White Paper

EIGER detectors in biological crystallography

by Andreas Förster, Stefan Brandstetter, Marcus Müller and Clemens Schulze-Briese
DECTRIS Ltd., Täfernweg 1, 5405 Baden-Dättwil, Switzerland

1. Introduction

Because of their superior properties, Hybrid Photon Counting (HPC) detectors have largely superseded their predecessors in macromolecular crystallography over the last few years. In contrast to earlier technology like CCDs and image plates, HPC detectors count individual photons and suppress noise. As a result, excellent data quality can be obtained even from marginal samples. Structures of increasingly large protein complexes, whose crystals tend to diffract poorly, are being solved, and novel methods have been proposed for the retrieval of phase information, which is normally lost in a crystallographic experiment.

This development was facilitated by enhancements to beam stability and brilliance at synchrotron facilities where most macromolecular crystallography is done. Further upgrades that amount to generational shifts have been proposed at several of these facilities [1]. New detectors are needed to take advantage of what promises to be dramatic advances. This White Paper introduces EIGER, the most highly performing detector for biological X-ray crystallography, of which three X 16M, three X 9M and four X 4M systems have already been installed at synchrotron facilities around the world.

2. Technical description and specifications

As an improvement over PILATUS, which is in regular operation at virtually all synchrotrons worldwide, DECTRIS has developed EIGER, a detector to make the most of state-of-the-art beamlines. It shares with PILATUS the ability to collect best data, uncorrupted by dark current or readout noise, but surpasses its predecessor in several important respects (Table 1). For example, the pixel size of EIGER has been reduced by 80% in area compared to PILATUS and now measures 75×75 μm^2. Even so, the point-spread function of measured X-ray photons is still only one pixel.

The dead-time between the reading of successive frames is 3.8 µs, which enables continuous readout with duty cycles beyond 99%. Rates of image acquisition range from 133 Hz for the large EIGER X 16M with more than 16 million pixels to 750 Hz for the EIGER X 4M. These high frames rates are critical for advanced data collection protocols that spread the permissible X-ray dose over as many frames as possible to obtain data with high multiplicity while minimizing the detrimental effects of radiation damage [2].

EIGER detectors exist in configurations for the synchrotron as well as for laboratory and industry applications, and in a number of sizes (Figure 1). All share maximal achievable count rates of up to $3 \cdot 10^6$ photons·s^{-1}·$pixel^{-1}$ ($5 \cdot 10^8$ photons·s^{-1}·mm^{-2}) and are fitted with Si sensors 450 µm thick for best quantum efficiency over a wide energy range (threshold energies of 2.7 - 18 keV).

	PILATUS3 X 2M, 6M	EIGER X 4M, 9M, 16M
Pixel size [µm^2]	172 x 172	75 x 75
Active area [pixels]	2-6 million	4-16 million
Dead time [µs]	950	3.8
Frame rate [Hz]	250-100	750-133

Table 1: Technical Specifications of PILATUS3 and EIGER. Key common features of large systems for macromolecular crystallography at synchrotrons are compared between PILATUS3 and EIGER.

3. System architecture

The active surface of EIGER detectors consists of tiled arrays of modules. The inactive area due to the gaps between the modules causes a slight reduction in multiplicity and potential problems with completeness in P1 symmetry. However, for the processing of the vast majority of crystallographic data, these gaps are immaterial. Nevertheless, the inactive area of EIGER detectors has been reduced over similarly sized PILATUS detectors.

EIGER X 1M and 4M systems are vacuum compatible and available with optional low-energy calibration. This makes them ideal instruments for measuring the anomalous signal from atoms natively occurring in biological macromolecules. A recent review by Liu and Hendrickson [3] estimated that nearly 95% of the structures in the PDB are candidates for native SAD protocols. EIGER detectors have been designed with this technique in mind.

4. System control

EIGER features a powerful application programming interface (API) that provides platform-independent access to the detector. It makes integration with beamline infrastructure and controling detectors straightforward. The API is based on the http/REST standard [4] uses JSON [5] and can be scripted easily, for example with Python, to query detector settings, change data collection parameters, update the pixel mask, and implement data collection protocols of arbitrary complexity.

Data can be monitored with the visualization software ALBULA [6] as they are acquired, streamed out as binary objects or saved to files. The data stream is particularly attractive where tight integration with computational infrastructure at the beamline is required because the data can be accessed and manipulated at a fundamental level.



