Accurate determination of segmented X-ray detector geometry

Oleksandr Yefanov,[1,*] Valerio Mariani,[1] Cornelius Gati,[1] Thomas A. White,[1] Henry N. Chapman,[1,2,3] and Anton Barty[1]

[1] *Center for Free-Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany*
[2] *Department of Physics, University of Hamburg, Luruper Chaussee 149, 22607 Hamburg, Germany*
[3] *Centre for Ultrafast Imaging, Luruper Chaussee 149, 22607 Hamburg, Germany*
[*] *oleksandr.yefanov@desy.de*

Abstract: Recent advances in X-ray detector technology have resulted in the introduction of segmented detectors composed of many small detector modules tiled together to cover a large detection area. Due to mechanical tolerances and the desire to be able to change the module layout to suit the needs of different experiments, the pixels on each module might not align perfectly on a regular grid. Several detectors are designed to permit detector sub-regions (or modules) to be moved relative to each other for different experiments. Accurate determination of the location of detector elements relative to the beam-sample interaction point is critical for many types of experiment, including X-ray crystallography, coherent diffractive imaging (CDI), small angle X-ray scattering (SAXS) and spectroscopy. For detectors with moveable modules, the relative positions of pixels are no longer fixed, necessitating the development of a simple procedure to calibrate detector geometry after reconfiguration. We describe a simple and robust method for determining the geometry of segmented X-ray detectors using measurements obtained by serial crystallography. By comparing the location of observed Bragg peaks to the spot locations predicted from the crystal indexing procedure, the position, rotation and distance of each module relative to the interaction region can be refined. We show that the refined detector geometry greatly improves the results of experiments.

OCIS codes: (120.1880) Detection; (120.3940) Metrology; (040.7480) X-rays, soft x-rays, extreme ultraviolet (EUV); (050.1940) Diffraction.

References and links

1. H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, R. B. Doak, F. R. N. C. Maia, A. V. Martin, I. Schlichting, L. Lomb, N. Coppola, R. L. Shoeman, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, L. Foucar, N. Kimmel, G. Weidenspointner, P. Holl, M. Liang, M. Barthelmess, C. Caleman, S. Boutet, M. J. Bogan, J. Krzywinski, C. Bostedt, S. Bajt, L. Gumprecht, B. Rudek, B. Erk, C. Schmidt, A. Hömke, C. Reich, D. Pietschner, L. Strüder, G. Hauser, H. Gorke, J. Ullrich, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kühnel, M. Messerschmidt, J. D. Bozek, S. P. Hau-Riege, M. Frank, C. Y. Hampton, R. G. Sierra, D. Starodub, G. J. Williams, J. Hajdu, N. Timneanu, M. M. Seibert, J. Andreasson, A. Rocker, O. Jönsson, M. Svenda, S. Stern, K. Nass, R. Andritschke, C.-D. Schröter, F. Krasniqi, M. Bott, K. E. Schmidt, X. Wang, I. Grotjohann, J. M. Holton, T. R. M. Barends, R. Neutze, S. Marchesini, R. Fromme, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, I. Andersson, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, and J. C. H. Spence, "Femtosecond X-ray protein nanocrystallography," Nature **470**(7332), 73–77 (2011).
2. S. Boutet, L. Lomb, G. J. Williams, T. R. M. Barends, A. Aquila, R. B. Doak, U. Weierstall, D. P. DePonte, J. Steinbrener, R. L. Shoeman, M. Messerschmidt, A. Barty, T. A. White, S. Kassemeyer, R. A. Kirian, M. M. Seibert, P. A. Montanez, C. Kenney, R. Herbst, P. Hart, J. Pines, G. Haller, S. M. Gruner, H. T. Philipp, M. W. Tate, M. Hromalik, L. J. Koerner, N. van Bakel, J. Morse, W. Ghonsalves, D. Arnlund, M. J. Bogan, C. Caleman, R. Fromme, C. Y. Hampton, M. S. Hunter, L. C. Johansson, G. Katona, C. Kupitz, M. Liang, A. V. Martin, K. Nass, L. Redecke, F. Stellato, N. Timneanu, D. Wang, N. A. Zatsepin, D. Schafer, J. Defever, R. Neutze, P. Fromme, J. C. H. Spence, H. N. Chapman, and I. Schlichting, "High-resolution protein structure determination by serial femtosecond crystallography," Science **337**(6092), 362–364 (2012).
3. M. Frank, D. B. Carlson, M. S. Hunter, G. J. Williams, M. Messerschmidt, N. A. Zatsepin, A. Barty, W. H.

