Time-Resolved Serial Femtosecond Crystallography at the European X-ray Free Electron Laser

Marius Schmidt (smarius@uwm.edu)
UW-Milwaukee
Suraj Pandey
UW-Milwaukee
Adrian Mancuso
EuXFEL
Richard Bean
EuXFEL

Method Article

Keywords: time resolved crystallography, photoactive yellow protein, Free Electron Laser

Posted Date: November 20th, 2019

DOI: https://doi.org/10.21203/rs.2.14634/v1

Abstract

This protocol introduces step by step into the collection of time resolved crystallographic data and their analysis at the European Free Electron Laser.

Introduction

The European XFEL produces intense X-ray bursts with MHz pulse repetition rates. These bursts can be exploited to rapidly collect time-resolved crystallographic data from femtoseconds to seconds (Pandey et al., 2019).

Reagents

Reagents to grow E. coli cultures

Reagents for overexpression and purification of protein

Reagents to produce protein microcrystals

Synthetic mineral oil to co-flow with protein slurry

Equipment

European XFEL

High repetition rate laser

Adaptive Gain Integrating Pixel Detector

Data Storage Facility

PC Linux workstation to analyze the data

Procedure

A. PYP overexpression and Purification

1. Use 50 µL of of glycerol stock of PYP expressing, recombinant *E. coli* to inoculate 100mL of ampicillin containing LB media in a 250ml flask. Grow overnight at 37^{o}C.

2. Prepare four 2 L flasks with 1 L ampicillin containing LB media and inoculate with ~20 mL, each from the 100 mL overnight culture. Let the culture grow at 37^{o}C until the OD_{600} reaches 0.6-0.8. Induce the cultures with IPTG (1mmol/L final concentration) and shake at 16 ^{o}C for 20 – 24 hrs.

3. Harvest the cells by centrifugation. Re-suspend the cells (20 g) in (50 mL) lysis buffer (20mmol/L HEPES,200mmol/L NaCl,5mmol/L Imidazole, pH 7.4) and add lysozyme and protease inhibitor. Sonicate 5 times for 1 min. Clarify the cell debris by centrifugation at 4 ^{o}C. Pour the supernatant into a beaker and slowly add activated pCA anhydride with stirring. Stir the beaker overnight at 4^{o}C.

Activated pCA anhydride:

Prepare a DCC solution by dissolving 6.25gm of DCC in 2.5ml DMF

Dissolve 0.328gm of p-coumaric acid in 7ml DMF.

Add 1ml of DCC solution into p-coumaric solution prepared in ii. and stir on ice in a closed container overnight.

4. Centrifuge the protein-pCA mixture to remove excess pCA. Load the supernatant on the column containing 40 mL of pre-charged and pre-calibrated Ni^{+2} resin. After loading the protein, rinse the column with Lysis buffer (see above) followed by wash buffer (20mmol/L HEPES,200mmol/L NaCl,10mmol/L Imidazole, pH 7.4). Elute the column with elution buffer (same as wash buffer but with 200 mmol/L imidazole) and collect the elute in a fresh breaker. Dialyze the elute overnight in dialysis buffer (20mmol/L HEPES, 200mmol/L NaCl, pH 7.4) to remove excess imidazole.

5. Concentrate the elute to around 10 mL and the check the protein concentration at 446nm (OD_{446}/45000 x 14700 x dilution factor). Add 20 µL enterokinase (New England Biolabs) and rock the mixture for 3 days at room temperature to cleave the His tag.

6. After completion of the enterokinase reaction, centrifuge the mixture to remove potential precipitate (if any). Re-load the protein onto the Nickel column and elute with lysis buffer (see above). This time collect the yellowish flow-through. This is the PYP without the His-tag. Wash the un-cleaved PYP from the column using elution buffer(see above) and freeze it after dialysis (see above) for later cleavage. Concentrate the HIS tag free PYP to 10 ml.

7. Purify the PYP by Anion Exchange Chromatography using a FPLC system. Elute the PYP with a NaCl gradient (25 mmol/L Tris pH 8 without NaCl to 1 mol/L NaCl). Collect all the tubes with a yellow solution. Measure absorption spectra for each of the tube. Pool those tubes where the ratio of the absorption at 446nm versus that at 280nm is at least 2. Dialyze the combined protein in storage buffer (pH 6.0, 50mmol/L sodium citrate).

8. Next day, concentrate the protein to 30-50mg/ml, aliquot into 400uL portions, and sterile filter using spin filters. Freeze at -80^{o}C.