Figure 1: EIGER X series of detectors. Shown are EIGER X 16M, EIGER X 9M, EIGER X 4M, EIGER X 1M (left to right).

By default, the detector is read out as an array of pixel values that covers the entire surface of the detector. For EIGER X 16M and EIGER X 9M, the firmware since version 1.6 enables the user to select for readout a subimage corresponding in size to an EIGER X 4M. As for the EIGER X 4M itself, the subimage can be read at frame rates of up to 750 Hz – with tremendous benefits for applications like grid screening and *in-situ* serial crystallography where individual images or highly incomplete datasets are collected of a large number of microcrystals randomly positioned on a sample holder [7] or injected into the beam [8] in a manner pioneered at X-ray free-electron lasers. The incomplete datasets are later merged to yield complete data of quality high enough to be used for structure determination by experimental phasing [9].

5. File format and compression

An EIGER X 1M produces up to 90,000 files in half a minute. Instead of individual files, which would put an inordinate strain on the file system at such frame rates, EIGER uses the HDF5 format to save entire datasets. A master file stores relevant metadata in compliance with the NeXus NXmx standard [10] and points to one or more data files that contain the images as compressed arrays of pixel values. Beamline-specific metadata, like goniometer geometry or identity of the beamline, can easily be added by the beamline control system.

Focusing on datasets instead of images and keeping all metadata in one place make it easy for facilities and users to keep track of, transfer and archive data. It is also in line with what other scientific communities are doing, which makes it possible to harness synergies in the development of local computing and network infrastructures. Using HDF5 [11], the *de-facto* standard for big data, will make it easier for macromolecular crystallography to transition into that world.

The processing of crystallographic data in HDF5 format is unproblematic as long as sufficient computational resources exist. Three strategies have been developed for working with the new format. Mosflm [12] and HKL2000 [13] obtain processing parameters from image headers and

Compression method	Angular increment [°]	Frame rate [Hz]	Image size [MB]	Compression factor	Data per 360° [GB]	Data rate [MB/s]
none	-	-	72	1	-	
CBF	0.25	4	18	4.0	25	72
bs-LZ4	0.25	4	11	**6.5**	15	44
bs-LZ4	0.1	10	6.5	**11.1**	23	65
bs-LZ4	0.05	20	5.5	**13.1**	39	110
bs-LZ4	0.02	50	3.5	**10.3**	62	175
bs-LZ4	0.01	100	2.7	**13.3**	95	270
bs-LZ4	0.01	133	2.5	**14.4**	88	333
CBF	0.01	133	18	2.0	633	2394
none	-	-	36	1	-	

Table 2: Compression of crystallographic data. Insulin datasets collected with EIGER X 16M at Swiss Light Source show that compression factors are higher with bit-shuffle LZ4 (bs-LZ4, green rows) than with CBF (blue rows). Compression with bs-LZ4 depends on the sparsity of the data, with images with less data (higher frame rates at constant flux) compressing better. The white rows show the sizes of uncompressed EIGER X 16M images. At frames rates below 50 Hz (above dashed line), EIGER X 16M acquires 32-bit images, at rates of at least 50 Hz (below dashed line), 16-bit images. The highest observed data rate was 333 MB/s (ca. 2.7 Gbit/s) with bs-LZ4 compression (red).

require EIGER data to be presented as crystallographic binary format (CBF) containing mini-CBF header information. A converter has been developed for this task [14]. XDS [15], on the other hand, can use H5ToXds, an image data extractor developed by DECTRIS [16], or a converter included with the autoPROC package by GlobalPhasing [17] to process HDF5 data. Extracted CBF images exist as temporary files only. Lastly, the newly developed program DIALS [18] can read and process HDF5 data directly. It was used to process the data the led to the first published structure solved with EIGER X 16M data [19].

Bit-shuffle filtered LZ4 [20, 21], the compression algorithm implemented in EIGER, is fast enough for real-time compression and expansion of images and achieves compression rates beyond the current standard CBF (Table 2, reference 22). Once the compression libraries are enabled as plugins for HDF5, their use is totally transparent.

6. Crystallographic performance

EIGER detectors have recently entered operation at pioneering synchrotrons. Early results indicate that EIGER is superior to PILATUS in three important ways [23]. First, the small pixel size allows users to take advantage of the small beam diameter and low beam divergence at modern synchrotrons to collect better data with less background. Second, thanks to the dramatically reduced dead time of EIGER, finer ϕ-slicing than is possible with PILATUS results in higher signal-to-noise ratios and lower R values. Since no penalty is associated with extreme fine ϕ-slicing, users should always ϕ-slice as finely as possible and sum image into appropriate slices prior to data processing.

