Fixed target FEL crystallography for phasing at 7.2 KeV

Chitra Rajendran[1], Ana Gonzalez[2], Aaron Brewster[§], Nicholas K Sauter[§]

[1]University of Regensburg, Institute of Biophysics and Physical Biochemistry, Regensburg, Germany
[2]MaxIV Laboratory, Sweden
[§] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

Abstract

Fixed target sample holders for serial crystallography allow precise positioning of crystals into the X-ray interaction for optimization of the hit ray. In addition, they offer the possibility of collecting additional shots from the same crystal if the first X-ray pulse has not destroyed the crystal and for large crystals it allows the possibility of translating the crystal and collecting the data, which could be very useful for collecting redundant and isomorphic data. Apart from this, to test the potential of inverse beam for phasing serial crystallography data we collected data from cryocooled crystals of multiheme protein H1 at the LCLS beamline XPP, using a defocused beam (50 x 50 microns2) by carrying out the experiment at cryotemperatures we expected to extend the lifetime of the crystal enough to tolerate two 40 femtosecond exposures to each crystal at full beam intensity, as had already been demonstrated by Cohen et al (2014) for myoglobin crystals. Combining the first shot data would result in a practically radiation damage free data containing some anomalous differences. The second shot illuminates the same crystal volume and should yield a partially damaged data set. Combining both data sets would, on the one hand, increase the multiplicity of the anomalous measurements, while, on the other hand, degrading the anomalous signal. By analysing the data and the anomalous signal it should be possible to determine if merging of the two data sets is advantageous or detrimental.

Aim of the experiment

To test the potential of inverse beam for phasing serial crystallography data we collected data from cryocooled multiheme protein H1 crystals at the LCLS beamline XPP, using a defocused beam (50 x 50 microns2).

Crystal preparation

Individual H1 crystals were harvested from hanging drop using a nylon loop and placed inside a drop of Paratone-N, and then placed inside a high-density sample mounting grid port (Figure 1)

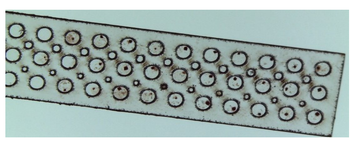

Figure 1 High density sample mounting grid filled with crystals of H1. The grid layout contains 75 ports of which 15 ports are of 125-µm diameter, 15 ports are of 200-µm diameter, and 45 ports are of 400-µm diameter

Data Collection

Two still shots were collected with unattenuated 50 micron beam at a nominal energy of 7.2 keV. At this energy setting, we expected to avoid collecting data below the iron absorption edge owing to the uncertainty in the energy of different XFEL pulses. For the second shot, the crystal was rotated by 180 degrees with respect to the initial orientation. The first "still" shot was followed by a succession of shots in burst mode, while the crystal rotates over a narrow angular wedge. These shots in burst mode could be used for autoindexing and diffraction characterization purposes, although they were not used for analysis of the anomalous signal. We were able to collect 200 images out of which only 69 images of the first shot and 70 images of the second shot were identified by *PRIME*.

Table 1: Crystallographic statistics

	First shot	Second shot	Both	Both (odd images)
Space group	H32			
Cell dimensions (Å)	a=b=135.28±0.78, c=213.09±1.96	a=b=135.35±0.73, c=213.00±1.58	a=b=135.30±0.79, c=213.05±1.82	a=b=135.36±0.78, c=212.97±1.70
Cell angles (°)	α=β=90°,γ=120°			
Wavelength (Å)	1.721±0.001			
Number of collected frames	107	93	200	101
Number of indexed frames	73	75	148	73
Number of frames accepted by prime	69	70	142	69
Resolution (Å) *	78.75-2.5 (2.54-2.50)	78.78-2.5 (2.54-2.50)	78.76-2.5 (2.54-2.50)	78.72-2.5 (2.54-2.50)
Number of observations	95404	85171	184896	90517
I/sigma *	10.6 (2.3)	8.6 (1.9)	8.3 (2.0)	10.16 (2.23)
CC1/2 (%) *	83.3 (55.9)	83.9 (56.2)	86.7 (59.5)	84.26 (59.22)
Completeness (%) *	94.4 (84.7)	91.9 (76.1)	99.0 (95.8)	92.9 (78.7)
Multiplicity *	3.9 (2.4)	3.5 (1.9)	7.1 (3.7)	3.7 (2.1)

* Numbers in parentheses represent the number in the highest resolution bin

Data processing (cctbxfel)

The still diffraction images were analyzed using the software packages IOTA (Lyubimov *et al.* 2016), *cctbx.xfel* (Sauter *et al.* 2013, Hattne *et al.* 2014) and *PRIME* (Uervirojnangkoorn *et al.* 2015). For each crystal, two images were collected, separated by a small rotation. We merged these data three times, keeping the first and second shots independent from another in two datasets, and merging all the data together in a third. Summary statistics are shown in Table 1.

Anomalous signal analysis

Anomalous data collected at XPP, LCLS at 7.2KeV was used to calculate the Fourier anomalous peak. The first shot only, the second shot only and the combined first and second shot (called both shots) were used separately to calculate anomalous peak table 2. It was not possible to solve the structure from the data sets collected at XPP as it was not always possible to find a distinctive anomalous peak at the iron positions from each of the heme groups present in the protein (The protein is a multiheme OCC containing 8 heme). This could be because of the low multiplicity of the data, especially for an anomalous data.

