rate. The gain is switched automatically and independently per pixel depending on the detected charge. The result of this approach is the combination of a linear response up to much higher photon rates and noise well below the limits set by Poisson counting statistics. We illustrate this for one recorded Bragg peak from a lysozyme crystal in Fig. 1.

Here we demonstrate that JUNGFRAU maintains the advantages of HPC detectors for routine MX applications and offers considerable improvements in data acquisition speed and low-energy phasing.

**Results**

**JUNGFRAU maintains low-noise performance.** The reliable detection of high-resolution weak reflections is the foremost requirement for X-ray detectors in MX. In this aspect, HPC detectors are nearly ideal because of their very low-noise detection with single-photon sensitivity and single-pixel point-spread function. We compared JUNGFRAU’s low-noise performance with that of EIGER, a widely used HPC detector, in the most common MX application: native data collection with 12.4-keV X-rays. These two detectors are particularly suited to a comparison of photon-counting versus charge-integrating methods because they have the same pixel size, same sensor area, same hybrid nature, and similar sensor thickness. We collected two datasets with the same thaumatin crystal under identical X-ray beam conditions (Methods): one with a JUNGFRAU 1-megapixel detector (JF1M; Supplementary Fig. 1), and the other with an EIGER 1-megapixel detector (E1M). We deliberately set the dose very low (0.6 kGy per dataset) to have Poisson-statistics-limited noise for the whole resolution range.

The two datasets had almost the same quality, with similar values of equivalent reflection agreement indicator $R_{meas}$ (Fig. 2a, Supplementary Tables 1 and 2). Both detectors recorded very weak intensities down to the one-photon level at resolutions of 2 Å and higher. The half-dataset correlations ($CC_{1/2}$) mean intensities, and mean signal-to-noise ratio of unmerged reflections ($\langle I/\sigma \rangle_{unmerged}$) were marginal lower for JF1M because of its slightly reduced duty cycle and thinner Si sensor (Fig. 2b–d). After normalization for detector duty cycle (JF1M 95% versus E1M 99.7%) and sensor thickness (JF1M 320 μm versus E1M 450 μm) (Methods), the intensities and $\langle I/\sigma \rangle_{unmerged}$ values were virtually the same in the whole resolution range for both detectors (Fig. 2c,d). As the maximum duty cycle of JUNGFRAU will be improved to the 99% level in the future, and as thicker sensors may be chosen during detector construction, it is expected that the performance of JUNGFRAU will approach the excellent results of EIGER for weak diffraction.

**JUNGFRAU enables data collection with full flux.** To test JUNGFRAU for high-count-rate applications, we conducted a series of experiments with increasing flux (beam transmission: 1%, 20%, 50%, and 100%) and rotation speed (1°, 20°, 50°, and 100° s$^{-1}$) using a thaumatin crystal at 6 keV (Methods). Compared with published results of similar experiments with an E1M detector, in which...
the data quality gradually deteriorated with increased flux owing to the count-rate limit<sup>11</sup>, the four JF1M datasets were of very similar quality as judged by \( R_{\text{meas}} \) and \( \langle I/\sigma \rangle_{\text{meas}} \) values for measurements obtained with beam transmissions at 1%, 20%, 50%, and 100%. The correlation of integrated intensities of reflections measured with beam transmission (tr.) of 100% with respective intensities measured at 1% transmission (Pearson correlation coefficient of 0.98). Red and blue dots represent reflections with photon rates below and above 200 Mcps mm<sup>-2</sup>, respectively. A.U., arbitrary units. d, The correlation of the estimated photon rate extracted from the single pixel of a reflection with the highest counts between the 1% and 100% transmission datasets (Pearson correlation coefficient of 0.93). The spread of the plot comes from the fact that depending on the slicing position of a reflection, the number of photons might differ for the pixel with the highest counts. The orange line represents an ideal linear response. The black curve is the theoretical behavior of a paralyzable counter with a dead time of 280 ns, and the horizontal dashed black line marks a corresponding count-rate limit.