9. Use a dry nitrogen shipper to send the frozen protein tubes to the European XFEL for onsite crystallization.

B. Protein Crystallization

1. Concentrate the protein to 100mg/ml and combine several tubes to a final volume of 2-3 ml.

2. Under vigorous stirring, add 4mol/L Sodium Malonate, pH 7 (around 14 mL) to the protein to a final concentration of 3.3 mol/L.

3. Stir the suspension in a closed vial for around 8 hrs and let it rest for 24 hrs at room temperature.

4. Centrifuge the slurry (in 15 mL conical tubes) and the microcrystals will swim up. Exchange the clear solution below the floating microcrystals with (7 mL) of 2.8 mol/L Sodium Malonate, pH 7.

5. Check crystal concentration and size homogeneity with a Neubauer cell-counting chamber under a microscope (400 x). It should be about 10^{10} crystals/mL.

6. Finally, filter the crystal slurry twice through a 10um steel filter. The crystals are ready to be loaded to a reservoir for injection.

C. Reservoir Loading and Reservoir Mounting

The protein slurry is loaded in a 2 mL reservoir. The reservoir has to be prepared so that no air is in between the Teflon plunger and the water line. The other end is connected to the protein line. By pressing on the water line with a 10 mL syringe filled with water, protein solution is pressed out until it leaks out of the protein line.

The reservoir is mounted on top of the beamline vacuum chamber and connected with the water line to an HPLC, and with the protein line to the injector nozzle capillary.

D. Injection

Injection of the microcrystalline slurry into the vacuum is achieved with a Gas Dynamic Virtual Nozzle (GDVN) mounted on a nozzle rod that reaches from ambient pressure into the vacuum chamber. Synthetic mineral oil (5% - 30%) is provided through a T-junction outside the nozzle rod to support injection and to improve the jetting behavior. Flow rates are determined by the maximum pressure allowed by the system. Using a 75 µm inner diameter nozzle capillary, flow rates should be at least 35 µL/min to cope with the high repetition rate X-ray pulses.

E. Instrument Setup – Upstream Interaction Region

This section of the protocol describes pre-experiment preparation steps and those taken at the beginning of, or during an X-ray beam-shift of 12 hours.

Pre-experiment preparation

1. Upstream interaction set up for injection experiment. Catcher, shroud (with entrance aperture, exit cone and laser windows) installed in sample chamber.

2. Establish accelerator bunch pattern to suit sample jetting behaviour

3. Verify X-ray focus at sample position by optical imaging of 20 µm thick Ce:YAG mounted to nozzle rod.

For each beam shift

1. Align SASE1 beamline offset mirrors to the recorded SPB/SFX instrument settings.

2. Verify the beam position against the aperture set by the entrance slits to the micron-scale KB mirror system and make adjustments with the translation and pitch of the second (downstream) offset mirror.

3. Verify beam position at the entrance to the interaction region where beam position is very sensitive to micron-mirror pitch movements and therefore assures focus at the sample.

4. Establish water jet and establish jet—X-ray overlap (as viewed by jet explosion) in the side-view interaction microscope

5. Verify jet—pump-laser overlap by reference to jet—X-ray overlap in the side-view interaction microscope

The upstream interaction region chamber is prepared for laser-pump XFEL probe experiments. The central shroud holds the sample injection GVDN nozzle, providing a degree of differential pumping and minimising sample contamination in the larger chamber. X-ray beam enters through an aperture in the upstream window, impinging on the sample jet, exiting through a cone designed to maximise the diffraction angle observable at the AGIPD detector. Optical laser beam is transmitted into the chamber from the instrument laser hutch before being directed and focused to the interaction point, almost perpendicular to the X-ray beam through additional windows in the shroud with wave-length appropriate anti-reflection coatings.

F. Laser Alignment

1. Pre-experiment preparation. Establish spatial and temporal overlap (time zero) of X-ray beam and pump laser at sample position with 100 mm thick $SrTiO_3$ substrate mounted onto nozzle rod. Establishes relative timing to ~0.1ps

2. Establish relative 'time zero' reading at the 'Pulse Arrival Time Monitor (PAM)' at the already established timing settings. Giving an easily verifiable timing signal accurate to within ~20 fs.