Third, because of a duty cycle of above 99%, high angular speeds can be achieved without loss of data quality as long as goniometer imprecision at high rotation speed and high-frequency fluctuations of X-ray beam intensity and position are negligible. At Swiss Light Source (SLS), one second of data collected at an angular velocity of 64°/s and an X-ray energy of 8.4 keV was used to experimentally phase native cubic insulin from the signal of the intrinsic sulfur atoms.

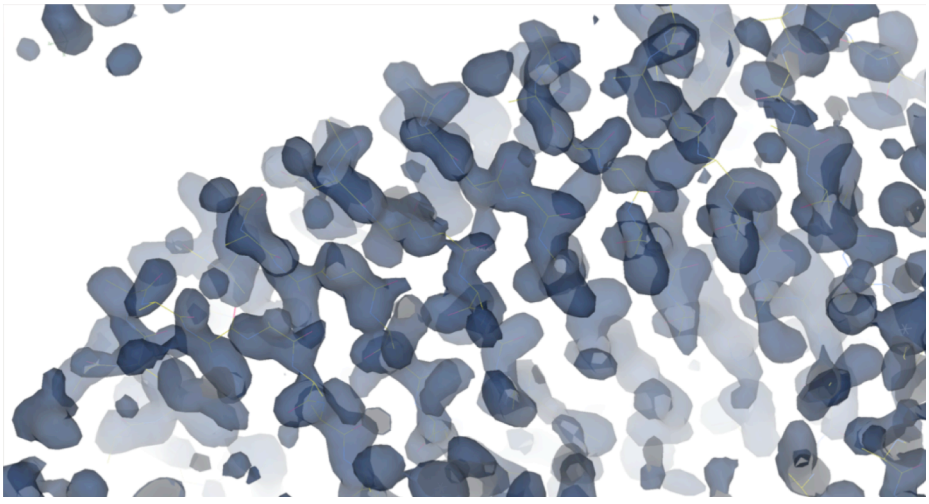


Figure 2: Native SAD with EIGER X 9M data. Experimental map of a leucine-rich repeat protein phased with the anomalous signal of nine zinc ions. Figure courtesy of William Shepard, Synchrotron SOLEIL.

The selection of results below summarizes the high level of performance EIGER has already achieved.

- Users of beamline PROXIMA-2 at Synchrotron SOLEIL solved the first new protein structure from EIGER X 9M data four weeks after delivery of the detector to the synchrotron [24].

- At Synchrotron SOLEIL, William Shepard solved the structure of a leucine-rich repeat protein from 250° of data collected in 2.1 s at a wavelength of 0.98 Å with an EIGER X 9M. Despite the low-symmetry space group of 21, the structure could be phased easily from the anomalous signal of nine Zn^{2+} ions (Figure 2).

- The crystal structure of CRISPR-Cpf1 in complex with guide RNA and target DNA was solved from EIGER X 16M data collected at SLS and published in Cell [19].

- Industrial users are collecting upwards of twelve full datasets per hour on the EIGER X 16M at SLS. Efficient compression with bit-shuffle LZ4 is critical for the success of high-throughput applications.

- Taking advantage of the small pixel size of the EIGER X 16M, data from a crystal with the largest unit cell dimension of 1070 Å was collected without overlaps at SLS (Figure 3).

- At Photon Factory, two EIGER X 4M detectors inside a helium chamber allow for the collection of data with minimal background at long wavelengths for the accurate measurement of anomalous differences from light atoms like sulfur and phosphorous. A maximum resolution of 2.3 Å can be achieved at a wavelength of 4.6 keV.

- The crystal structure of PsoE, a biosynthetic enzyme for a fungal secondary metabolite, was solved from EIGER X 4M data collected at Photon Factory [25].

