ESRF	**Experiment title:** CryoEM structure Determination of the ZIKV NS5 in complex with RNA promoter element 5SLA	**Experiment number**: MX-2235
Beamline:	**Date of experiment**: from: 29/11/2019 to: 02/12/2019	**Date of report**: 30/07/19
Shifts:	**Local contact(s)**: Michael Hons	*Received at ESRF:*

Names and affiliations of applicants (* indicates experimentalists):

Dr Nuria Verdaguer (main proposer). Dr Diego Ferrero *. Mr. Victor Ruiz Arroyo*.

Molecular Biology Institute of Barcelona. IBMB-CSIC. Parc Cientific de Barcelona. Baldiri Reixac 15. 08028 BARCELONA. SPAIN

Report:

- **Overview:**

Zika virus (ZIKV) is a Flavivirus with single stranded RNA (ssRNA) genome that encodes for 3 structural and 5 non-structural proteins (NS). The largest protein is NS5 (100kDa) which is the responsible of copying the ssRNA viral genome. Flanking the RNA genome ends, there are two highly structured untranslated regions, the most relevant RNA structure for viral replication is located at 5' and forms a small stem loop (5SLA) that strongly binds NS5 and has been reported to act as replication promotor. We optimized the *in vitro* formation of the ZIKV NS5-5SLA derived complex and the vitrification conditions to obtain spreaded complex particles and proper ice thickness. Several Cryo-EM grids (Quantifoil 1.2 1.3 Cu/Rh) were sent to ESRF and movies were collected at CM-01 using the FEI Titan Krios electron microscope equipped with K2 direct electron detector set in counting mode at a magnification of x165000 (pixel size of 0.827 Å/pixel). We collected 3642 movies of 80 frames each during 8 seconds with a dose per frame of 0.75 e/A^2/frame.

- **Quality of data:**

The best cryo-EM grids sent to ESRF were seriously damaged during the shipment, nevertheless some of them were still intact for data acquisition. Movies collected were motion corrected using MotionCorr in the on-the-fly pipeline installed at CM-01 beamline. Corrected micrographs present a high heterogenicity in the particle distribution, showing regions completely crowded of particles and regions with good particle concentration.

The quality of the images is considerably high, reaching the 2.14 Å resolution in the CTF estimation (Figure 1).

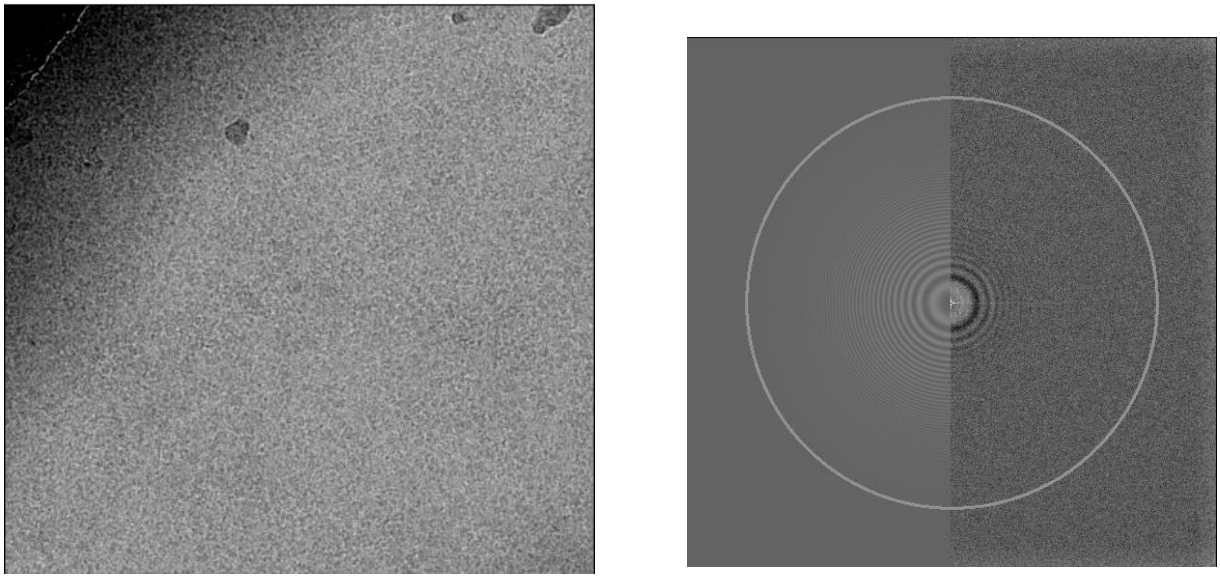

Figure1: In the left panel, a representative micrograph of NS5-5SLA complex and its CTF estimation at the right panel.