For each beam shift

1. Adjust laser phase-shifter timing to previously established SPB/SFX settings

2. Verify laser timing with PAM reading.

3. Verify required laser pulse pattern relative to X-ray pulse pattern, and delay relative to X-ray pulse.

Figure 2 shows oscilloscope traces from the X-ray pulse pattern (magenta), pump laser pulse pattern (cyan) and a 10Hz trigger at the position of the 1^{st} X-ray pulse (yellow). The yellow pulse is not a pre-trigger, but synchronized with the first X-ray pulse of the train. It is simply a 10 Hz (fixed) trigger that we use as a reference for the oscilloscope.

G. Data Collection

Diffraction data are recorded with the AGIPD 1M. The AGIPD internal veto pattern configured to match the delivered bunch pattern. For crystallographic experiments, the detector is positioned with the front face protruding through the large downstream gate-valve to capture the highest diffraction angles.

H. Data Processing at EuXFEL

Data processing requires separating diffraction patterns with Bragg reflections from blanks without Bragg reflections. Diffraction patterns with Bragg reflections need to be indexed and the intensities extracted. Cheetah (Barty et al, 2014) is used for hitfinding and with CrystFEL (White et. al, 2012) all other steps toward a dataset can be performed.

Cheetah

Two scripts preprocess-dk and quick_calib.pro are used for the calculation of the dark calibration file. Both scripts are available onsite at the EuXFEL.

Run “preprocess-dk hg-run mg-run lg-run” in a Linux terminal where hg-run, mg-run and lg-run are run numbers for high gain, medium gain and low gain. This will create a folder.

Use IDL to execute quick_calib.pro and select the “dark_joined_constants_xfel.h5” file (generated in i.)

A folder with 16 h5 files and 16 image files will be created.

Use the file “Cheetah-AGIPD00-calib.h5” as a dark calibration file in the “ini” file of Cheetah.

Use Cheetah for Hit-finding with peak finding algorithm number 8 with parameters minSNR=8, minADC=200, minPixel = 1 and minPeaks=25 as also described in Wiedorn et al, 2018.

CrystFEL is used for all subsequent steps.

Choose XGANDALF from CrystFEL suite for indexing. The process is described in detail in the CrystFEL tutorial: http://www.desy.de/~twhite/crystfel/tutorial.html.

Optimize the detector geometry iteratively using geoptimiser in CrystFEL as described: http://www.desy.de/~twhite/crystfel/manual-geoptimiser.html

Use the ‘ambigator’ module in CrystFEL to remove indexing ambiguity of the PYP crystal spacegroup $P6_3$.

Use the ‘partialator’ module in CrysFEL to scale all data, and separate the data based on the pulse-ID into datasets at different time-delays, see also http://www.desy.de/~twhite/crystfel/tutorial.html

I. Data Analysis and Structure Determination

1. Data Sets.

After the TR-SFX data sets of both, the (dark state) reference state and those collected a time-delay t after the reaction initiation are available. Each data set consists of a list of Bragg reflections with indices h,k,l and their associated intensities, and an estimated of their measurement errors. The intensities need to be converted to amplitudes by using the ccp4 program ‘truncate’.

2. Difference Amplitudes and Difference Maps

First, calculate structure factor amplitudes (Fc_ref) from a very well refined reference (dark state) model. Scale the observed reference structure factor amplitudes (Fo_ref) to the Fc_ref using proper scaling software provided by the ccp4 suite of programs. This brings the Fo_ref to the absolute scale. Scale the time-resolved TR-SFX amplitudes (Fo_t) to the reference amplitudes as accurately as possible. Now, both

the Fo_t and the Fo_ref are on the absolute scale. Calculate difference amplitudes, deltaF_obs = Fo_t - Fo_ref. From these, calculate a weighted difference map (deltaRho_obs) using phases from the reference model (Schmidt, 2019, Pandey et al., 2019).