**JUNGFRAU improves data accuracy.** We performed a native-SAD phasing experiment to assess the quality of data obtained with the JUNGFRAU detector, because this method relies on very accurate measurements of reflection intensities to derive phases<sup>27,28</sup>. We measured a thaumatin crystal with 6-keV X-rays using the JUNGFRAU and E1M detectors (Methods). We used two settings for E1M: one with the default 50% threshold (E1M-50), and the other with a 60% threshold (E1M-60) to simulate a situation in which the lowest possible threshold is higher than 50% of the photon energy (<6 keV).

\[ \frac{\text{R}_{\text{meas}}}{\text{obs}} \text{ JF tr. 100\% (A.U.)} \]
\[ \frac{\text{R}_{\text{meas}}}{\text{mrgd}} \text{ JF tr. 1% (A.U.)} \]
\[ \frac{\langle I \rangle}{\langle \sigma \rangle} \text{ JF tr. 1\%} \times 100 \text{ (Mcps mm}^{-2}\text{)} \]

\[ \text{CC}_{\text{obs}} \]
\[ \text{CC}_{\text{week}} \]

**Fig. 3** | Comparison of measurements with different photon rates acquired with the JUNGFRAU detector. a, b, The \( R_{\text{meas}} \) and \( \langle I/\sigma \rangle_{\text{meas}} \) values for measurements obtained with beam transmissions at 1%, 20%, 50%, and 100%. e, d, The correlation of the estimated photon rate extracted from the single pixel of a reflection with the highest counts between the 1% and 100% transmission datasets (Pearson correlation coefficient of 0.93). The spread of the plot comes from the fact that depending on the slicing position of a reflection, the number of photons might differ for the pixel with the highest counts. The orange line represents an ideal linear response. The black curve is the theoretical behavior of a paralyzable counter with a dead time of 280 ns, and the horizontal dashed black line marks a corresponding count-rate limit.

**Fig. 4** | Comparison of 6-keV thaumatin crystal data measured with JF1M and E1M detectors (two threshold settings for E1M). a, The crystallographic \( R_{\text{meas}} \) as a function of resolution. b, The anomalous signal \( \langle S_{\text{anom}} \rangle \) as a function of total rotation range. The dashed magenta line represents the threshold for structure solvability. c, SHELXD substructure determination from 200 trials with 60° JF1M data. The correct solution with high \( CC_{\text{obs}} \) and high \( CC_{\text{week}} \) is indicated by a red dot. d, SHELXD substructure determination from 5,000 trials with 60° E1M-50 data.

For the direct comparison, all measurements were made at the same position of the same crystal with identical data-collection parameters (Methods). For this thaumatin crystal, the typical size of a diffraction spot was a few pixels on average and was smaller at low resolution than at high resolution because of the parallax in the diffraction geometry (Supplementary Fig. 2).

The recorded JF1M data were of high quality as evaluated by \( R_{\text{meas}} \) (Fig. 4a, Supplementary Tables 1 and 5) and \( \langle I/\sigma \rangle \) (Supplementary Fig. 3). The \( R_{\text{meas}} \) of 2.5% measured at the lowest-resolution shell reflects the excellent consistency between individual measurements. \( R_{\text{meas}} \) gradually increased with the resolution to 5% at 2.7 Å, with a characteristic bump around 6 Å due to an intensity distribution typical of most protein crystals. In contrast, the E1M-50 data were noticeably worse at low resolution, with \( R_{\text{meas}} \) of 5%. The data quality deteriorated further in the E1M-60 data. Such differences had a considerable effect on the average density in the anomalous difference Fourier map for sulfur atoms \( \langle S_{\text{anom}} \rangle \) (Fig. 4b). \( \langle S_{\text{anom}} \rangle \) is a useful metric for structure solvability in SAD phasing, and a value greater than 10σ usually indicates sufficient signal for structure solution<sup>29</sup>. We obtained \( \langle S_{\text{anom}} \rangle \) values of 10.2σ, 8.9σ, and 8.3σ for 75° with JF1M, E1M-50, and E1M-60, respectively. Two to three times more data were required to elevate \( \langle S_{\text{anom}} \rangle \) above 10σ for E1M data (Fig. 4b). Indeed, the substructure was solved with SHELXC/D<sup>25</sup> with merely 60° JF1M data (Fig. 4c), whereas the same 60° data from E1M-50 did not produce a structure solution (Fig. 4d).