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

Benner, K. Chu, A. T. Graf, S. P. Hau-Riege, R. A. Kirian, C. Padeste, T. Pardini, B. Pedrini, B. Segelke, M. M. Seibert, J. C. H. Spence, C.-J. Tsai, S. M. Lane, X.-D. Li, G. Schertler, S. Boutet, M. Coleman, and J. E. Evans, "Femtosecond X-ray diffraction from two-dimensional protein crystals," IUCrJ **1**(Pt 2), 95–100 (2014).
4. P. Hart, S. Boutet, G. Carini, M. Dubrovin, B. Duda, D. Fritz, G. Haller, R. Herbst, S. Herrmann, C. Kenney, N. Kurita, H. Lemke, M. Messerschmidt, M. Nordby, J. Pines, D. Schafer, M. Swift, M. Weaver, G. Williams, D. Zhu, N. van Bakel, and J. Morse, "The CSPAD megapixel x-ray camera at LCLS," Proc. SPIE **8506**, 85040C (2012).
5. L. Strüder, S. Epp, D. Rolles, R. Hartmann, P. Holl, G. Lutz, H. Soltau, R. Eckart, C. Reich, K. Heinzinger, C. Thamm, A. Rudenko, F. Krasniqi, K.-U. Kühnel, C. Bauer, C.-D. Schröter, R. Moshammer, S. Techert, D. Miessner, M. Porro, O. Hälker, N. Meidinger, N. Kimmel, R. Andritschke, F. Schopper, G. Weidenspointner, A. Ziegler, D. Pietschner, S. Herrmann, U. Pietsch, A. Walenta, W. Leitenberger, C. Bostedt, T. Möller, D. Rupp, M. Adolph, H. Graafsma, H. Hirsemann, K. Gärtner, R. Richter, L. Foucar, R. L. Shoeman, I. Schlichting, and J. Ullrich, "Large-format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources," Nucl. Instrum. Methods Phys. Res., Sect. A **614**, 483–496 (2010).
6. T. Kameshima, S. Ono, T. Kudo, K. Ozaki, Y. Kirihara, K. Kobayashi, Y. Inubushi, M. Yabashi, T. Horigome, A. Holland, K. Holland, D. Burt, H. Murao, and T. Hatsui, "Development of an X-ray pixel detector with multi-port charge-coupled device for X-ray free-electron laser experiments," Rev. Sci. Instrum. **85**(3), 033110 (2014).
7. B. Henrich, J. Becker, R. Dinapoli, P. Goettlicher, H. Graafsma, H. Hirsemann, R. Klanner, H. Krueger, R. Mazzocco, A. Mozzanica, H. Perrey, G. Potdevin, B. Schmitt, X. Shi, A. K. Srivastava, U. Trunk, and C. Youngman, "The adaptive gain integrating pixel detector AGIPD a detector for the European XFEL," Nucl. Instrum. Methods Phys. Res., Sect. A **633**, S11–S14 (2011).
8. C. B. Wunderer, A. Marras, M. Bayer, J. Correa, S. Lange, I. Shevyakov, S. Smoljanin, M. Viti, Q. Xia, M. Zimmer, G. Cautero, A. Gianoncelli, D. Giuressi, R. H. Menk, L. Stebel, H. Yousef, N. Tartoni, J. Marchal, N. Rees, J. Thompson, R. Turchetta, I. Sedgwick, D. Das, B. Marsh, and H. Graafsma, "Percival: an international collaboration to develop a MAPS-based soft X-ray imager," Synchrotron Radiat. News **27**(4), 30–34 (2014).
9. P. Denes, "Two-dimensional imaging detectors for structural biology with X-ray lasers," Philos. Trans. R. Soc. Lond. B Biol. Sci. **369**(1647), 20130334 (2014).
10. A. Barty, R. A. Kirian, F. R. N. C. Maia, M. Hantke, C. H. Yoon, T. A. White, and H. Chapman, "Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data," J. Appl. Cryst. **47**(3), 1118–1131 (2014).
11. T. A. White, R. A. Kirian, A. V. Martin, A. Aquila, K. Nass, A. Barty, and H. N. Chapman, "CrystFEL: a software suite for snapshot serial crystallography," J. Appl. Cryst. **45**(2), 335–341 (2012).
12. J. Hattne, N. Echols, R. Tran, J. Kern, R. J. Gildea, A. S. Brewster, R. Alonso-Mori, C. Glöckner, J. Hellmich, H. Laksmono, R. G. Sierra, B. Lassalle-Kaiser, A. Lampe, G. Han, S. Gul, D. DiFiore, D. Milathianaki, A. R. Fry, A. Miahnahri, W. E. White, D. W. Schafer, M. M. Seibert, J. E. Koglin, D. Sokaras, T. C. Weng, J. Sellberg, M. J. Latimer, P. Glatzel, P. H. Zwart, R. W. Grosse-Kunstleve, M. J. Bogan, M. Messerschmidt, G. J. Williams, S. Boutet, J. Messinger, A. Zouni, J. Yano, U. Bergmann, V. K. Yachandra, P. D. Adams, and N. K. Sauter, "Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers," Nat. Methods **11**(5), 545–548 (2014).
13. D. P. DePonte, U. Weierstall, K. Schmidt, J. Warner, D. Starodub, J. Spence, and R. B. Doak, "Gas dynamic virtual nozzle for generation of microscopic droplet streams," J. Phys. D Appl. Phys. **41**(19), 195505 (2008).
14. R. G. Sierra, H. Laksmono, J. Kern, R. Tran, J. Hattne, R. Alonso-Mori, B. Lassalle-Kaiser, C. Glöckner, J. Hellmich, D. W. Schafer, N. Echols, R. J. Gildea, R. W. Grosse-Kunstleve, J. Sellberg, T. A. McQueen, A. R. Fry, M. M. Messerschmidt, A. Miahnahri, M. M. Seibert, C. Y. Hampton, D. Starodub, N. D. Loh, D. Sokaras, T. C. Weng, P. H. Zwart, P. Glatzel, D. Milathianaki, W. E. White, P. D. Adams, G. J. Williams, S. Boutet, A. Zouni, J. Messinger, N. K. Sauter, U. Bergmann, J. Yano, V. K. Yachandra, and M. J. Bogan, "Nanoflow electrospinning serial femtosecond crystallography," Acta Crystallogr., Sect. D: Biol. Crystallogr. **68**(11), 1584–1587 (2012).
15. U. Weierstall, D. James, C. Wang, T. A. White, D. Wang, W. Liu, J. C. H. Spence, R. Bruce Doak, G. Nelson, P. Fromme, R. Fromme, I. Grotjohann, C. Kupitz, N. A. Zatsepin, H. Liu, S. Basu, D. Wacker, G. W. Han, V. Katritch, S. Boutet, M. Messerschmidt, G. J. Williams, J. E. Koglin, M. Marvin Seibert, M. Klinker, C. Gati, R. L. Shoeman, A. Barty, H. N. Chapman, R. A. Kirian, K. R. Beyerlein, R. C. Stevens, D. Li, S. T. A. Shah, N. Howe, M. Caffrey, and V. Cherezov, "Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography," Nat. Commun. **5**, 3309 (2014).
16. W. Liu, D. Wacker, C. Wang, E. Abola, and V. Cherezov, "Femtosecond crystallography of membrane proteins in the lipidic cubic phase," Philos. Trans. R. Soc. Lond. B Biol. Sci. **369**(1647), 20130314 (2014).
17. A. E. Cohen, S. M. Soltis, A. González, L. Aguila, R. Alonso-Mori, C. O. Barnes, E. L. Baxter, W. Brehmer, A. S. Brewster, A. T. Brunger, G. Calero, J. F. Chang, M. Chollet, P. Ehrensberger, T. L. Eriksson, Y. Feng, J. Hattne, B. Hedman, M. Hollenbeck, J. M. Holton, S. Keable, B. K. Kobilka, E. G. Kovaleva, A. C. Kruse, H. T. Lemke, G. Lin, A. Y. Lyubimov, A. Manglik, I. I. Mathews, S. E. McPhillips, S. Nelson, J. W. Peters, N. K. Sauter, C. A. Smith, J. Song, H. P. Stevenson, Y. Tsai, M. Uervirojnangkoorn, V. Vinetsky, S. Wakatsuki, W. I. Weis, O. A. Zadvornyy, O. B. Zeldin, D. Zhu, and K. O. Hodgson, "Goniometer-based femtosecond crystallography with X-ray free electron lasers," Proc. Natl. Acad. Sci. U.S.A. **111**(48), 17122–17127 (2014).
18. W. Liu, D. Wacker, C. Gati, G. W. Han, D. James, D. Wang, G. Nelson, U. Weierstall, V. Katritch, A. Barty, N. A. Zatsepin, D. Li, M. Messerschmidt, S. Boutet, G. J. Williams, J. E. Koglin, M. M. Seibert, C. Wang, S. T. A. Shah, S. Basu, R. Fromme, C. Kupitz, K. N. Rendek, I. Grotjohann, P. Fromme, R. A. Kirian, K. R. Beyerlein, T. A. White, H. N. Chapman, M. Caffrey, J. C. H. Spence, R. C. Stevens, and V. Cherezov, "Serial femtosecond