7. Conclusion

Several large EIGER detectors have already been delivered to synchrotron facilities around the world. An EIGER X 16M has been operational at SLS in Villigen, Switzerland, since October, and two more have been installed at National Synchrotron Light Source II in Brookhaven, USA, and at Advanced Photon Source in Argonne, USA. Three EIGER X 9M detectors have been installed, one at Synchrotron SOLEIL in Saclay, France, one at Advanced Photon Source in Argonne, USA, and one

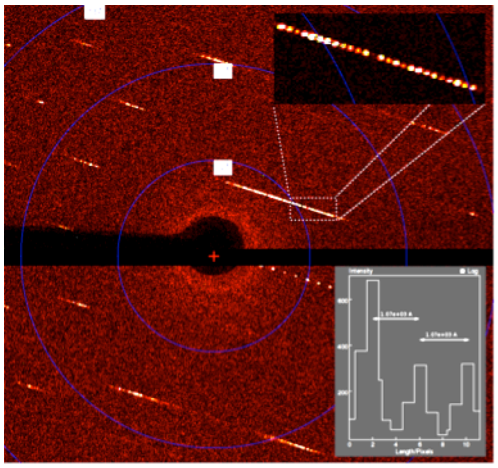


Figure 3: Diffraction from a crystal with extreme unit cell dimensions. The combination of large active area, small pixel size and single-pixel point-spread function allows EIGER X 16M to resolve individual spots from a crystal with a unit cell dimension of above 1070 Å up to a resolution of 2.3 Å at the edge of the detector. Figure courtesy of Arnau Casañas, Swiss Light Source.

at SPring-8 in Hyogo Prefecture, Japan. In addition, EIGER X 4M detectors are in operation at Photon Factory in Tsukuba, Japan, and the European Synchrotron Radiation Facility in Grenoble, France.

With EIGER, the crystallographic community will acquire a tool to make today's cutting-edge applications like micro-crystallography, serial crystallography at room temperature and grid-scanning at microfocus beamlines more robust and accessible to a wide user base. At the same time, inventive users will develop novel methods and original approaches that exploit the potential of EIGER in unexpected ways. As with any leap in instrumentation technology in X-ray crystallography, the result will be structural information on formerly intractable systems and a deeper understanding of biology.

References

The links below were retrieved on 9 May 2016.

1. Reich ES (2013), *Ultimate upgrade for US synchrotron*, Nature 501:148
2. Weinert T *et al.* (2015), *Fast native-SAD phasing for routine macromolecular structure determination*, Nat Methods 12:131
3. Liu Q and Hendrickson WA (2015), *Crystallographic phasing from weak anomalous signal*, Curr Opin Struct Biol 34:99
4. Richardson L and Ruby S (2007), *RESTful web service*, O'Reilly Media, ISBN 978-0-596-52926-0
5. http://jsonapi.org/
6. https://www.dectris.com/Albula_Overview.html
7. Zander U *et al.* (2015), *MeshAndCollect: an automated multi-crystal data-collection workflow for synchrotron macromolecular crystallography beamlines*, Acta Cryst D71:2328
8. Nogly P *et al.* (2015), *Lipidic cubic phase serial millisecond crystallography using synchrotron radiation*, IUCrJ 2:168
9. Huang C *et al.* (2016), *In meso in situ serial X-ray crystallography of soluble and membrane proteins at cryogenic temperatures*, Acta Cryst D72:93
10. http://download.nexusformat.org/sphinx/classes/applications/NXmx.html
11. https://www.hdfgroup.org/why_hdf/
12. Battye TGG *et al.* (2011), *iMosflm: a new graphical interface for diffraction-image processing with MOSFLM*, Acta Cryst D67:271
13. Otwinowski Z and Minor W (1997), *Processing of x-ray diffraction data collected in oscillation mode*, Methods Enzymol 276:307
14. https://github.com/biochem-fan/eiger2cbf
15. Kabsch W (2010), *XDS*, Acta Cryst D66:125
16. https://www.dectris.com/hdf5.html
17. http://www.globalphasing.com/autoproc/
18. Waterman DG *et al.* (2016), *Diffraction-geometry refinement in the DIALS framework*, Acta Cryst D72:558
19. Yamano T *et al.* (2016), *Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA*, Cell 165:949
20. https://github.com/kiyo-masui/bitshuffle
21. https://github.com/Cyan4973/lz4
22. http://www.bernstein-plus-sons.com/software/CBF/
23. Casañas A *et al.*, *Manuscript in preparation*
24. http://www.synchrotron-soleil.fr/Soleil/ToutesActualites/2015/EIGER122015
25. Yamamoto T *et al.* (2016), *Oxidative trans to cis isomerization of olefins in polyketide biosynthesis*, Angew Chem Int Ed Engl 55:6210