To understand the origin of the discrepancy in quality between data obtained with JF1M and that obtained with E1M, we quantified the uniform response at the sub-pixel level by mapping the deviation of intensities in fractional coordinates of 1 pixel on the basis of the refined position of reflections (\( \Delta_e \)). Then we calculated an average pixel map with the normalized \( \Delta_e \) (Methods; equation (4)). In the case of JF1M the pixel map was essentially...
featureless, indicating no significant bias in intensity measurement regardless of where the reflection was located within the pixel (Fig. 5a), as expected for a charge-integrating detector. However, in the case of E1M-50 there was a systematic difference between reflections centered in the middle of a pixel and those near the corners (Fig. 5b), and the magnitude of the effect increased with the detector threshold (Fig. 5c). Because most diffraction spots of the crystal were elongated in the vertical direction (Supplementary Fig. 2), the effect was much stronger in the horizontal direction in the E1M pixel maps. It is likely that the nonuniformity in EIGER can be attributed to this corner effect, inaccuracy in threshold calibration, and count-rate corrections at low energy. To estimate the contribution of these effects to crystallographic $R_{\text{meas}}$, we introduced $R_{\text{pix}}$ as a measure of systematic errors caused by the nonuniformity across pixels by averaging out random errors (Methods; equation (5)). $R_{\text{pix}}$ values were <1% for JF1M (Fig. 5d). For E1M data, $R_{\text{pix}}$ had a resolution-dependent behavior because the detector nonuniformity was more visible for sharp low-resolution spots. It increased gradually from 3 Å toward lower resolution and became a main contributor to the higher $R_{\text{meas}}$ in the low-resolution range (Fig. 5e,f).

We also verified the sub-pixel and inter-pixel uniformity in JF1M by means of detector-shifting experiments in which we measured datasets with JF1M shifted by one-third and two-thirds of a pixel in a diagonal direction orthogonal to the beam direction. When we combined two datasets—one with and one without the JF1M shift—we achieved a data accuracy equal to that measured with the same amount of data collected with only one detector position (Fig. 5g). In similar detector-shift experiments with E1M, the data accuracy improved substantially when we combined data from two detector positions to average out the nonuniform response within and between pixels with E1M (Fig. 5h,i). This analysis confirms that JUNGFRAU has good uniform responses within pixels, which permits the measurement of reflection intensities with high accuracy even at low X-ray energies and with diffraction peak sizes similar to the size of the pixel.

**JUNGFRAU expedites experimental phasing.** Accurate measurement of reflection intensities with high incoming photon rates, made possible by JUNGFRAU, enables efficient use of the full flux provided by an undulator beamline for experimental phasing with anomalous diffraction, the success of which stringently depends on the data accuracy. We chose one of the most challenging phasing methods—native SAD—to demonstrate JUNGFRAU’s distinct advantages.

First, we demonstrated that a flash of low-energy X-rays less than 1 s in duration is sufficient for native-SAD phasing with a thaumatin...
Fast native-SAD phasing with an unattenuated beam at both 6 keV and 12.4 keV with JF1M. For each case, results of the substructure search with SHELXD are shown on the left, with the correct solutions with high CC<sub>all</sub> and high CC<sub>weak</sub> marked as red dots, and the electron density map is shown on the right. a, Thaumatin with 60° of data measured in 600 ms at 6 keV. Density map obtained after density modification and automated tracing. b, Lysozyme with 500° of data measured in 5 s at 12.4 keV. Density map obtained after density modification and automated tracing with SHELXE. c, PepN with 600° of data measured in 1 min at 6 keV. Density map obtained after density modification and automated tracing with SHELXE.

Crystal as the model system. We collected a total of 60° of data from one crystal at 6 keV with a rotation speed of 100° s<sup>-1</sup>. The entire exposure lasted 0.6 s. With these data, all sulfurs were identified readily with SHELXD<sup>30</sup>, and the electron density map is shown in Fig. 6a. Thaumatin with 60° of data measured in 600 ms at 6 keV. Density map obtained after density modification, automated tracing, and refinement with CRANK2. b, Lysozyme with 500° of data measured in 5 s at 12.4 keV. Density map obtained after density modification and automated tracing with SHELXE. c, PepN with 600° of data measured in 1 min at 6 keV. Density map obtained after density modification and automated tracing with SHELXE.

crystal as the model system. We collected a total of 60° of data from one crystal at 6 keV with a rotation speed of 100° s<sup>-1</sup>. The entire exposure lasted 0.6 s. With these data, all sulfurs were identified readily with SHELXD<sup>30</sup>, and the electron density map is shown in Fig. 6a. Thaumatin with 60° of data measured in 600 ms at 6 keV. Density map obtained after density modification, automated tracing, and refinement with CRANK2. b, Lysozyme with 500° of data measured in 5 s at 12.4 keV. Density map obtained after density modification and automated tracing with SHELXE. c, PepN with 600° of data measured in 1 min at 6 keV. Density map obtained after density modification and automated tracing with SHELXE.

to the accumulation of weak anomalous signals (Supplementary Tables 1 and 6).