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

crystallography of G protein-coupled receptors," Science **342**(6165), 1521–1524 (2013).
19. C. Gati, G. Bourenkov, M. Klinge, D. Rehders, F. Stellato, D. Oberthür, O. Yefanov, B. P. Sommer, S. Mogk, M. Duszenko, C. Betzel, T. R. Schneider, H. N. Chapman, and L. Redecke, "Serial crystallography on in vivo grown microcrystals using synchrotron radiation," IUCrJ **1**(2), 87–94 (2014).
20. F. Stellato, D. Oberthür, M. Liang, R. Bean, C. Gati, O. Yefanov, A. Barty, A. Burkhardt, P. Fischer, L. Galli, R. A. Kirian, J. Meyer, S. Panneerselvam, C. H. Yoon, F. Chervinskii, E. Speller, T. A. White, C. Betzel, A. Meents, and H. N. Chapman, "Room-temperature macromolecular serial crystallography using synchrotron radiation," IUCrJ **1**(4), 204–212 (2014).
21. T. A. White, A. Barty, F. Stellato, J. M. Holton, R. A. Kirian, N. A. Zatsepin, and H. N. Chapman, "Crystallographic data processing for free-electron laser sources," Acta Crystallogr., Sect. D: Biol. Crystallogr. **69**(7), 1231–1240 (2013).
22. T. R. M. Barends, L. Foucar, S. Botha, R. B. Doak, R. L. Shoeman, K. Nass, J. E. Koglin, G. J. Williams, S. Boutet, M. Messerschmidt, and I. Schlichting, "De novo protein crystal structure determination from X-ray free-electron laser data," Nature **505**(7482), 244–247 (2013).
23. M. Schmidt, K. Pande, S. Basu, and J. Tenboer, "Room temperature structures beyond 1.5 Å by serial femtosecond crystallography," Struct. Dyn. **2**(4), 041708 (2015).
24. L. Redecke, K. Nass, D. P. DePonte, T. A. White, D. Rehders, A. Barty, F. Stellato, M. Liang, T. R. M. Barends, S. Boutet, G. J. Williams, M. Messerschmidt, M. M. Seibert, A. Aquila, D. Arnlund, S. Bajt, T. Barth, M. J. Bogan, C. Caleman, T.-C. Chao, R. B. Doak, H. Fleckenstein, M. Frank, R. Fromme, L. Galli, I. Grotjohann, M. S. Hunter, L. C. Johansson, S. Kassemeyer, G. Katona, R. A. Kirian, R. Koopmann, C. Kupitz, L. Lomb, A. V. Martin, S. Mogk, R. Neutze, R. L. Shoeman, J. Steinbrener, N. Timneanu, D. Wang, U. Weierstall, N. A. Zatsepin, J. C. H. Spence, P. Fromme, I. Schlichting, M. Duszenko, C. Betzel, and H. N. Chapman, "Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser," Science **339**(6116), 227–230 (2013).
25. D. Li, L. Vogeley, A. Keogh, N. Howe, P. Stansfeld, M. Sansom, P. Fromme, R. Fromme, B. Basu, I. Grotjohann, C. Kupitz, K. Rendek, U. Weierstall, N. Zatsepin, V. Cherezov, W. Liu, S. Bandaru, N. English, C. Gati, A. Barty, O. Yefanov, H. Chapman, K. Diederichs, M. Messerschmidt, S. Boutet, G. Williams, M. Seibert, and M. Caffrey, "Ternary structure reveals mechanism of a membrane diacylglycerol kinas," Nat. Commun. (posted 24 October 2014, in press).
26. Y. Kang, X. E. Zhou, X. Gao, Y. He, W. Liu, A. Ishchenko, A. Barty, T. A. White, O. Yefanov, G. W. Han, Q. Xu, P. W. de Waa, J. Ke, M. H. Tan, C. Zhang, A. Moeller, G. M. West, B. D. Pascal, N. Van Eps, L. N. Caro, S. A. Vishnivetskiy, R. J. Lee, K. M. Suino-Powell, X. Gu, K. Pal, J. Ma, X. Zhi, S. Boutet, G. J. Williams, M. Messerschmidt, C. Gati, N. A. Zatsepin, D. Wang, D. James, S. Basu, S. Roy-Chowdhury, C. E. Conrad, J. Coe, H. Liu, S. Lisova, C. Kupitz, I. Grotjohann, R. Fromme, Y. Jiang, M. Tan, H. Yang, J. Li, M. Wang, Z. Zheng, D. Li, N. Howe, Y. Zhao, J. Standfuss, K. Diederichs, Y. Dong, C. S. Potter, B. Carragher, M. Caffrey, H. Jiang, H. N. Chapman, J. C. Spence, P. Fromme, U. Weierstall, O. P. Ernst, V. Katritch, V. V. Gurevich, P. R. Griffin, W. L. Hubbell, R. C. Stevens, V. Cherezov, K. Melcher, and H. E. Xu, "Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser," Nature **523**(7562), 561–567 (2015).
27. T. Barends, "De novo protein crystal structure determination from X-ray free-electron laser data," experimental data, http://www.cxidb.org/id-22.html.
28. P. A. Karplus and K. Diederichs, "Linking crystallographic model and data quality," Science **336**(6084), 1030–1033 (2012).
29. S. Boutet, "High-resolution protein structure determination by serial femtosecond crystallography," experimental data, http://www.cxidb.org/id-17.html
30. W. Lui, "Serial femtosecond crystallography of G protein-coupled receptors," experimental data, http://www.cxidb.org/id-21.html

