ESRF BAG report MX2263

Title of BAG	The Iberian cryo-EM Community		
Main Proposer	Rafael Fernández-Leiro, Spanish National Cancer Research Centre - CNIO		
BAG Members (PIs/Affiliation)	Ana Luisa Carvalho, FCT; Alberto Marina Moreno, IBV; Margarida Archer, ITQB; Beatriz Herguedas Francés, BIFI; Carmen San Martín, CNB; Clara Marco, IBV; Célia Romão, ITQB; Cristina Vega, CIB; Daniel Lietha, CIB; Daniel Luque, ISCII; Ernesto Arias palomo, CIB; Elin Moe, ITQB; Ignasi Fita, IBMB; Israel Sánchez, IBMB; Iban Ubarretxena, EHU; Jose Maria Casasnovas, CNB; Javier Garcia-nafria, BIFI; Jose Luís Llacer Guerri, IBV; Jaime Martin-Benito Romero, CNB; Jose María Valpuesta, CNB; Jose Castón, CNB; Manuel Palacín, IRB; Pedro Matias, ITQB; Miquel Coll, IRB; María Joao Romao, FCT; Marcelo Guerin, BIOGUNE; Nuria Verdaguer, IBMB; Oscar Llorca, CNIO; Patricia Casino Ferrando, UV; Pedro Pereira, IBMC; Rafael Fernandez-Leiro, CNIO; Santiago Ramon-Maiques, IBV; Sean Connell, BIOGUNE; Sandra Ribeiro, IBMC; Teresa Santos-Silva, FCT; Armando Albert, IQFR; Juan Hermoso, IQFR; Julia Sanz Aparicio, IQFR; María José Sanchez Barrena, IQFR; Jerónimo Bravo, IBV.		
Proposal Reference Number	MX2263	**Date of Report**	February 25th 2022
Sessions Reported	6		
Number of EMDB/PDB submissions since last report	0		
Number of publications since last report	0		

Use of beam time

This report covers the period since last report on February 2021 until now. This includes six sessions from MX2263 not reported in the previous document and 5 sessions from MX2269. The last three sessions scheduled for MX2269 (February/March 2022) will be reported in the next period as they are too recent or in the future.

Session	Beam-line	Date	User
MX2263-5	CM01	27 - 29 January 2021	Célia Romão
MX2263-6	CM01	05 - 07 March 2021	Ernesto Arias Palomo
MX2263-7	CM01	23 - 25 April 2021	Clara Marco Marín
MX2263-8	CM01	14 - 16 May 2021	Oscar Llorca
MX2263-9	CM01	21 - 23 June 2021	Dani Luque
MX2263-10	CM01	30 - 02 July 2021	Jose María Valpuesta

Beam-time reports

MX2263-5 - *Deinococcus radiodurans* DNA polymerase I/Single particle cryo-EM (Dr Célia Romão)

Session	Date	Micrographs	Ptcls	Fractions	Å/px	Images/hole
MX2263-5	27-29/01/2021	6473		60	0.827	2

D. radiodurans DNA polymerase I (DrPolI) is a monomeric 102 kDa protein. Despite the limitation for single particle cryo-EM has been the low signal-to-noise ratio due to the protein size. Our preliminary sample preparation analysis revealed good quality 2D class-averages (Fig. 1a), before data collection on the Titan Krios beam CM01.

From the 4 grids sent to ESRF, we selected for data collection the one with best ice quality that was prepared with apo-protein. In the initial collected images analysis on session MX2263-5, possible particles were detected. But after pre-processing data, 2D classes did not show DrPolI particles (Fig. 1b). Different picking methods were used from Scipion and cryoSPARC. Based on the 2D classification results, we consider that the grid was "empty" or almost "empty". Unfortunately, a problem in the EM sample preparation was not detected in the early steps of data collection due to the difficulty to detect this protein. The expected size of DrPolI monomer is 9-10 nm.

A B

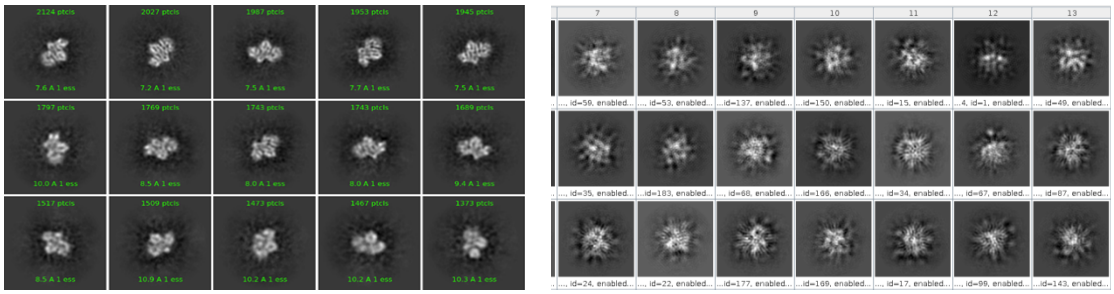

Fig. 1. MX2263-5. (A) Expected 2D classes (previous data collection). **(B)** Obtained 2D classes (data from MX2263-5 session).