Next, we selected Escherichia coli aminopeptidase N (PepN), representing a more challenging example. PepN is a 101-kDa protein (891 residues) containing 30 sulfurs, 1 bromine, and 1 zinc atom. Because the average diffraction power of PepN is much weaker than that of test proteins such as thaumatin and lysozyme, we limited the rotation speed of data collection to 10° s<sup>-1</sup> to ensure sufficient exposure per diffraction image with a flux of 2.7 × 10<sup>11</sup> photons per second at 6 keV. A 600° dataset, collected in 1 min (Supplementary Tables 1 and 6), allowed a straightforward structure solution using SHELXC/D/E (Fig. 6c). If five to ten times more flux were available, we would expect that the same structure could be solved within 5–10 s using 50–100° s<sup>-1</sup> rotation.

**Discussion**

The diffraction pattern of a macromolecular crystal contains thousands of sharp Bragg peaks with large variations in intensity. Structure solution, especially by experimental phasing methods, requires accurate measurement of strong Bragg peaks, and the atomic features of a structure are determined by precise recording of weak Bragg peaks at high resolution. Strong and weak intensities pose different challenges for measurement. Ideally a detector should have a uniform response across a large surface area on both the pixel and the sub-pixel level, a high dynamic range, high sensitivity to the single-photon level, a single-pixel point-spread function, and continuous readout. None of the previous generations of detectors addressed all of these points. JUNGFRAU meets the challenge by using a charge-integrating readout chip and direct-detection hybrid pixel detector technology to provide a low-noise performance over the whole dynamic range of 10<sup>4</sup>–10<sup>12</sup> keV photons per frame per pixel and a uniform response within and across pixels.

From a data accuracy and precision point of view, the requirements are most stringent in experimental phasing because the small anomalous differences (≤1%) between Bragg peaks related by Friedel’s law and crystal symmetry lead to reliable evaluation of phases. The size of Bragg spots is similar to or smaller than the pixel size of the detector, their measurement accuracy will be compromised when they are measured with detectors with a nonuniform sub-pixel response such as HPC detectors. The smaller the spot is in comparison to the detector pixel size, the more severe the effect will be. In practice, one can mitigate this shortcoming conveniently by collecting true high-multiplicity data with a multi-axis goniometer<sup>28</sup>, but at the cost of increased X-ray dose, and thus increased radiation damage<sup>33</sup> and experiment time. We have demonstrated that JUNGFRAU permits accurate measurement of photons regardless of where they land on the detector surface, which allows the user to obtain highly accurate data and thus achieve experimental phasing with minimum X-ray dose and reduced multiplicity (Fig. 4). Therefore, the JUNGFRAU detector holds great promise, especially for native-SAD phasing with X-rays in the range of 3–5 keV, where the calibration of HPC detectors is particularly challenging. Furthermore, the measurements can be carried out faster with high flux because JUNGFRAU is not count-rate limited. We demonstrated this unique combination of accuracy and speed in native-SAD experiments (Fig. 6a–b) with a rotation speed of 100° s<sup>-1</sup>, a speed that was considered of no practical use in the past but now can be exploited for the development of novel data-collection strategies.

Time-resolved crystallography with Laue methods was made possible by third-generation high-energy synchrotron facilities<sup>47</sup>. However, Laue methods require large crystals and specialized beamlines. The emerging serial synchrotron crystallography (SSX) technique has introduced novel crystal-delivery techniques and automated data-collection methods with fast frame-rate detectors<sup>31–37</sup>. To further improve the efficiency of SSX methods and
the time resolution, one can increase the available flux density 100–1,000-fold through the use of wide-band-pass X-rays. Then integrating detectors become indispensable. JUNGFRAU technology meets these challenges nicely and should allow the study of biologically relevant dynamics down to mircosecond time scales in a pump-probe fashion at synchrotrons. There are challenges to the implementation of JUNGFRAU at MX beamlines. ‘Dark’ runs (that is, without X-rays) are required to calculate pedestals for each gain and need to be included in the data-acquisition sequence with minimum overhead. The raw data need to be corrected and converted to photons before the data volume can be reduced by frame summation. This requires handling of high data rates (4 GB s⁻¹ per 1 million pixels) for real-time data analysis. Researchers at the Paul Scherrer Institute are actively developing solutions to match the robustness and simplicity of operating HPC detectors. Such challenges are essentially the same for XFEL serial crystallography applications.