1. Introduction

The advent of X-ray Free Electron Lasers has necessitated the development of integrating X-ray detectors capable of full frame readout at the pulse repetition rate of the X-ray source. This new class of X-ray source produces X-ray pulses of a few tens of femtoseconds duration containing sufficient photons to record the diffraction pattern from individual micron-sized, or smaller, crystals in a single shot [1,2]. Each pulse can measure a different crystal, and diffraction patterns can be measured at the pulse repetition rate, which so far has been up to 120 Hz. Detectors developed for such applications typically consist of several small detector modules tiled together to cover a larger area, for example the CSPAD detector [3,4], pnCCD [5], MPCCD [6], AGIPD [7] and Percival [8] detectors. Other segmented X-ray detector systems are either planned or under development, including detectors for synchrotron applications.

The use of multiple independently moveable modules makes a detector system highly flexible and enables damaged modules to be replaced without having to replace the whole detector. Moreover large-area fast hybrid detectors are currently made from tiles due to the limited size of readout electronics [9]. Individual modules can be arranged in an offset manner

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

surrounding a central hole through which the direct beam can pass [4,5,7], obviating the need for a beamstop, and may be movable so that the overall detector layout can be optimized for each experiment. However, movable modules introduce the necessity of determining the relative locations of each module with respect to the beam-sample interaction point each time the detector configuration is changed. It also makes the process of generating an image from the measured data more complicated. One needs a prescription describing where each pixel of each module is located in real space. Such prescription we call detector geometry, and can be saved in one of many representations including a separate text file (for example the geometry file format used for CrystFEL [10,11]) or lookup table for the coordinates of each pixel (such as the format used for Cheetah [10]).

Accurate determination of detector geometry is crucial for almost every type of experiment, including X-ray crystallography, coherent diffractive imaging (CDI), small angle X-ray scattering (SAXS) and spectroscopy. The required accuracy depends on the particular experiment, but it is not uncommon to require that pixel locations be known to an accuracy of better than one pixel. A simple yet robust detector geometry calibration procedure is therefore required.

Here, we describe an effective method for determining the geometry of a segmented X-ray detector using X-ray serial femtosecond crystallography (SFX) data. The detector geometry is progressively refined by comparing the location of measured Bragg peaks to the location of peaks predicted once the crystal reciprocal lattice vectors have been determined using auto-indexing [11,12]. This method can be applied using either a dedicated calibration sample or, in the case of an SFX experiment, from the serial crystallography data itself. Indeed, this technique can be applied *post facto* for improvement of previously collected serial crystallography data.

2. Defining sensor module positions

Consider the CSPAD detector used at LCLS [4] as an example of a tiled detector system. The 2.2 megapixel version of the CSPAD detector is composed of 32 separate detector modules tiled together, with 8 modules mounted on each of four moveable quadrants so that the central hole can be optimized for each experiment as illustrated in Fig. 1. The question we wish to answer is: what is the spatial relationship between the X-ray interaction point and all pixels in the detector? Although we use the CSPAD detector as an example, the problem of determining the relative positions of sensor elements in a detector is a general one relevant to a range of X-ray area detectors.

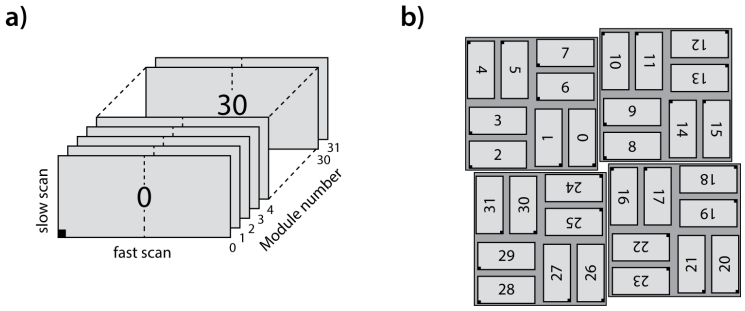

Fig. 1. The 2.2 megapixel CSPAD detector used at LCLS is composed of 32 separate detector elements tiled together to cover a large area with an adjustable central hole. (a) each sensor is read out independently (b) the task at hand is to determine the relative location of sensor elements relative to each other across the whole detector.

We start by assuming that each sensor module consists of an array of pixels whose locations relative to each other are accurately known, ideally to within 1/10th of a pixel. This level of accuracy can typically be assumed from the fabrication of the detector. We refer to these arrays as "panels" or "tiles", which can further be grouped together in "rigid groups". A conventional single module CCD would be considered as a single panel. A full CSPAD

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

detector could be considered as 32 panels grouped into four rigid groups corresponding to the four quadrants, which can be moved relative to one another during an experiment to change the size of the central hole. The positions of the rigid groups will be refined by our procedure, which allows for refinement at different levels of granularity. For example, each panel could be placed in a separate rigid group to refine their positions individually, or every panel could be placed in one single rigid group to refine only the overall sample-detector distance and beam center coordinates. We also assume that the location and orientation of individual modules is roughly known, for example from the mechanical design or from a manual alignment using a powder diffraction pattern.