3. Estimate of the Population Transfer from Difference Maps

The extent of reaction initiation, or the population transfer into the reaction, can be low, on the order of 5 - 10% for many TR-SFX experiments. The smaller this extent, the smaller the signal to noise ration of the difference features. As a rule of thumb, 10 sigma difference features refer to about a 15% population transfer in most TR crystallographic experiments. Since this sigma level is also dependent on the experimental noise in the amplitudes, a more accurate estimate is urgently needed. Extrapolated maps provide a way such an estimate. Extrapolated maps are calculated by extrapolated structure factors which are obtained by adding a multiple N of the deltaF_obs to the Fc_ref. In an extrapolated map, the electron density of the reference state disappears at positions where there is strong negative difference density in the difference map. If N is too large, too much negative density is added, and strong negative density in the extrapolated map becomes apparent at all positions of strong negative density in the difference map. This can be utilized in a semi-automatic way to determine a characteristic, optimal N_c. The method is outlined by Schmidt, 2019, and described in detail by Pande et al., 2019. Software is available from Pande et al., 2019. In short, the negative electron density in extrapolated maps is integrated within a spherical volume of interest, and the result plotted as a function of N. The integrated negative electron density values start to diverge at N_c. For best results, the extrapolated maps are calculated with the F000 term included. The F000 term is usually available from the reference model and equals to the total electron count in the unit cell. N_c is related to the population transfer as 100/N x 2 in %. E.g. if N_c = 20, the population transfer is 10%. The factor of tow arises from the difference-Fourier approximation.

4. Initial Structural Refinement of a Model against Time-Resolved Data.

An extrapolated electron density map is calculated using the characteristic N_c. The model of the reference state (M_ref) is displayed with the extrapolated map in ‘coot’. A stepped refinement in coot optimizes the agreement of the atomic positions of the model with the extrapolated map. A new model is generated this way. The torsional restraint in the stepped refinement should be switched off. One may want to play with switching on and off other restraints. As a result, when the structural changes are smallish, the new model is obtained fully automatically. If the structural changes are larger, the initial model has to be altered by hand as in conventional structure determination after molecular replacement. For photoactive yellow protein (PYP), this approach generates a model fully automatically on time-scales < 100 ps, where the structural changes are smaller. For PYP, initial processes of the trans-to-cis isomerization are automatically modeled and a time-resolved structure (M_t) is determined.

5. Calculated Difference Electron Density

From M_t and M_ref a difference map (deltaRho_c) with difference amplitudes and difference phases (deltaPhi_c) is calculated. Major features in the difference map deltaRho_c should match to the observed difference map, deltaRho_obs. This is the major criterion to justify that the extrapolated map refinement has been successful.

6. Phased extrapolated maps and Final Model

The deltaPhi_c obtained from the previous step can be combined with the deltaF_obs to calculate phased extrapolated maps by adding N times the deltaF_obs to the Fc_ref as vectors in the complex plane. The extrapolated amplitudes calculated this way can be used for further (final) structural refinement of M_t. R-factors from this refinement are usually acceptable and conventional 2mFo-DFc maps calculated after refinement are very clean. Software and Linux shell scripts to support steps 1 to 3 and 4 to 5 are provided by Pande et al., 2019. This concludes the protocol.

Troubleshooting

GDVN nozzle clogging: replace nozzle, replace tubing and capillaries.

Schmutz on nozzle: impregnate nozzle with Novec, inject 5% of synthetic oil, move the X-ray interaction region as far away as possible from the nozzle.

Data acquisition problems: restart DAQ, install an earlier version

Time Taken

4 shifts of 12 hours

Anticipated Results

structures along the reaction coordinate, and chemical kinetic mechanism from the same set of X-ray data.

References

S. Pandey , R. Bean, T. Sato, I. Poudyal, J. Bielecki, J. Cruz Villarreal, O. Yefanov, V. Mariani, T. A. White, C. Kupitz, M. Hunter, M. H. Abdellatif, S. Bajt, V. Bondar, A. Echelmeier, D. Doppler, M. Emons, M. Frank, R. Fromme, Y. Gevorkov, G. Giovanetti, M. Jiang, D. Kim, Y. Kim, H. Kirkwood, A. Klimovskaia, J. Knoska, F. H. M. Koua, R. Letrun, S. Lisova, L. Maia, V. Mazalova, D. Meza, T. Michelat, A. Ourmazd, G. Palmer, M. Ramilli, R. Schubert, P. Schwander, A. Silenzi, J. Sztuk-Dambietz, A. Tolstikova, H. N. Chapman, A. Ros, A. Barty, P. Fromme, A. P. Mancuso, *M. Schmidt* (2019) Time-Resolved Serial Femtosecond Crystallography at the European XFEL. Nature Methods (revised). #contributed equally.

M. Schmidt(2019) Time-Resolved Macromolecular Crystallography at Pulsed X-ray Sources, International Journal of Molecular Sciences, 20, 1401; doi:10.3390/ijms20061401.