Current HPC detectors produce data of high quality for the majority of MX applications, but they have their limitations. The improvement in data accuracy and data-collection speed that we obtained with the JF1M detector is remarkable. The ultimate obtainable data quality from a given crystal depends on many factors, but it is evident that detectors like JUNGFRAU will be pivotal in helping scientists get close to this limit. We expect that detectors like JUNGFRAU will prompt the development of low-noise instruments in the next-generation MX beamlines to capitalize on the full potential of the next-generation synchrotron sources like diffraction-limited storage rings in the coming decade.

References
Methods

General experiment setup. Experiments were conducted at the X06SA protein crystallography undulator beamline, Swiss Light Source, at beam energies of 12.4 keV and 6 keV. The beam size was adjusted to 80 x 80 μm and the flux for a nonattenuated beam was 1.6 x 10^10 photons (ph)/s and 2.7 x 10^10 ph/s for 12.4 and 6 keV, respectively. For 12.4-keV measurements we used the default beamline settings, whereas for 6 keV we detuned the monochromator by 0.002° to remove higher harmonics. The beamstop was placed 7 mm from the sample, which shadowed reflections with resolutions less than 10 Å for 12.4-keV X-rays. The beamline was equipped with a motorized stage that allowed movement of the JUNGFRAU and EIGER detectors in three directions. Sample-to-detector distances could be varied from 40–120 mm, and the two perpendicular directions could be set within 20 mm from the detector center. The motor resolution was 2.5 μm. Crystal centering and EIGER data collection were controlled with DA+ software8. The JUNGFRAU data collection was carried out with customized programs.

JF1M detector characteristics. The unique feature of the JUNGFRAU detector is its dynamic gain-switching, with three gain levels accommodating both single-photon sensitivity and high dynamic range. The JUNGFRAU detector is modular, and each module has an active area of 4 x 8 cm² with eight application-specific integrated circuits (ASICs) and contains ~500,000 pixels of 75-μm pitch. The sensor geometry is identical to that of EIGER. Modules are independent in terms of readout; each has a dedicated 10 Gb/s Ethernet link and can be arranged into various geometric shapes. Currently silicon of 320-μm thickness is used for this JUNGFRAU sensor. A thicker sensor, such as the 450-μm-thick sensor of EIGER, could also be used.

The JUNGFRAU system used in this experiment consisted of two modules forming a single JUNGFRAU system (JF1M). The ASIC of JUNGFRAU is designed to keep the readout noise below Poisson σ, which is roughly doubled at the foreseen operation frame rate of 2.3 kHz. Other parameters, such as the internal ASIC voltages, sensor bias voltage, and timing, were standard as also used in XEELs. We used a dedicated computer to control the detector and to store frames during data collection. The frame rate (1.136 kHz) and frame size (1 million pixels in 16-bit) required a wide bandwidth of 2.3 GB/s to prevent frame loss.

The JUNGFRAU detector is designed to keep the readout noise below Poisson statistics and to have single-photon sensitivity at energies as low as 2 keV. The readout noise is estimated as 200 electrons for the high gain with an integration time of 840 μs. When the detector is operated at an XEEL with an integration time of 10 μs, the noise is reduced to 70 electrons.

The maximum number of counts is determined by the charge range of the low gain. Because the induced charge from a single photon is proportional to its energy, the dynamic range is effectively doubled at 6 keV compared with that at 12 keV. When the detector is operated at a 1 kHz frame rate, the dynamic range is about 12 and 25 Mcps pixel⁻¹ (2,100 and 4,400 Mcps mm⁻²) at 12 and 6 keV, respectively, and is roughly doubled at the foreseen operation frame rate of 2.3 kHz.

JUNGFRAU data format and image processing. The result of each JUNGFRAU measurement is a raw image. For each pixel, the gain level (2-bit) and digitized accumulated charge (14-bit) are recorded. For conversion of raw signal to photon energy, six constants are needed per pixel: for each of the three gain levels, one needs to know the amplification factor (i.e., the ratio of arbitrary detector charge units and energy) and the pedestal (i.e., the offset corresponding to the energy) and is roughly doubled at the foreseen operation frame rate of 2.3 kHz.

