	Experiment title: DNA conformation within the complexes DNA-DPS grown in vitro	**Experiment number**: MX/2057
Beamline: ID23-1 ID23-2	**Date of experiment**: from: 16 June 2018 to: 17 June 2018	**Date of report**:
Shifts: 3	**Local contact(s)**: Max Harunobu Nanao	*Received at ESRF:*

Names and affiliations of applicants (* indicates experimentalists):

LOIKO Nataliya*, KOVALENKO Vladislav*, KRUPYANSKIY Yury*

Semenov Institute of Chemical Physics Laboratory of biopolymer dynamics Structure of matter RAS

Report:

Here we present the successful experience of Dps protein structure determination by multi-crystal data collection. This method using synchrotron radiation was successfully applied to determine three-dimensional structure of a Dps protein from *E. Coli* bacteria at 2.0 Å resolution. Final data set was obtained by combining 261 partial diffraction data sets measured from triclinic space group crystals with the sizes about 5 μm.

A structural and functional unique oligomeric nucleoid-associated protein from Escherichia coli bacteria, called DPS (Almiron *et al.,* 1992), was discovered in 1992. This protein protects cell from oxidative stress, inactivating potentially dangerous ions Fe^{+2}, and provide innercellular process of nucleoid crystallization (Almiron *et al.,* 1992). Although Dps structure does not contain classic modules for DNA nucleotide sequences recognition, it's considered, that interaction between DNA and Dps implemented by electrostatic bonds formation between negative charged DNA surface and positive charged amino acid residues, localized at N-terminus site of Dps monomers (Chiancone *et al.,* 2010). It still stays unclear, how bacterial chromosome ordered compaction control by Dps protein is performed.

Our initial task was to solve the structure of Dps-DNA crystal complex, grown *in vitro*, to find optimal crystallization conditions we varied DNA and protein mass concentrations ratio (from 6:1 to 1:6), use different types of precipitants (low masses PEGs, ammonia sulfate, glycerol, MPD) and varied length of DNA (24 b.p., 1000 b.p., circular vector pBluescript SK+/BamHI, QIAGEN Plasmid Maxi Kit, 2958 b.p and 10000 b.p.). But unfortunately, all of grown crystals contained only protein in its structure. Although our initial task was failed, there were one crystallization conditions that gave us a result worthy of scientific interest. Crystals obtained by hanging drop vapor diffusion method, mixing Dps protein solution (3.1 mg/mL) and pBluescript SK solution

(1.04 mg/mL) in 6:1 mass ratio and after that diluted with ammonium sulfate as a precipitant in a volume ratio of 1:1, were growing during six months and occupied all drop volume forming dense layer of small (3-7 microns in linear dimension) crystals (Fig. 1).

Preliminary quality estimates of Dps crystals were made at an ID23-1 beamline equipped with the detector PILATUS6-M, the radiation wavelength was 0.972 Å. The minimum beam size at this beamline is achieved with an aperture of 10 μm and was too far from optimal for Dps crystals. Main measurements were made at micro-focus beamline ID23-2 equipped with micro-diffractometer MD3Up and detector PILATUS3 X 2M. The beam size on the sample collimated with X-ray mirrors is 10 x 4 μm², the radiation wavelength is 0.873 Å.

All experiments were carried out at 100 K. The crystal solution from the crystallization drop was transferred to a sample holder and quickly frozen in liquid nitrogen. The *DOZOR* and *MESH-BEST* (Melnikov *et al.,* 2018) software programs were used to automatically recognition of the diffraction from protein crystals and form a final map of crystal position distribution in the holder. The collection parameters were determined based on estimates of the resolution limit equal to 2.0 Å produced by the software program *DOZOR*. The full angle of rotation per crystal varied from 7 ^{0} up to 20 ^{0} for different loops with crystals; the angle increment of a diffraction image was the same in all cases equal to 0.1 ^{0}. Exposure time was determined so that absorbed radiation dose per crystal was around 10 MGy. The dose rate was estimated by *RADDOSE* (Karthik *et al.,* 2010). During the measurements, the exposure time and the full angle of rotation varied with an attempt to find the most optimal conditions. The correct selection of the exposure was controlled based on the results of rapid processing of successive diffraction patterns from individual crystals using *DOZOR*.