The most obvious way in which to determine the relative location of sensor modules is to project a known intensity distribution, such as a regular array of spots or fringes, onto the detector. X-ray diffraction from a crystalline sample produces a highly regular array of Bragg peaks, albeit regular in three dimensions rather than two, whose intensity is sufficiently well localized to form a positional reference. The directions of the diffracted beams, and hence the positions where they should form Bragg peaks on the detector, can be accurately calculated once the crystal orientation relative to the X-ray beam is known. We exploit this property of crystal diffraction to accurately refine the detector geometry. The process of geometry refinement allows correction of not only the position of each sensor module (X_{cen}, Y_{cen}), but also the rotation angle of each sensor module (α) and sample-detector distance for either each panel or the whole detector.

3. Data collection and initial processing

The collection of serial crystallography data has been described in detail elsewhere and can be accomplished in a range of ways including using flowing liquid suspensions [13], electrospinning [14], viscous extrusion [15,16] and fixed targets [14,17]. Diffraction data can be measured using either X-ray Free Electron Laser [1,2,15,18] or Synchrotron [9,19,20] sources. The common feature of serial crystallography experiments is that a single diffraction pattern is collected from each of a large number of individual crystals in a random orientation.

An initial estimate of the detector geometry is required for our technique. This can be created manually (for simple cases with small numbers of panels), by using optical measurements of each tile position made using a microscope, or using the detector's response to a visible-light laser spot translated with high precision. We have used the manual approach for some commercial detectors, results from the optical method for the CSPAD detector (where the positions of the tiles were measured with moderate accuracy within each quadrant), and results from the laser method for the MPCCD detector. In many cases, the geometry from a previous experiment can be re-used as the initial estimate for later experiments.

The first step in geometry calibration is to verify the starting detector geometry by forming a "virtual" powder diffraction pattern by summing all measured diffraction patterns [8] and, if necessary, altering the panel positions such that concentric powder diffraction rings line up with one other. This check enables coarse errors in detector geometry to be corrected. However, powder diffraction is insensitive to rotation of each sensor module about the beam center and only weakly sensitive to the size of the gap in the middle of the detector.

Detector frames containing diffraction patterns from individual crystals are first scanned for observable Bragg peaks ("detected peaks") [21], and patterns indexed using the freely available software package CrystFEL [11]. The process of crystal indexing determines the orientation and unit cell parameters of each crystal, enabling the best fit of the locations of Bragg peaks from a particular lattice ("predicted peaks"). We thereby obtain a list of detected Bragg peak positions and a corresponding list of predicted peak locations for each crystal measured. These lists are contained in CrystFEL's output data stream. Our detector geometry refinement program takes this data stream and the initial detector geometry described earlier as input, and produces updated detector geometry as its output.

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

4. Geometry calculation

In our current implementation, three tile motions can be corrected for: in plane translation, in-plane rotation, and displacement along the beam axis. We first correct for in-plane rotation and displacement along the beam, and afterwards for in-plane translation. Correcting the rotations/distance first makes calculation of the translations more robust as the distribution of possible translations is narrowed.

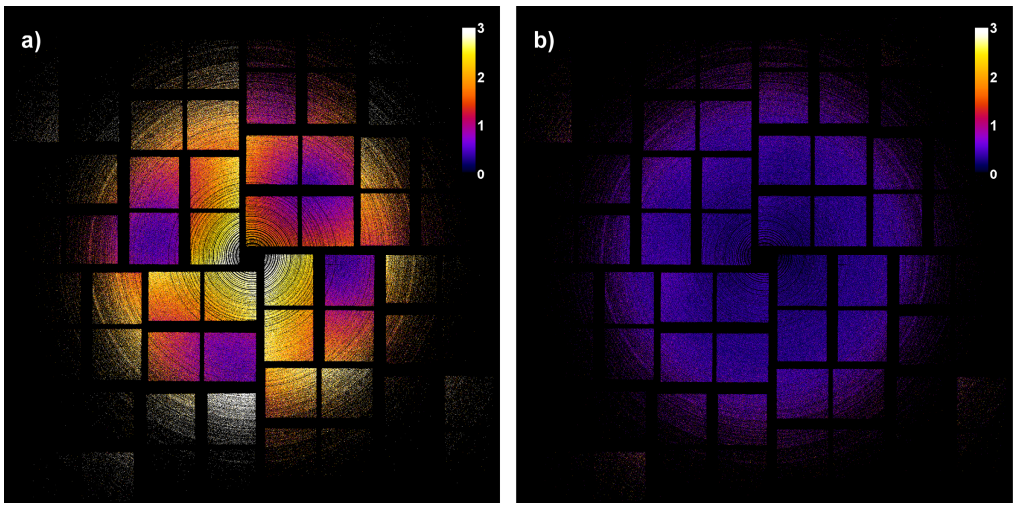

Fig. 2. Error map of the CS-PAD detector (a) before the refinement and (b) after the refinement. Absolute error color-coded in number of pixels. Experiment [22].