M. O. Wiedorn, D. Oberthur, R. Bean, R. Schubert, N. Werner, B. Abbey, M. Aepfelbacher, L. Adriano, A. Allahgholi, N. Al-Qudami, J. Andreasson, S. Aplin, S. Awel, K. Ayyer, S. Bajt, I. Barak, S. Bari, J. Bielecki, S. Botha, D. Boukhelef, W. Brehm, S. Brockhauser, I. Cheviakov, M. A. Coleman, F. Cruz-Mazo, C. Danilevski, C. Darmanin, R. B. Doak, M. Domaracky, K. Dorner, Y. Du, H. Fangohr, H. Fleckenstein, M. Frank, P. Fromme, A. M. Ganan-Calvo, Y. Gevorkov, K. Giewekemeyer, H. M. Ginn, H. Graafsma, R. Graceffa, D. Greiffenberg, L. Gumprecht, P. Gottlicher, J. Hajdu, S. Hauf, M. Heymann, S. Holmes, D. A. Horke, M. S. Hunter, S. Imlau, A. Kaukher, Y. Kim, A. Klyuev, J. Knoska, B. Kobe, M. Kuhn, C. Kupitz, J. Kupper, J. M. Lahey-Rudolph, T. Laurus, K. Le Cong, R. Letrun, P. L. Xavier, L. Maia, F. Maia, V. Mariani, M. Messerschmidt, M. Metz, D. Mezza, T. Michelat, G. Mills, D. C. F. Monteiro, A. Morgan, K. Muhlig, A. Munke, A. Munnich, J. Nette, K. A. Nugent, T. Nuguid, A. M. Orville, S. Pandey, G. Pena, P. Villanueva-Perez, J. Poehlsen, G. Previtali, L. Redecke, W. M. Riekehr, H. Rohde, A. Round, T. Safenreiter, I. Sarrou, T. Sato, *M. Schmidt*, B. Schmitt, R. Schonherr, J. Schulz, J. A. Sellberg, M. M. Seibert, C. Seuring, M. L. Shelby, R. L. Shoeman, M. Sikorski, A. Silenzi, C. A. Stan, X. Shi, S. Stern, J. Sztuk-Dambietz, J. Szuba, A. Tolstikova, M. Trebbin, U. Trunk, P. Vagovic, T. Ve, B. Weinhausen, T. A. White, K. Wrona, C. Xu, O. Yefanov, N. Zatsepin, J. Zhang, M. Perbandt, A. P. Mancuso, C. Betzel, H. Chapman, A. Barty (2018) Megahertz serial crystallography, Nature Communications 9 (1), 4025., DOI: 10.1038/s41467-018-06156-7.

K. Pande, C. D. M. Hutchison, G. Groenhof, A. Aquila, J. S. Robinson, J. Tenboer, S. Basu, S. Boutet, D. P. DePonte, M. Liang, T. A. White, N. A. Zatsepin, O. Yefanov, D. Morozov, D. Oberthuer, C. Gati, G. Subramanian, D. James, Y. Zhao, J. Koralek, J. Brayshaw, C. Kupitz, C. Conrad, S. Roy-Chowdhury, J. D. Coe, M. Metz, P. L. Xavier, T. E. Grant, J. E. Koglin, G. Ketawala, R. Fromme, V. Šrajer, R. Henning, J. C. H. Spence, A. Ourmazd, P. Schwander, U. Weierstall, M. Frank, P. Fromme, A. Barty, H. N. Chapman, K. Moffat, J. J. van Thor, *M. Schmidt*(2016) Femtosecond Structural Dynamics Initiates the Trans/Cis Isomerization in Photoactive Yellow Protein, Science 352, 725-729.

J. Tenboer, S. Basu, N. Zatsepin, K. Pande, D. Milathianaki, M. Frank, M. Hunter, S. Boutet, G. Williams, J. E. Koglin, D. Oberthuer, M. Heymann, C. Kupitz, C. Conrad, J. Coe, S. Roy-Chowdhury, U. Weierstall, D. James, D. Wang, T. Grant, A. Barty, O. Yefanov, J. Scales, C. Gati, C. Seuring, V. Srajer, R. Henning, P. Schwander, R. Fromme, A. Ourmazd, K. Moffat, J. Van Thor, J. H. C. Spence, P. Fromme, H. Chapman, *M. Schmidt* (2014) Time Resolved Serial Femtosecond Crystallography Captures High Resolution Intermediates of Photoactive Yellow Protein, Science 346 (6214) 1242-1246.

Acknowledgements

Support of BioXFEL (NSF-1231306), a National Science Foundation Science and Technology center is acknowledged.