Processing for such partial datasets requires attention since the reflections can be weak and that indexing with a small wedge can be more complicated. This is particularly true if the crystals happen to be in a low symmetry and unknown space group. When rotation data is collected at the ESRF automatic processing runs. For multi crystal data collection, these results can be used as an initial guess to process all the partial datasets which have been collected. Only for 71 out of 404 crystals of Dps automatic processing produced some results. In total six possible space groups were determined, three of them had two types of lattice parameters. Further, data sets were processed by software *XDS* (Kabsch, 2010) using all the nine possible variants found in the preliminary analysis to avoid any decision bias. The most successful and plausible space groups resulted to be triclinic group P1 (a= 87.9, b= 89.5, c= 156.5, α=105.7 β=92.3, γ=117.1). Processing all the 404 individual datasets in this space group with XDS was successful for close to 90% of all collected datasets (374 partial datasets). The success rate for different loops does not show the dependence on the width of the collected wedge on the partial dataset, but rather on the individual quality of each small crystal.
Low symmetry datasets are a challenging test case for merging procedures. As shown above, the partial datasets have low completeness and most of the reflections are extremely dim, especially at higher resolutions. Adding this issue to the low number of common reflections makes it difficult to assess isomorphism between the partial datasets. We compared thus intensities using hierarchical cluster analysis, where the comparison between the partial datasets is based on their pairwise correlation coefficient, as discussed in (Giordano *et al.,* 2012). To summarize, distance between the datasets is defined as $d(a,b) = \sqrt{(1 - cc(a,b)^2)}$. We used the software ccCluster (Santoni *et al.,* 2017) to perform the distance calculations and the consequent HCA and merging of the data. We found a cluster at threshold 0.65 containing 255 of datasets. This threshold value means that the average correlation between those 255 datasets is 0.75. Merging all of those together in XSCALE provides a 98% complete dataset with quite good statistics. Dendrogram for thus experiment is shown in Fig. 2. Some datasets have low to none superimposition with the partial datasets, thus being impossible to compare by HCA methods. Once a convincing cluster solution was found, we manually added some of those back into it. We managed this way, after a tedious manual analysis, to find a group of 6 more datasets that improve the overall statistics of the final dataset, even if by a marginal factor.
The structure of Dps was successfully determined and refined based on merged multi-crystal data set. Phase problem was solved by molecular replacement method, using the dodecameric structure of DPS, pdb access code 1JTS as initial model. The structure was refined using the REFMAC5 (Murshudov *et al.,* 2011). The COOT program (Emsley *et al.,* 2010) was used for visual inspection, manual model refinement, and geometry verification. Final refinement statistics are reported in Table 4. All of the solvent molecules in the structure were modeled as water, with the exception of 46 SO_4 ions bounded on the inner surface of the dodecamers. The coordinates have been deposited in the RCSB Protein Data Bank with the accession number 6QVX. At Fig. 3 depicted electron density map on outer surface of Dps dodecamer.

Despite of our initial task for receiving of DNA-Dps co-crystalls was failed we did manage to solve Dps structure from small sized crystalls with the lowest symmetry space group, wide sread of unit cell constants of partial datasets and prefereble orientation of crystals lattiece in sample holder. The successful solution of Dps structure confirms the applicability of the Mesh&Collect pipeline for crystals of micron dimensions with the lowest possible symmetry of crystal lattice. Automatic search and selection of the crystals with measured diffraction is performed with high efficiency: more than 90% of sets were processed and about 65% of sets were used in final combination of data.