4.1. Error map of the difference between predicted and measured peak locations

For each diffraction pattern we have a list of measured peak locations $\{M_i = (x_m, y_m)\}$ and a list of predicted peak locations $\{C_j = (x_c, y_c)\}$, calculated from the image coordinates after application of a geometry estimate we used to obtain indexing results. For each measured peak M_i the list of predicted peaks $\{C_1, C_2, ..., C_n\}$ is searched to find the peak C_j that lies closest to the measured peak location M_i, discarding outliers where no predicted peaks were found within the distance R_{search}. Selection of an appropriate value of R_{search} is made such that it is large enough to find at least one predicted peak close to each measured peak, while not being so large as to select wrong peak. In most cases R_{search} can be set to half of the minimal inter-Bragg distance. The closest predicted peak is assumed to correspond to the same reflection as the observed peak, so that the distance between the measured and predicted peak is the peak position error

$$\begin{aligned} E_P &= M_i - C_j \\ &= (x_m - x_c, y_m - y_c) \\ &= (\Delta x, \Delta y) \end{aligned}$$

The error in peak position E_P is attributed to the pixel P, closest to the detected peak M_i. Repeating this process for each indexed crystal diffraction pattern, we create a list of *npeaks* peak location errors for each pixel P

$$\{E_1^P, E_2^P, ..., E_{npeaks}^P\}$$

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

Once all diffraction patterns have been processed, the median value of the error list for each pixel is used as the error for that pixel

$$E_P = \text{median}\left(\left\{E_1^P, E_2^P, ..., E_{npeaks}^P\right\}\right)$$

We use the median value at each pixel location rather than an average to discard outliers resulting from incorrect pairing of predicted and measured peak locations.

To increase the fidelity of error determination, only pixels with many error measurements (*npeaks*) are considered. We typically require $npeaks \geq 3$, although this number can be increased for experiments with many indexed patterns or decreased if number of indexed patterns is low. Setting *npeaks*>1 reduces the effect of outliers in the case where incorrect predicted peaks are associated with some measured peaks.. Another requirement is that each rigid group has to have at least *npix* pixels for which an error estimate has been made. For the CSPAD, we usually use $npix \geq 100$ for each tile (with total number of pixels 185*388). Again this is done to prevent changing the geometry of a rigid group based on too few measurements, in the event that some of the error estimates could be incorrect.

Figure 2(a) shows the map of the peak position absolute error ($\sqrt{\Delta x^2 + \Delta y^2}$) for one experiment before geometry refinement. Note that some pixels do not have an error estimate due to lack of measurements. Those pixels are just treated as part of a rigid group, while the geometry of each rigid group is refined using the pixels with calculated error. The strategy for the correction of rigid groups with not enough measurements (*npix*) is described later.

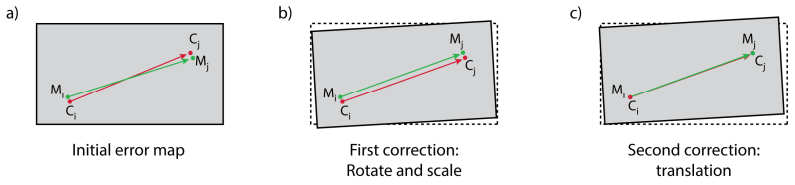

Fig. 3. (a) Vectors between each pair of observed and predicted peak locations are calculated; (b) rotation and scale are corrected first, followed by (c) translation.

4.2 Correction for rotation and scale

For each pixel P on the detector with an error estimate we have both the current location (x, y), calculated from current geometry estimate, and the averaged error between observed and predicted peaks $E_P = (\Delta x, \Delta y)$. For each pair of pixels $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ within a rigid group we form two vectors, the first being the vector between the current location estimate:

$$\begin{aligned} V_1 &= P_2 - P_1 \\ &= (x_2 - x_1, y_2 - y_1) \end{aligned}$$

and the second being the vector between predicted peak positions:

$$\begin{aligned} V_2 &= (P_2 + E_2) - (P_1 + E_1) \\ &= (x_2 + \Delta x_2 - x_1 - \Delta x_1, y_2 + \Delta y_2 - y_1 - \Delta y_1) \end{aligned}$$

The angle α between these two vectors is the error in module rotation given by

$$\alpha = \arccos\left(\frac{V_1 \cdot V_2}{|V_1||V_2|}\right)$$

while the scale factor s is

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

$$s = \frac{|V_1|}{|V_2|}.$$

This calculation is performed for all pairs of pixels within a rigid group. In order to obtain maximum accuracy in angle and scale factor coefficients, we require the distance between pixels to be more than one quarter of the diagonal length of the rigid group. Finally, the median value of all scale and angle coefficients within a rigid group are assigned to the rigid group under analysis, where the median value is used to exclude the effect of outliers.

The rotation angle α determined for a rigid group can be applied to all panels within that rigid group. However, the scale factor can be attributed to three different physical parameters: incorrect wavelength, incorrect sample-detector distance or incorrect unit cell parameters of the crystal. The wavelength is usually known either from monochromator settings (at a synchrotron) or calculated from electron beam and undulator parameters on a shot-to-shot basis (at FELs). The unit cell size of the crystals is usually known for the types of crystals used as a calibration sample. When both these parameters are well known, the scale error is attributed to error in detector distance.

We note that at high scattering angles, such as those used in protein crystallography experiments, the influence of incorrect unit cell and wavelength can be distinguished from incorrect detector distance due to curvature of the Ewald sphere. This can be used for independent determination of the unit cell parameters and sample detector distance, for known wavelength. Such a task is not a part of described algorithm, but it was successfully tested on several experimental data sets.

4.3 Correction for translation

After determination of the distance and angle correction for each panel, the error map is updated by applying the angular and scale corrections to it. From the corrected error map we now extract the translation of each rigid group. The translational correction is the median value of translation error $(\Delta x, \Delta y)$ of all pixels within the rigid group - see Fig. 3.

4.4 Correction for panels with not enough measurements

In some cases, there are too few peaks, or perhaps no peaks at all, within a rigid group to arrive at a sufficiently accurate value of the correction parameters. This is often the case for detector panels located further from the detector center since not all samples scatter well enough to produce Bragg peaks at high scattering angles. Different strategies for refining the geometric corrections of such "empty" panels were analyzed such as duplicating the corrections of the closest panel or a panel with highest number of measurements. When refining the position of each panel of the CSPAD individually, we found it most reliable to take the mean correction factors from the successfully refined tiles in the same quadrant and apply it to the "empty" panels, indicating that the alignment of panels within each quadrant was known quite well from optical metrology. The amount of data needed for the refinement strongly depends on the detector design. We found that we needed several thousand indexed patterns to refine the position of each panel of the CSPAD individually (32 rigid groups), but when the geometry of each quadrant was well characterized using some previously measured data and hence the geometry could be refined only at the level of quadrants (4 rigid groups), the number of indexed patterns was drastically reduced to a couple of hundred.