MX2263-6 - *Regulation and target selection for DNA transposition* (Ernesto Arias-Palomo)

Session	Date	Micrographs	Ptcls	Fractions	Å/px	Images/hole
MX2263-6	5-7/03/2021	3,078	137,787	46	1.052	1

DNA transposons are nucleic acid fragments that can transfer horizontally between different DNA molecules. This activity can modify gene expression, promote organismal evolution, and spread antibiotic resistance and virulence factors. How transposable elements are regulated has been a long-standing question, however. Notably, transposase control often relies on nucleotide cofactors to regulate transposase activity or to choose appropriate target DNAs. Here we are studying IS21, one of the most widespread families of transposons, which is composed of a canonical DDE transposase, IstA, and an essential ATPase regulator, IstB. We have previously determined that IstB forms decamers that sharply bend short fragments of DNA. However, we have recently observed that IstB can form higher-order filaments/oligomers

on longer DNA substrates, reminiscent of those formed by the Mu and Tn7 regulators MuB and TnsC.

In this experiment we have analysed the structure of IstB bound to a large duplex DNA (120 bp). This complex form elongated filament-like structures that offer a limited range of views, and we collected some tilted images (1963 images out of 3078 were collected at 35°) to alleviate this issue. We picked the particles using Topaz, and after several rounds of 2D and 3D classification we have been able to obtain a 4.4 Å resolution structure (Fig. 2). We are still processing the data but in the current state we are already able to model the protein and the DNA using our previous results. Notably, this structure provides key insights that help to understand the molecular determinants of the nucleic acid in this process and how IstB primes the target DNA to promote the strand transfer reaction.

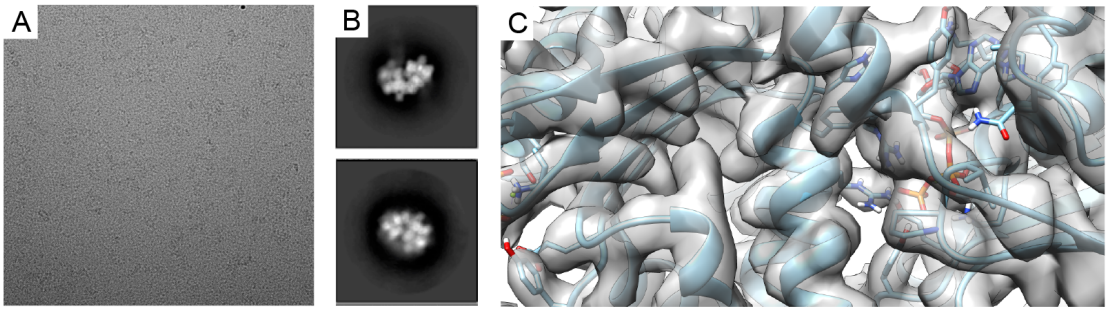

Fig. 2. MX2263-6. (A) Micrograph. **(B)** Reference-free 2D classes of the complex. **(C)** Detail of the reconstruction and model.

MX2263-7 - *Structural insight into human Δ¹-pyrrolin-5-carboxylate synthetase (P5CS) deficiency (P5CSD) by using single particle cryo electron microscopy of the bifunctional enzyme that causes this genetic disease.* (Vicente Rubio and Clara Marco-Marín)

Session	Date	Micrographs	Ptcls	Fractions	Å/px	Images/hole
MX2263-7	23-25/04/2021	4767	249.405	42	1.05	1

P5CSD is due to decrease or loss-of-function mutations in the bifunctional enzyme P5CS, an oligomer of a single polypeptide with two catalytic domains (one per reaction) that catalyse the first two steps of de novo synthesis of both ornithine and proline. Interestingly, we found that this genetic disorder presents a puzzling variety of four genetic syndromes defined by a combination of their clinical manifestations and their type of inheritance, dominant or recessive. The type of inheritance and the type of syndrome depend on the mutation(s) present. We have interpreted the disease as a single continuum of severity and the dominance of some mutations as dominant negative effects that would result from disturbed structural organization of the enzyme oligomer with impaired channelling of the intermediate and unstable product, glutamate-5-phosphate from the kinase to the reductase domain of adjacent subunits.