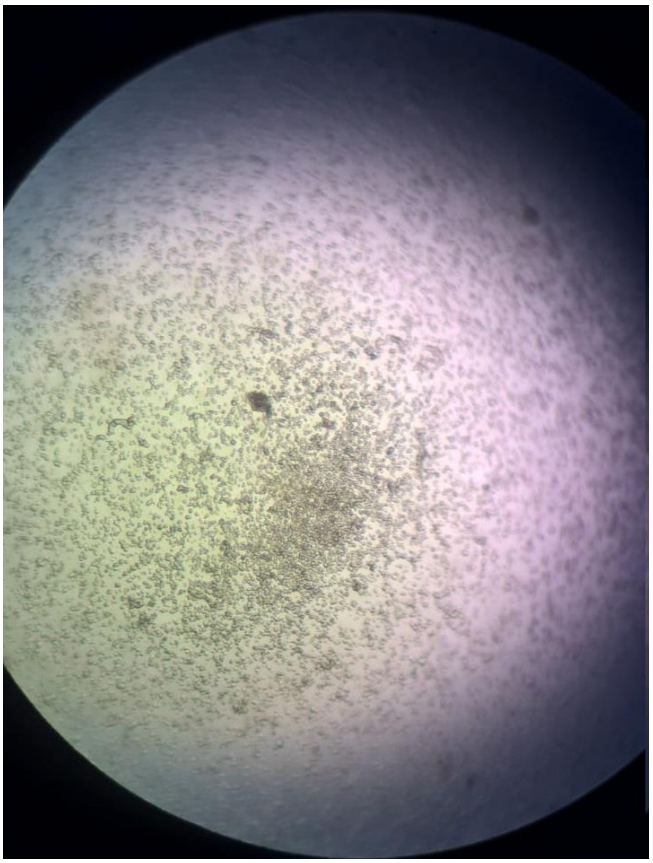

Figure 1 View of the drop with the Dps crystals through an optical microscope.

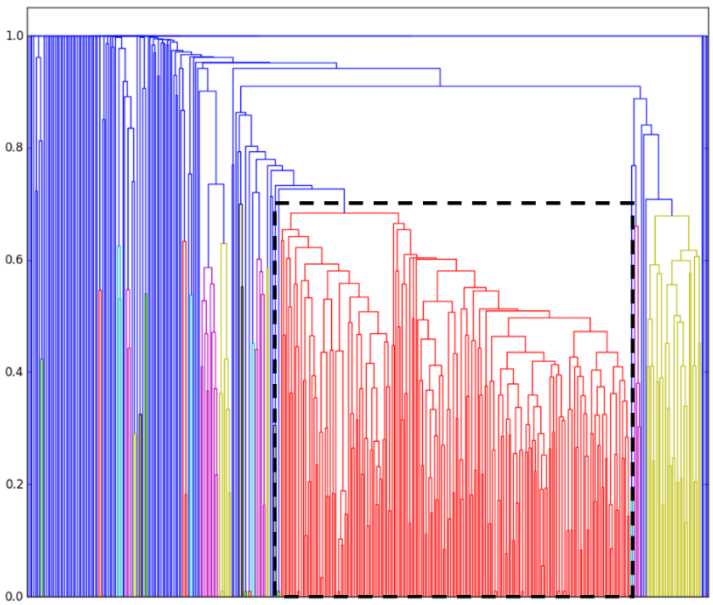

Figure 2 The dendrogram based on 374 single-crystal sets, it's colored according to chosen clustering threshold of 0.65. The red cluster, used to solve the structure is contoured by the dashed rectangle. It contains 255 partial datasets, collected with different oscillation ranges.