Figures

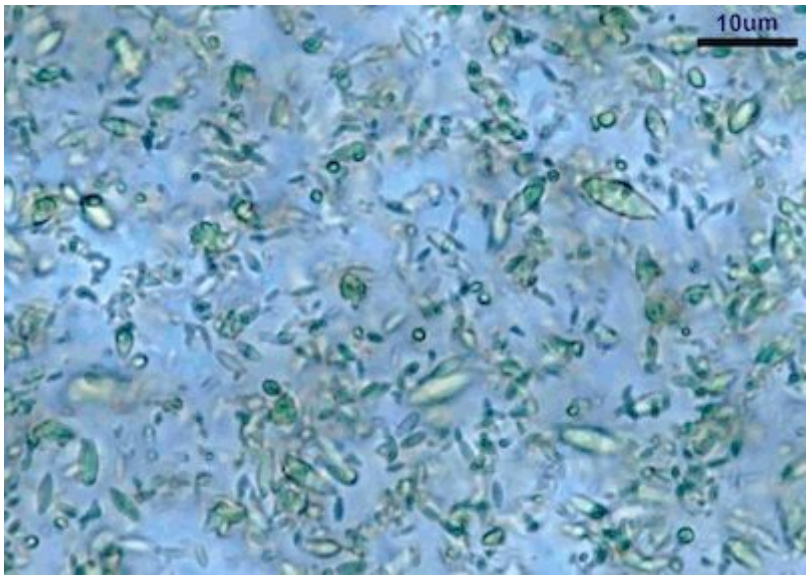

Figure 1

Microcrystals of PYP

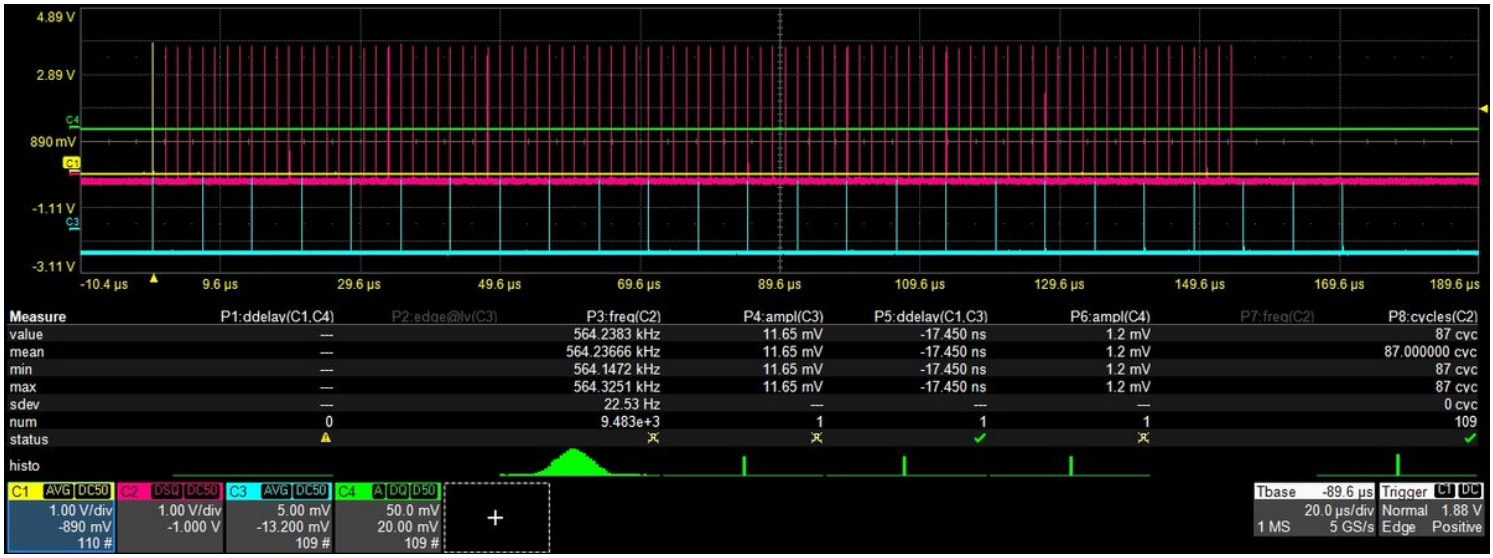

Figure 2

Oscilloscope trace to determine laser to X-ray pulse synchronization.