In order to understand such channelling, and also to clarify the mechanism of the negative dominance in the P5CS mutation-associated disease spectrum, we determined the structure for the apo form of human P5CS at 3.4 Å overall resolution which forms tetramers organized

as dimers of dimers, and that can associate into octamers (4.23 Å) and dodecamers (4.68 Å). To better delineate the P5CS active centers, we prepared samples of P5CS in the presence of substrates or substrate analogs, hoping to be able to visualize some loops with no density in the maps for the apo structure that are expected to be involved in the binding of substrates, that could facilitate the understanding of P5CS functioning particularly of the intermediate channelling.

From these samples we collected 4767 images from which we extracted 249.405 particles after 2D-classification. Analysis of these revealed that the protein had a preferential orientation, with most of the particles corresponding to top views where only one of the catalytic domains (reductase domain) can be seen. Interestingly, in our previous structure of the apo form of P5CS, this particular view was rare to be seen and in consequence the local resolution for this domain was not better than ~ 4,5-5 Å. So, in order to boost the resolution of this particular domain we decided to combine the particles from this dataset and from our previous dataset (Fig. 3) and we obtained a map with better resolution for the reductase component, allowing us to refine the position of some secondary structure elements of the reductase domain. We are currently finishing the analysis of the structures and in the process of manuscript preparation.

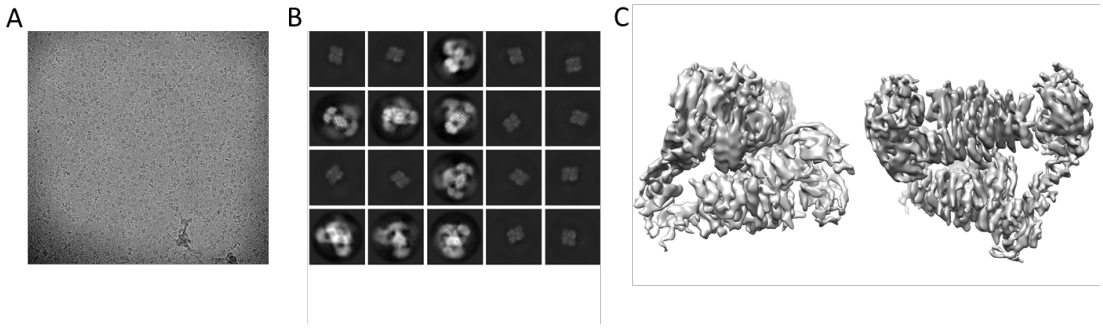

Fig. 3. MX2263-7. (A) Typical micrograph. **(B)** Representative 2-D clases. **(C)** Two views of the P5CS octamer map.

MX2263-8 - Cryo-EM characterization of the histone modifying complex (Oscar Llorca)

Session	Date	Micrographs	Ptcls	Fractions	Å/px	Images/hole
MX2363-8	14-16/05/2021	7283	631K	45	1.052	1

Chromatin remodelling regulates the dynamic balance between genome packaging and accessibility, through the concerted action of chromatin remodelers and histone-modifying enzymes. The histone modifications specifically recruit other proteins, dictating whether particular regions of chromatin are actively transcribed or are held in an inactive state. Thus, specific histone modifications are crucial for transcription initiation and elongation, as well as DNA repair. Moreover, aberrant histone modifications can impact cell cycle progression, apoptosis, tumorigenesis and stem cell differentiation. Despite the intensive research on the histone modification machinery, its structure and the molecular mechanisms driving its specificity are still unclear.

In this session we collected a dataset on a histone-modification complex. 7283 movies were collected, motion corrected, and automatically picked with Topaz software yielding a pool of 631K particles. To deal with the conformational and compositional heterogeneity of the dataset, specific image processing software was used, such as Relion, cryoSPARC and CryoDRGN. This approach yielded a 3D reconstruction of the NCP complex (Fig. 4) that shows an extra density bound to the protein nucleosome face and the DNA in a flexible manner. We are currently improving the quality of the cryo-EM model to dissect the interactions involved, which will be the subject of a manuscript in preparation.

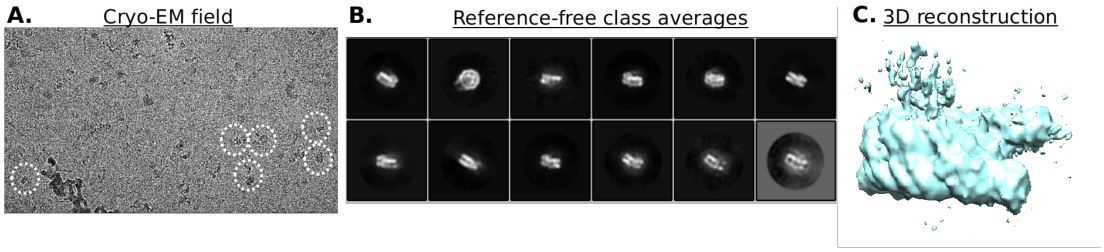

Fig. 4. MX2263-8. Cryo-EM characterization of the histone modifying complex. (A). EM field showing single particles in different orientations. **(B)**. Reference-free class averages. **(C)**. 3D model of the NCP complex at 7.3 Å resolution.