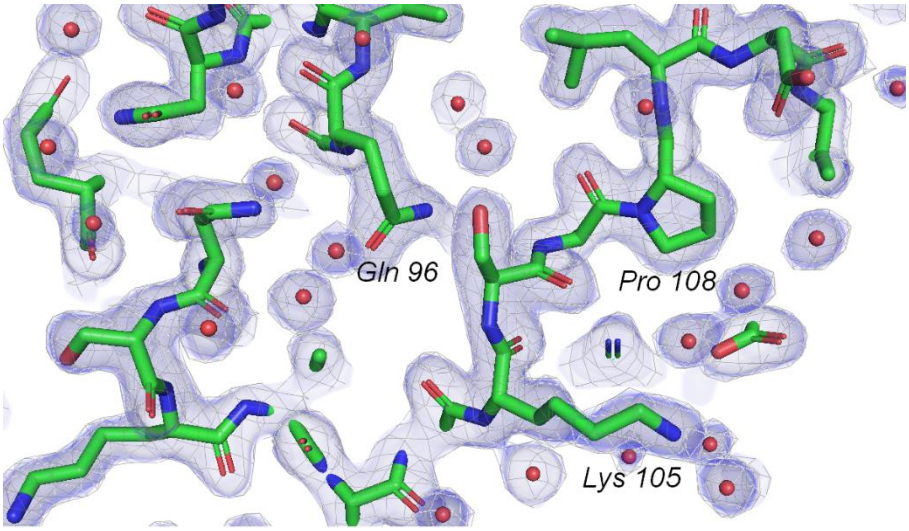

Figure 3 Electron density map on outer surface of Dps dodecamer.

Table 1 Data and refinement statistics for final structure of Dps crystal; Data in parentheses are for the highest resolution shell.

Space group	P1
Unit cell (a, b, c Å; α, β, γ °)	89. 90.6 157.1; 92.2 105.0 117.9
Resolution range (Å)	50.0 – 2.0 (2.2-2.0)
Total No. of reflections	5664896
Number of unique reflections	272262
No. of reflections, test set	259017 (19152)
No. of reflections, test set	13287 (925)
Completeness (%)	98.7 %
Multiplicity	20.8 (19.1)
Half-set correlation $CC_{1/2}$	93.8(39.0)
I/σ(I)	4.9 (1.7)
B-factor, Wilson plot ($Å^2$)	14.5
Final R-factor	0.210
Final R-free	0.261
Number of atoms	33105 (3243 water molecules)
<B> from atomic model ($Å^2$)	19.0

R.m.s.d bond lengths (Å)	0.0093
R.m.s.d angles (°)	1.58
Ramachandran plot analysis Most favoured regions	98.5%
Allowed regions	1.5%

Research was performed within frameworks of the state tasks for ICP RAS 0082-2014-0001 (state registration #AAAA-A17-117040610310-6) and 0104-2018-0029.

References

- Almirón M., Link A., Furlong D., Kolter R. (1992). *Genes. Dev.* Vol. **6**. – pp. 2646–2654
- Chiancone E., Ceci P. (2010). *Frontiers in Bioscience* **15** 122-131
- Emsley P., Lohkamp B., Scott W.G., Cowtan W.G. (2010). *Acta Cryst*. D**66**: 486–501.
- Giordano R., Leal R. M. F., Bourenkov G. P., McSweeney S. & Popov A. N. (2012). *Acta Cryst.* D**68**, 649–658
- Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
- Melnikov I., Svensson O., Bourenkov G., Leonard G. and Popov A. (2018). *Acta Cryst.* D**74**, 355-365
- Murshudov G.N., Pavol Skubak, Andrey A. Lebedev, Navraj S. Pannu, Roberto A Steiner, Robert A. Steiner, Robert A. Nicholls, Martyn D. Winn, Fei Long, and Alexei A. Vagin, (2011). *Acta Cryst.* D**67**, 355-367
- Santoni G., Zander U., Muleller-Dieckmann Ch., Leonard G. and Popov A. (2017) *J Appl Crystallogr* **50** (Pt6) 1844-1851 Doi: 10.1107/S1600576717015229