Dynamic range study. A thauatin crystal (Thau1; 480 x 240 x 180 μm³) was measured at 12.4 keV with a flux of 3.5 x 10^10 ph/s (0.25% beam transmission). The datasets with full rotation (360°) were measured at a rotation speed of 50 s⁻¹ with both JF1M and E1M detectors. The accumulated dose was about 0.6 keV of crystal damage, which was below the detection limit, while detectors were exchanged, so both measurements were made with the same position of the crystal and same X-ray beam conditions. The JF1M and E1M detectors were positioned approximately 60 mm from the sample and operated at a 1.136 kHz frame rate and 500 Hz, respectively.

Sub-pixel uniformity study. Measurements were carried out with both JF1M and E1M detectors operated with the same frame rate of 1.136 kHz and positioned 45 mm from the crystal. Two settings were used for the E1M: one with the default 50% threshold (E1M-30), and the other with 60% (E1M-60). The integration times of JF1M and E1M were 840 μs and 877 μs, corresponding to duty cycles of 95.5% and 99.7%, respectively. A large thaumatin crystal (Thau3) of about 360 x 240 x 240 μm³ was measured at 6 keV with 15% beam transmission (flux of 2.5 x 10^12 ph/s). The same crystal volume was illuminated with the same X-ray beam through the entire experiment.

Protein crystal preparation. Lysozyme was dissolved at 50 mg/ml in 50 mM sodium acetate, pH 4.5, and crystallized in 5% PEG MME 5000, 2 M NaCl, 50 mM sodium acetate, pH 4.5, 25% ethylene glycol. Thaumatin was suspended at 50 mg/ml in water and crystallized in 24% sodium potassium tartrate, 100 mM Bis-Tris propane, pH 6.5. Both lysozyme and thaumatin were obtained from Sigma-Aldrich. PepN was purified and crystallized with inhibitor 11 according to a published protocol44.

X-ray data collection. Low-noise performance. A large thauatin crystal (Thau1; 480 x 240 x 180 μm³) was measured at 12.4 keV with a flux of 3.5 x 10^10 ph/s (0.25% beam transmission). The datasets with full rotation (360°) were measured at a rotation speed of 50 s⁻¹ with both JF1M and E1M detectors. The accumulated dose was about 0.6 keV of crystal damage, which was below the detection limit, while detectors were exchanged, so both measurements were made with the same position of the crystal and same X-ray beam conditions. The JF1M and E1M detectors were positioned approximately 60 mm from the sample and operated at a 1.136 kHz frame rate and 500 Hz, respectively.

Sub-pixel uniformity study. Measurements were carried out with both JF1M and E1M detectors operated with the same frame rate of 1.136 kHz and positioned 45 mm from the crystal. Two settings were used for the E1M: one with the default 50% threshold (E1M-30), and the other with 60% (E1M-60). The integration times of JF1M and E1M were 840 μs and 877 μs, corresponding to duty cycles of 95.5% and 99.7%, respectively. A large thaumatin crystal (Thau3) of about 360 x 240 x 240 μm³ was measured at 6 keV with 15% beam transmission (flux of 2.5 x 10^12 ph/s). The same crystal volume was illuminated with the same X-ray beam through the entire experiment.

All diffraction data were collected with a rotation speed of 10° s⁻¹, that is, a step of 0.0088° in 0.00888 s. Five 360° datasets were measured for each detector. We carried out the first two runs with the detector in an initial position, the third with the detector shifted by 25 μm (1/3 pixel) in both x and y directions from the initial position, the fourth with the detector shifted by 50 μm (2/3 pixel) in both x and y directions, and finally one with the detector shifted by 225 μm (3 pixels) in both x and y directions relative to the initial position. Only results of the first three experiments are presented. The total dose accumulated through the experiment was estimated as less than 0.5 MGy, well below the damaging dose limit for cryo-cooled crystals.

Fast native-SAD phasing. For native SAD at 6 keV X-ray energy, we used a beam size of 80 x 80 x 80 μm³ and the full flux of 2.7 x 10¹⁰ ph/s. A thaumatin crystal with a size matching the beam size (Thau2; 80 x 80 x 80 μm³) was measured for a 360° angular range at 100 s⁻¹ rotation speed. The frames with 0.088° angular increment were used for data processing directly, without summing. At the same setting of 6 keV for a PepN crystal of 100 x 80 x 80 x 80 μm³ was used to measure a 360° angular range at a rotation speed of 10° s⁻¹. Ten frames were summed to make one image covering a 0.088° rotation width.

