	Experiment title: CHIKV nsP1 in complex with S-adenosyl methionine, a substrate for the capping reaction	**Experiment number**: MX2261
Beamline: CM01	**Date of experiment**: from: 25/11/2020 to: 27/11/2020	**Date of report**: 18/02/2021
Shifts:	**Local contact(s)**: Michael Hons	*Received at ESRF:*

Names and affiliations of applicants (* indicates experimentalists):

Rhian Jones AFMB Universite Aix-Marseille

Juan Reguera AFMB Universite Aix-Marseille

Report:

The aim of this experiment was to solve the structure of CHIKV nsP1 in complex with an S-adenosyl methionine (SAM) ligand; one of the precursors of the capping reaction. The experiment was performed by Michael Hons, who communicated with us via Skype about the experimental set-up due to coronavirus restrictions. Four replicate grids of nsP1 with an S-adenosyl methionine (SAM) substrate (Quantifoil R2/2 holey 300 mesh) that had been pre-screened on the TALOS Artica at the CNB Madrid were loaded into the microscope by Michael. Unfortunately, the first grid used for a previous data collection on a TALOS had too low a number of suitable squares for an overnight data collection, as many areas of the grid seemed to be dry. We assume that this is something that happened in transportation of the grids as significantly more contamination was also apparent in the ATLAS screen relative to that aquired at the CNB. Due to this, Michael had to spend a significant amount of time in screening the other three grids to find suitable areas of ice for data collection, and we are extremely grateful for his patience and for salvaging the experiment.
The evening of the 25th an automated data collection was set up with the following parameters; 4 images were collected per hole at a magnification of 165,000 (corresponding to an image pixel size of 0.827Å/pix), spot

size 5, defocus range: 1-2.5µm, exposure time: 4s distributed over 40 frames, with a dose of 1.095 $e^-/Å^2$ per frame yielding a total dose per movie of 42.36$e^-/Å^2$.

A total of 5158 movies were collected, and on the fly processing was performed using the CM01 pipeline; monitored via the ExiMX interface. Movie drift was first corrected using MotionCorr2 and CTF correction performed with gCTF. The images exhibited minimal drift over the dataset, and high overall resolution. In addition, ice contamination was markedly improved in the grid Michael collected from relative to the one used for data collection on the TALOS, evident from comparison of the power spectra. Particles are clearly visible at up to -0.8um defocus.

Data were transferred to the lab, and the movies were processed from within Relion 3.0. Following motion correction (Motion Corr2) and CTF correction (CTFFind 4), 599 of the 5158 movies were discarded due to drift, ice contamination, poor contrast or highly astigmatic CTFs. Around 1000 particles in a selection of orientations were selected from 100 micrographs for 2D classification, yielding templates for Relion's autopicking procedure. 366,156 particles were selected from the dose-weighted micrographs in auto-picking (roughly 80 per micrograph), and 214,084 were retained following 2D classification. 2D classification yielded classes clearly corresponding to single dodocameric rings of nsP1, and double dodecameric rings (24mers, Figure 1) where the two rings were separated by a belt of amphipol. The presence of the double and single ring species of nsP1 has previously been documented in negative stain and cryo-EM imaging. The proportion of single rings relaive to double rings is much lower in this dataset relative to the dataset collected for apo nsP1 at the ESRF, consistent with what was observed from processing of the TALOS data collection with this batch of grids. Both single and double rings exhibited a range of particle orientations corresponding well with projections from our final models.

Figure 1: Class averages calculated in Relion 3.1, showing views corresponding to single dodecamric rings and double rings.