Molecular replacement

Though the anomalous peak that was obtained from the SFX data were weak, we were interested in determining the structure. Therefore, we carried out molecular replacement for datasets referred as first shot, second shot using MOLREP within the CCP4i (Potterton L*, et al.* 2004) suite using wild type structure with PDB ID 4Q05 as search model. The refinement statistics are shown in table 3 below.

Table 4 Structure solution and refinement

	bothanom	firstanom
Wavelength (Å)	1.7	1.7
Resolution range (Å)	67.59 - 2.5 (2.59 - 2.5)	56.42 - 2.5 (2.589 - 2.5)
Space group	R 3 2:H	R 3 2:H
Unit cell	135.175 135.175 212.678 90 90 120	135.13 135.13 212.651 90 90 120
Unique reflections	25837 (2471)	24690 (2220)
Completeness (%)	98.92 (96.56)	94.60 (86.91)
Wilson B-factor	9.63	8.62
Reflections used in refinement	25833 (2469)	24684 (2217)
Reflections used for R-free	1999 (191)	1999 (180)
R-work	0.3035 (0.3726)	0.3268 (0.3812)
R-free	0.3695 (0.4327)	0.3901 (0.4370)
Number of non-hydrogen atoms	4500	4480
macromolecules	4054	4054
ligands	359	359
solvent	87	67
Protein residues	531	531
RMS (bonds)	0.010	0.010
RMS (angles)	1.57	1.53
Ramachandran favored (%)	93.83	91.91
Ramachandran allowed (%)	6.17	7.51
Ramachandran outliers (%)	0.00	0.58
Rotamer outliers (%)	0.00	0.00
Clashscore	13.53	13.78
Average B-factor	15.12	16.55
macromolecules	15.52	16.99
ligands	10.84	12.28
solvent	14.59	13.24

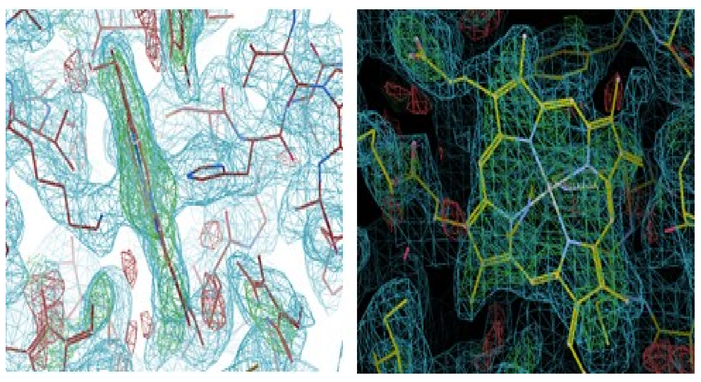

Figure 2 Electron density maps calculated from still images. (*A*) $2Fo - Fc$ (blue, contoured at 1.0 σ) for H1 at the site of one of the heme, following molecular replacement using data processed by cctbx.xfel. $2Fo - Fc$ density (blue, contoured at 1.0σ) for H1, showing the density corresponding to the heme group, which was omitted from the refinement.

Table 2 Anomalous peak

Fe site	Fourier anom peak (times sigma)			
	first shot data	Second shot data	"odd mixed" data	all data
601		2.2		
602	2.9	2.5	2.5	2.1
603	2.2	2.8		3.0
604			2.5	
605		3.7		3.3
606		3.6	2.6	3.5
607	3.0	2.0		3.0
608	2.5	3.4	3.0	3.1

Table 3 Correlation of anomalous signal calculated from the refined model and the processed data

	CC correlation coefficient of the measured anomalous differences to the differences calculated from the model			
Resolution	First short only	Second shot only	Both (all)	Both (odd):
2.5A	0.08	0.09	0.11	0.09
3.0A	0.09	0.10	0.13	0.09
4.0A	0.10	0.12	0.15	0.11

Conclusion

As we had very weak anomalous signal, It was not always possible to find a distinctive anomalous peak at the iron positions from each of the heme groups present in the protein (The protein is a multiheme OCC containing 8 heme) because of the low multiplicity of the data, especially for an anomalous data. We could on the other hand determine the structure by molecular replacement and find all the eight hemes in the protein. Electron density was clear enough to view all the eight hemes present in the protein. We are in the process of obtaining more redundant data and use other target proteins as well for the proof of principle.

Acknowledgements

We thank all the SSRL and LCLS staffs for their support and specially Henrick Lemke for beam tuning at XPP and Aina Cohen for the software support.

References

Cohen AE *et al.* (2014) Proc Natl Acad Sci U S A Dec 2;111(48):17122-7
Hattne *et al.* 2014 *Nature Methods*, **11**, 734–736
Uervirojnangkoorn *et al.* 2015 *eLife* **4**
Lyubimov *et al.* 2016 *J Appl Crystallogr* **49**: 1057-1064
Sauter *et al.* 2013 *Acta Crystallogr. D* **69**: 1274-1282