4.5 Suggested refinement procedure

An optimal sample to refine the detector geometry would fill the entire detector with Bragg peaks. If necessary, the detector could be temporarily moved further back from the interaction region to measure such a calibration data set; this data set could be used to determine the relative positions of all detector tiles before moving the detector back to the desired position for the experiment. The detector center coordinates (x,y) and sample-detector distance in the new position can be further refined from data that does not extend to the edges of the detector

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

while leaving the relative position of detector modules unchanged. It may be desirable to use a well-diffracting calibration sample for cases where the sample of interest diffracts weakly or for experiments that do not produce Bragg peaks.

A typical geometry calibration procedure starts with the best geometry estimate from available metrology data, for example optical metrology or design drawings. Modules are then coarsely aligned using virtual powder diffraction patterns until indexing of patterns is possible. Once diffraction patterns index, module locations can be refined using the procedure described in this paper. Repeating the indexing and geometry refinement steps two to three times is usually adequate for convergence. The number of successfully indexed patterns usually increases after each iteration, progressively improving the fidelity of the geometry refinement.

4.6 Assessment of the geometry correction

After applying all the corrections, the error in each pixel location is usually drastically reduced to a sub-pixel value, as shown in Fig. 2(b). An improvement in peak prediction can also be seen by inspecting individual diffraction patterns, comparing the location of observed and predicted peaks as shown in Fig. 4. The residual RMS error for each rigid group calculated for the initial and final geometry Fig. 2 can be used as a metric for the quality of the geometry estimate, both for each individual rigid group and for the detector geometry as a whole. These two error maps are very useful guides as to whether the algorithm has succeeded.

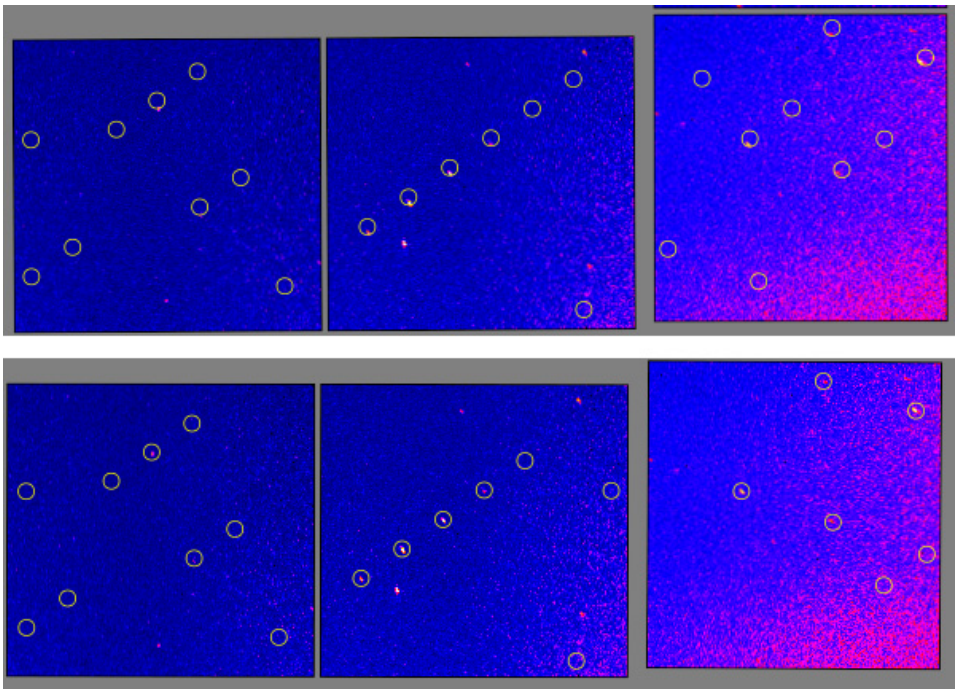

Fig. 4. Enlarged section of a protein crystal diffraction pattern before (top) and after (bottom) geometry refinement. The dots are actual measured peaks and circles represent the positions of predicted peaks.

5. Improvement in SFX data quality

Accurate definition of detector geometry has a strong effect on both indexing rate and final data quality. For example, a significant proportion of the improvement in data quality reported for one data set [23] is attributed to improvements in detector geometry specification made using the procedures described here. In order to isolate the effects of geometry alone on the

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

quality of SFX data, we processed several data sets with CrystFEL, using identical protocols, before and after application our geometry correction procedure. The first important metric is the indexing rate, i.e. the fraction of diffraction patterns for which the auto-indexing procedure succeeds.

Table 1. The improvement of indexing rate after the geometry refinement.

Protein	Number of crystal hits	Indexed before geometry refinement	Indexed after geometry refinement
Gd:Lysozyme [22] (4N5R)	391,000	42% (164,000)	56% (218,000)
Cathepsin B [24] (4HWY)	358,000	56% (201,000)	59% (210,000)
DgkA [25]	180,000	60% (109,000)	78% (140,500)
Rhodopsin-Arrestin [26] (4ZWJ)	11,000	20% (2,200)	80% (8,800)

For data from two published data sets, Gd:Lysozyme [22] (Dataset 1, Ref. [27]) and Cathepsin B [24], the indexing rate improved noticeably after geometry refinement. For Gd:Lysozyme [22], the indexing rate increased from 42% to 56%. Similarly for Cathepsin B crystals [24] an improvement from 56% to 59% in indexing rate was observed. The improvement is even more noticeable for crystals of large membrane proteins: for DgkA [25] the indexing rate increased from 60% to 78%, while for Rhodopsin[25] the indexing rate improved from 20% to 80% due to improvements in detector geometry alone. These results are summarised in Table 1. In all cases indexing rates are significantly higher when a refined geometry is used.