- **Status and progress of evaluation**

Despite of the particle heterogenicity and the presence of crowded areas along the micrographs, corrected micrographs were picked using a previous subset of NS5-5SLA 2D averages. We used several autopicking programs as Relion, Gautomatch, Cryosparc and Cryolo, obtaining the best results using a set of manually picked 2D class averages as input for Gautomatch and seting a picking threshold of 0.4 with a particle diameter of 150 A (Figure 2).

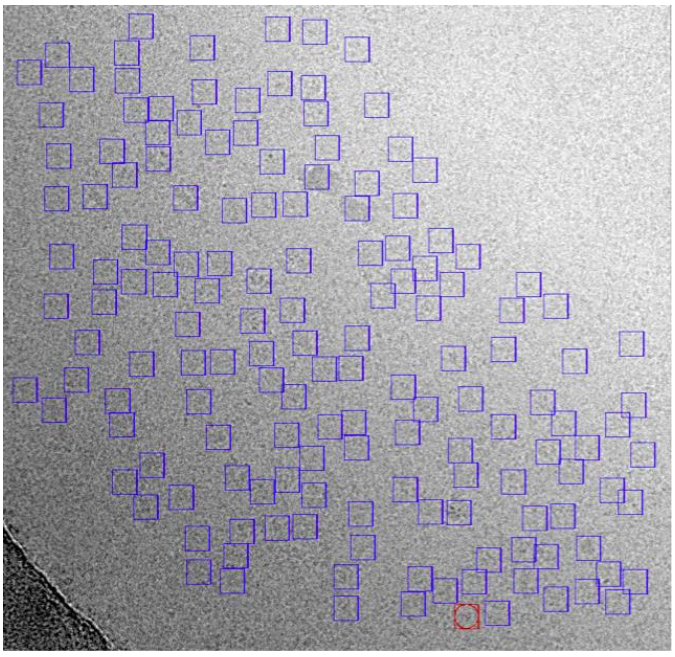

Figure 2: A representative micrograph picked with Gautomatch program.

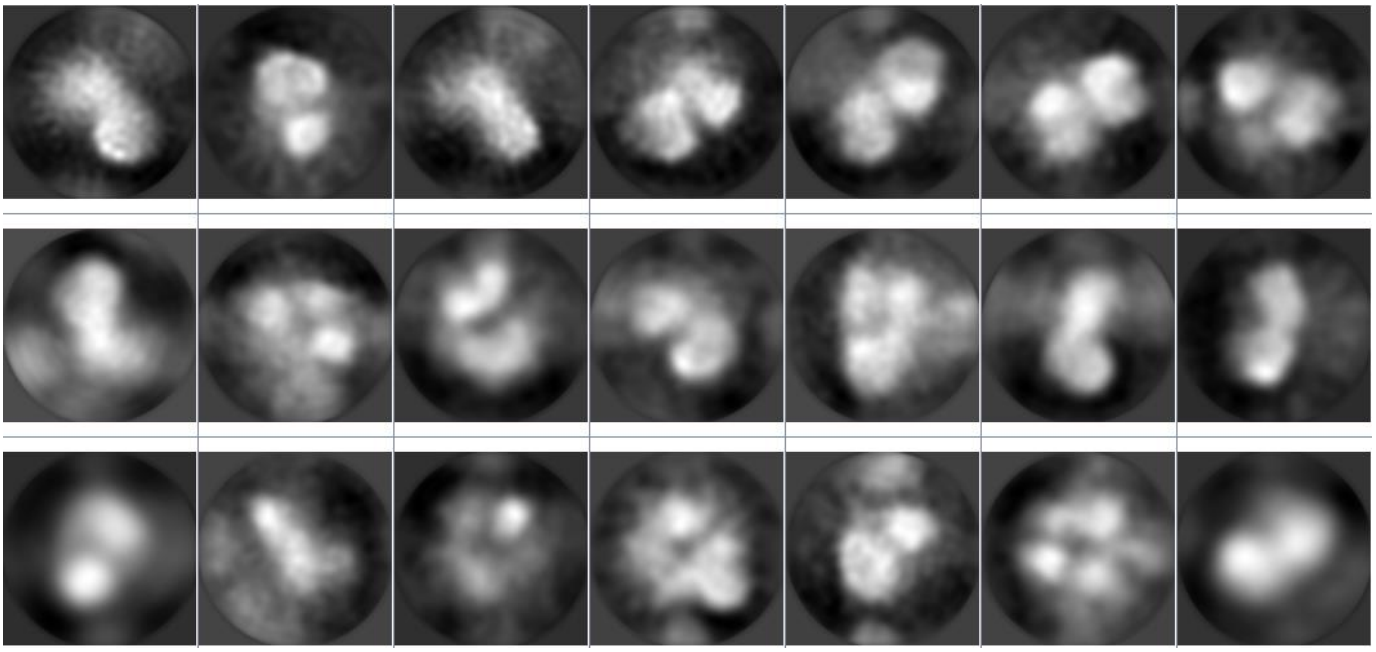

Figure 3: 2D class averages obtained after automatic picking of NS5-5SLA corrected micrographs