For native SAD at 12.4 keV, a lysozyme crystal of 80 x 80 x 80 μm³ was used. The lysozyme dataset was measured with 100 s⁻¹ rotation speed and 100% beam transmission (1.6 x 10¹⁰ ph/s).
X-ray data processing, structure determination, and refinement. MX data quality is dependent on phi-slicing\(^ {25} \), and in principle a slower rotation speed allows for finer slicing at a given detector frame rate, which could result in a bias toward slower rotation speeds (up to the point where data-processing software can correctly account for the extremely weak signal and low background). Therefore, before data processing, we carried out frame summation to ensure that images obtained at various rotation speeds corresponded to a similar rotation angle (0.088° for 1.136 kHz and 0.100° for 1.000 kHz).

Frames were processed with XDSS\(^ {25} \) software with standard settings. To improve position refinement for the pixel map calculation, we used the segment refinement feature of XDSS to account for imprecisions in module positions and the gap size in JF1M and E1M. To allow a direct comparison of intensities for the dynamic range and low-noise performance experiments, we fixed scaling factors for integration in XDSS at 1.0. We divided intensities calculated in the XDS_ASCII.HKL by the Lorenz-polarization correction factor, to recover the total photon count of a reflection for presentation in Fig. 3a.

The calculation of data-quality indicators (R\(_ {\text{mean}}\), R\(_ {\text{merge}}\) and (\( \langle \text{I}\rangle \))) was performed on the basis of XDSS and XSQUEEZE outputs using custom Python scripts for plotting in finer-resolution shells. In the low-noise performance experiment, the normalization of intensity was calculated with the ratio of the duty cycles and the ratio of the absorptions of the Si sensor at 29 of 32° (Fig. 2c). \( \langle \text{I}\rangle \) was normalized with the square root of the ratios (Fig. 2d). The duty cycle was 95% and 99.7% and the sensor thickness was 320 \( \mu \text{m} \) and 450 \( \mu \text{m} \) for JF1M and E1M, respectively.

Experimental phasing with native SAD was carried out with SHELXC/D/E\(^ {29} \) via HKL2MAP GUI\(^ {30} \) or with the CRANK2 pipeline\(^ {31} \). The mean peak height for anomalous data (\( \langle \text{I}_{\text{anomal}}\rangle \)) was calculated with ANODE\(^ {28} \). The structures were refined with phenix.refine\(^ {17} \) and deposited in the Protein Data Bank (see “Data availability”).

Sub-pixel uniformity characterization. To explore the systematic errors of the detector on the sub-pixel level, we grouped all the reflections according to where they impinged relative to a pixel center. In this task we benefited from the fact that XDSS provides the predicted reflection center to a precision of 1/10 of a pixel. For each reflection we considered only the fractional part of its position in pixel units, ignoring its integer part. For example, if a spot was predicted to fall at \( x = 450.1 \) pixel and \( y = 363.5 \) pixel, we considered its “in-pixel position” as \( x = 1, y = 5 \). Because in XDSS the coordinate system in-pixel position \( x=0, y=0 \) corresponds to the center of a pixel, we shifted the positions by half a pixel to put the origin of the coordinate system at a corner.

To quantify such spatial effects, we first calculate the deviation from the mean for each observation:

\[
\Delta = \frac{n}{n-1} (I_{\text{obs}} - I_{\text{mean}})
\]

where \( n \) is multiplicity, \( I_{\text{obs}} \) is the measured intensity, and \( I_{\text{mean}} \) is the mean intensity for all symmetry equivalent reflections (including the one in question). The extra term, \( n/(n-1) \) corrects for underestimation of the difference between the observation and the mean. \( I_{\text{mean}} \) is then simply

\[
R_{\text{mean}} = \frac{\sum_{n \in \text{ref}} |I_{\text{obs}} - I_{\text{mean}}|}{\sum_{n \in \text{ref}} I_{\text{obs}}}
\]

where \( n \) is the multiplicity. Reflections that were observed only once are ignored in the summation.