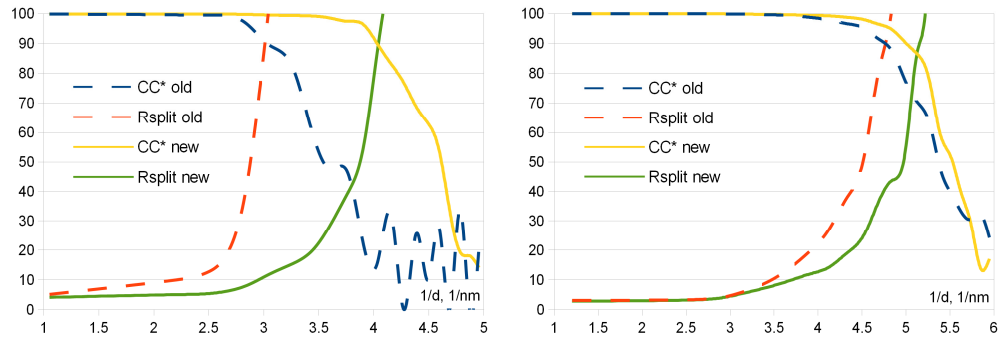

Fig. 5. CC* and Rsplit plots for DgkA (left) and CatB (right) before and after refinement.

Post-merging data quality metrics CC* [28] and R_{split} [11] are also improved by the use of a refined geometry. Figure 5 shows the resolution dependence of CC* and R_{split} before and after the refinement for two different protein crystal samples. It is interesting to note that even though the indexing rate for Cathepsin B crystals did not change appreciably (see Table 1), the resolution as determined by CC* was still noticeably improved. We attribute this to the fact that the predicted peak locations match the actual peak locations more accurately when a refined detector geometry is used, reducing the probability that a real Bragg reflection could be missed by the integration procedure. The only difference between processing of these two data sets was the use of a refined geometry specification, clearly identifying the determination of accurate detector geometry as a major contributor to high resolution SFX data quality (or lack thereof).

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

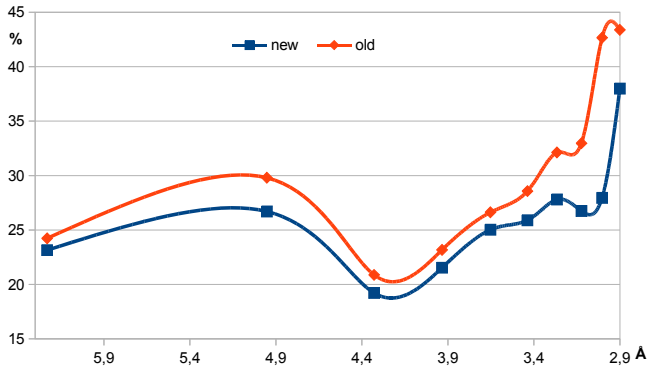

Fig. 6. Graph of Rfree against resolution (d) for 5HT2B before (red) and after (blue) geometry refinement.

More importantly, the structure of biological specimens can be reconstructed more accurately using data processed with a refined detector geometry. Reprocessing of published Lysozyme data [2] (Dataset 2, Ref. [29]) to the same resolution as before (1.9 Å) after geometry refinement and using a more recent version of CrystFEL (0.5.2 released in 2014 vs an old version from 2011) showed Rfree = 19.7%, an improvement from 22.9% in the original paper, while the data resolution was improved from 1.9 Å to 1.7 Å (overall Rfree = 20.5% for the refinement at 1.7 Å). As this improvement combined the effects of both geometry and software version, we performed one more test using crystals of a G-protein coupled receptor, 5HT2B [18] (Dataset 3, Ref. [30]) considering the effects of geometry refinement alone. Refinement using Phenix.refine to a resolution of 2.9 Å led to an improvement of Rfree from 26% to 23.8% after refining the detector geometry. The difference between the two processed data sets was only the corrections to the detector geometry, with Fig. 6 showing improvement for all resolution shells.

Conclusions

We have described a simple scheme by which the detector module geometry can be optimized using serial crystallography data, resulting in sub-pixel accuracy in the determination of the positions and orientations of detector modules. Geometry refinement alone significantly improved high resolution SFX data, identifying this as a major contributor to overall SFX data quality. The technique is general, and can be applied to a wide range of segmented detectors used, or planned for use, at synchrotron or X-ray free electron laser facilities. Indeed the calibration described here can improve the quality of protein crystallography data obtained using synchrotron detectors that nominally have a well-defined initial geometry, by refining the sample-detector distance and beam center, as tested using data measured at Petra III and ESRF. The method can be applied to any SFX data set, including those collected in the past without the use of a dedicated calibration sample. To date, we have applied this procedure to over 30 different data sets and the results were always improved. The optimization procedure has been implemented in the CrystFEL software suite as the program “geoptimiser”, available starting from CrystFEL version 0.6.0. Shot-to-shot jitter of an FEL beam position and phase can introduce sub-pixel shift of each diffraction pattern – this effect is corrected by shifting the center of each shot by CrystFEL. Future improvements of “geoptimiser” will include refinement of the out-of-plane rotation of the rigid groups, as well as improvement in user friendliness, like better automatic diagnostics if the refinement procedure succeeded and automatic recommendations for user which parameters to change.

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015

Acknowledgments

We acknowledge support of the Helmholtz Association through project- oriented funds and the Virtual Institute “Dynamic Pathways in Multidimensional Landscapes”; the DFG through the Gottfried Wilhelm Leibniz Program; the European Research Council under the European Union's Seventh Framework Programme ERC Synergy Grant 609920 “AXSIS” and Marie Curie FP7-PEOPLE-2011-ITN Grant 317079 “Nanomem”; the BMBF through Projects 05E13GU1 and 05K2012; the Graduate College “GRK 1355” at the University of Hamburg, the BioXFEL Science Technology Center (award 1231306); the US National Institutes of Health (NIH), National Institute of General Medical Sciences grants R01 GM095583, U54 GM094599, and R01 GM097463. Parts of the sample injector used at LCLS for this research was funded by the NIH, P41GM103393, formerly P41RR001209. Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

#245734 Received 17 Jul 2015; revised 18 Sep 2015; accepted 9 Oct 2015; published 22 Oct 2015