A total of 435687 particles were automatically picked with Gautomatch using a particle diameter of 150 px and an extracted box of 220x220 px. Using Relion we performed a reference free 2D classification (Figure 3) to get rid of the uncomplexed NS5 protein and picking artifacts, finally, a subset of 16201 particles were selected to generate a 3D ab-initio model using Relion.
Several 3D classification and 3D refinements were performed with Relion to reach the current model that fits a NS5 protein and exhibit an extra density that could correspond to the RNA (Figure 4).

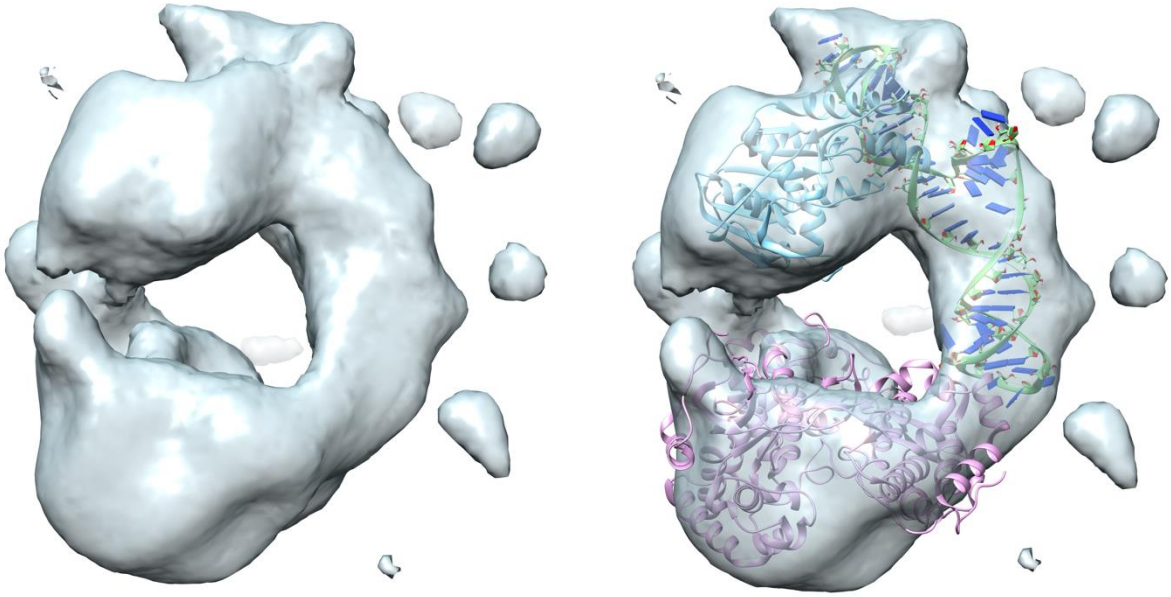

Figure 4: At left, current cryo-EM map of NS5 in complex with 5SLA RNA, at right, the crystallographic ZIKV NS5 model and a predicted 5SLA model fited in the same cryo-EM map.

As we aim to reach higher resolution, we kept working in the picking process trying other picking programs as Relion autopicking and Cryosparc. We found that Relion does not produce reliable results with either of the two picking protocols, neither with LoG Laplacian nor with 2D average template assisted picking. Both picking yield artifactual 2D class averages when submitted to 2D classification. Cryosparc was not either able to pick our complex¡ particles and just picked NS5 particles which yielded 2D class averages that corresponded to the uncomplexed protein when particles were subjected to reference free 2D classification.

Currently we are optimizing picking parameters to run crYOLO autopicker. First pickings using a small dataset of micrographs with this program produced a promising set of particles that after being subjected to a free reference 2D classification in Relion, yielded a good quality 2D class averages.

- **Results**

Despite of the low resolution of our current maps we can confirm that ZIKV NS5 protein binds the RNA promoter 5SLA, linking both the MTase and the RdRP domains. However, to improve the resolution of the final reconstruction we are taking a step back to optimize the picking process. Our future experiments will revolve around the picking program crYOLO aiming to obtain better particle selection. In parallel, we will further fine tuning optimize vitrification conditions specially as changing glow discharge times or testing other grid materials, in order to improve the quality of the experimental data.