Next we bin all reflections according to their in-pixel position, and for each position \( x,y \) we calculate

\[
\Delta_{x,y} = \frac{\sum_{n \in \text{ref}} |I_{\text{obs}} - I_{\text{mean}}|}{\sum_{n \in \text{ref}} I_{\text{obs}}}
\]

where \( n \) is the number of reflections that fall into a particular \( x,y \) in-pixel position. To allow comparison between in-pixel positions, \( \Delta_{x,y} \) can be also normalized similarly to \( R \) factors:

\[
\Delta_{x,y}^{\text{norm}} = \frac{\Delta_{x,y}}{\frac{\sum_{n \in \text{ref}} I_{\text{obs}}}{\sum_{n \in \text{ref}} I_{\text{obs}}}}
\]

Because \( \Delta_{x,y}^{\text{norm}} \) is calculated without the absolute value of \( \Delta \), being taken before averaging, random differences in intensity measurements should cancel out—a value close to zero of \( \Delta_{x,y}^{\text{norm}} \) should indicate that there is no systematic error introduced at in-pixel position \( x,y \). However, if reflections in a particular bin are systematically larger or smaller than the ones in other bins, \( \Delta_{x,y}^{\text{norm}} \) should indicate it by a positive or negative value, respectively. \( \Delta_{x,y}^{\text{norm}} \) values for each in-pixel position can then be presented on a map that indicates the degree of the nonuniformity across 1 pixel. Pixel maps calculated with low-resolution reflections (\( \Delta > 10 \) \( \AA \)) are presented in Fig. 5a–c.

With the value of \( \Delta_{x,y} \) known, one can calculate the effect that charge sharing has on the \( R \) factor value, by calculating the mean of the absolute values of \( \Delta_{x,y} \):

\[
R_{\text{merge}} = \frac{\sum_{n \in \text{ref}} |I_{\text{obs}} - I_{\text{mean}}|}{\sum_{n \in \text{ref}} I_{\text{obs}}}
\]

where \( n \) is the number of all reflections with multiplicity of at least two. Because \( |a| + |b| \geq |a+b| \), \( R_{\text{merge}} \) is an upper limit for \( R_{\text{mean}} \) and comparison of the two values can indicate the share of systematic errors due to sub-pixel nonuniformity in relation to the total uncertainty.

For calculations we apply a standard cutoff for reflection intensities of \( I > 3\sigma \). Because we are interested in systematic deviations of reflection intensities, we also include misfits, marked in XDSS ASCII.HKL, with negative \( n \) values, in all statistics calculations presented in Fig. 5d–f (\( R_{\text{mean}} \) and \( R_{\text{merge}} \)).

Photon count-rate estimation. The peak photon rate for a reflection observation was approximated as the following:

\[
\text{Rate} = \frac{\text{Max} C}{(0.075 \text{ mm})^2} \frac{v}{\Delta \phi}
\]

where \( \text{Max} \) is the highest count observed in a single pixel from a single frame for a particular reflection (column MAXC in INTEGRATE.HKL from XDSS), \( v \) is the rotation speed in degrees per second, \( \Delta \phi \) is the rotation range of a single image in degrees, and 0.075 mm is the pixel pitch. This number is only the lower estimation of the peak rate, because while the crystal rotates, the intensity of a reflection varies according to its rocking curve, especially if \( \Delta \phi \) is larger than the mosaicity (as in our case). However, if one compares two datasets collected with the same \( \Delta \phi \), the incoming photon rate should be similar in both. The spread in observed values might come from the different spread of counts inside a peak (charge sharing).

In Fig. 3d we present correlations of peak rates of two JF1M datasets collected on the same crystal at 1° s\(^{-1} \) and 100° s\(^{-1} \) rotation speeds with corresponding beam transmissions of 1% and 100% (see above for exact experimental details). To ensure that equivalent \( \Delta \phi \) frame summation was performed on the slower dataset, for the correlation plot, we chose only reflections with identical Miller indices from both datasets, and we applied no symmetry equivalence. In this way, peak rates calculated in 1° s\(^{-1} \) JF1M data, multiplied by 100, are an approximation of the ‘true’ rates for 100° s\(^{-1} \) data. These rates are then compared with the measured rate values in 100° s\(^{-1} \) data.

For reference, we calculate peak rate values using a theoretical model for a paralyzable counter, where the relation between true count rate \( I \) and observed count rate \( I_{\text{obs}} \) is given as

\[
I = I_{\text{obs}} e^{-\tau}
\]

where \( r \) is an energy-dependent sensor dead time. The value used in Fig. 3d was taken as 180 ms, which is an experimental value determined for 6-keV photons for the PSI manufactured EIGER\(^ {32} \).

References