3D classification was used with low pass filtered single or double ring apo nsP1 structures obtained previously at the ESRF to check for the presence of multiple conformations/ bound states, with particles that were binned 2-3 times from an original box size of 310 pixels. Classifications with different numbers of classes consistently yielded a single major class corresponding to the single rings (particles) and double rings (69,000 particles), with smaller classes of "junk" particles, suggesting that there is only one major conformation for the rings. All classifications were performed without imposing symmetry. A double class containing 69,000 particles was rextracted at the original pixel size and autorefined to 2.88Å (2.7Å post processing using a loose mask). Reconstruction following a single round of particle polishing and CTF refinement on a per-particle basis did not significantly imporve resolution. However, density for the methyltransferase domains that form the crown of nsP1 is more poorly defined relative to the core of the structure (Figure 3a), suggesting that this domain is flexible or that there may be subtle differences in conformation between individual monomers that

are averaged out through imposing symmetry. In the active site pockets within each methyltransferase domain for every monomer in the ring, we observe an additonal peak in the maps that is likely to correspond to the adenosine base of the SAM (Figure 3b). However, the peak is quite poorly defined, particularly in the region beyond the adenosine base, and many of the side chains in the surrounding region are relatively poorly defined in the maps. Masked refinement in C1 yields a reconstruction at 3.75Å (3.4Å post processing), where some minor improvements in the overall density are observed, but the density within the active site remains ambiguous. We assume this is due to averaging of ligand occupanices between monomers, overall lower local resolution in this region, or partial occupancy of the SAM site. In order to distinguish between these possibilities, various data processing strategies are ongoing. Symmetry expansion will be performed within Relion with focussed classification or refinement to try and assess differences between individual monomers and improve the ligand density.

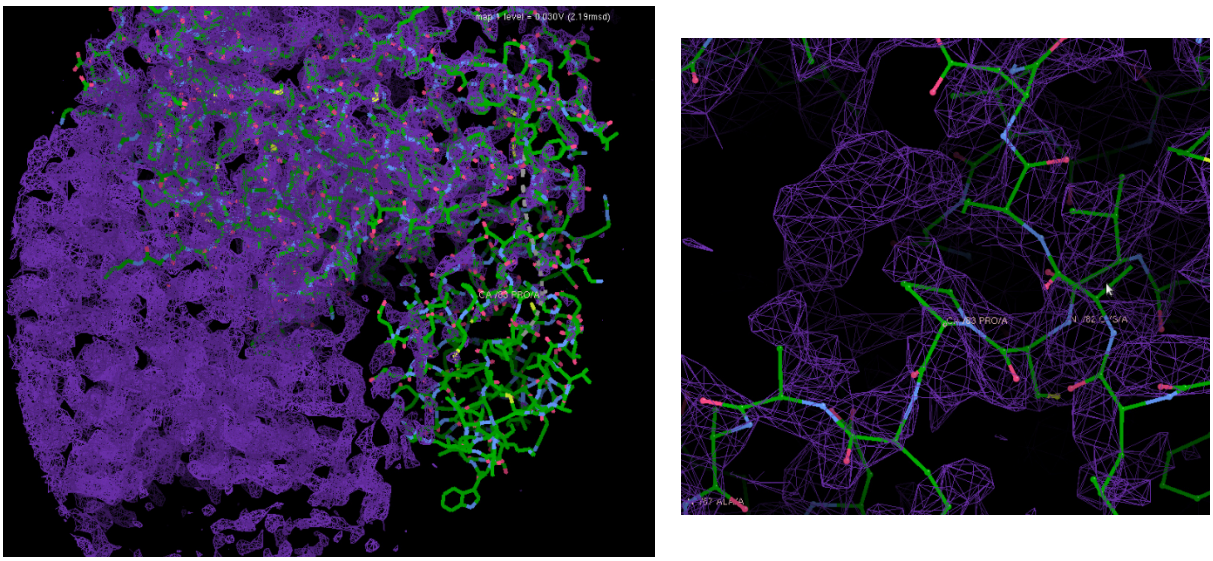

Figure 3: Map features viewed in Coot. A: Representation of the D12 symmetrised map at a contour level of 2.19 RMSD (0.03V) demonstrates that the local resolution of the methyl transferase domains in the crown of the rings is significantly lower than in the core of the rings. B: Detail of the methyl-transferase active site showing extra potential that is likely to correspond to the SAM ligand.

Michael has also been assisting us with the EM processing, and has obtained maps with notably improved occupancy and resolution through using non-uniform refinement in cryosparc and local B factor sharpening with deepemhancer (nominal overal resolution of 2.22 Å). Michael's guidance in data collection and processing has been indispensable, and he is now collaborating with us on this project as an EM expert.