MX2263-9 - *Rotavirus group A vDLP polymerase VP1 loaded* (Daniel Luque)

Session	Date	Micrographs	Ptcls	Fractions	Å/px	Images/hole
MX2263-9	21-23/06/2021	8,961	58,268	50	0.675	2

One of the aims of our laboratory is the study of the structural variability in the Rotavirus (RV) genus. RV is the main causative agent of infectious gastroenteritis in infants and young children in developed and developing countries. RV are divided into 9 different groups named alphabetically from RVA to RVI. The RVA, B, C and H. Most of the information that is available about RV has been obtained from the study of the RVA, since the RV of other groups present problems of adaptation to in vitro growth, which in many cases has not yet been developed. The expression of the structural proteins of RV belonging to the B, C, D, F, G, H, and I groups, allow us to obtain VLP of these RV. The structure of these VLP will be similar to the structure of the particles obtained during the viral infection. The determination of their structure at atomic or near-atomic resolution will allow us to know for the first time the structure of the reference viruses of the different RV groups, analyse their structural variability, where should reside the differences of tropism and host of these viruses.

During the MX2263-9 session we collected, using the K3 direct detector in super resolution mode, ~9K good quality movies of control RVA vDLPs loaded with the VP1 RNA-dependent RNA-polymerase. After movie averaging and dose weighting ~58 K particles where selected and extracted. 2D and 3D classification was performed to select ~52K particles datasets that after 3D refinement have rendered a 3.3 Å resolution density map (Fig. 5). At the moment, the presence and structure of non-symmetric components (VP1) are being analysed by symmetry relaxation and localize 3DR.

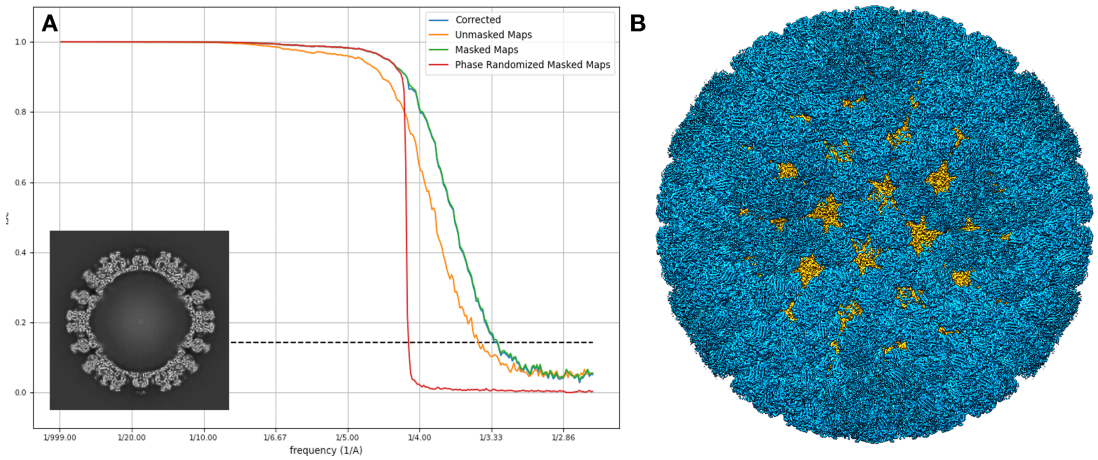

Fig. 5. MX2263-9. RVA VLP + VP1 RNA-dependent RNA-polymerase. (A) FSC resolution curve. A central section of the density map is shown in the inset. **(B)** Radial colored surface of the 3.3 Å resolution 3DR. VP2 and VP6 are couloured in yellow and blue, respectively.

MX2263-10 - *Structural characterization of the 26S:Bag1 complex by CryoE* (Jose María Valpuesta)

Session	Date	Micrographs	Ptcls	Fractions	Å/px	Images/hole
MX2363-10	30-02/07/21	14079	1,600,000	40	0.65	2

We have biochemically and biophysically determined that the cochaperone Bag1, a cofactor that tunes the activity of Hsp70 chaperones, is able to interact with the proteasome and drive the Hsp70-bound substrate to degradation. We have also found that the proteasome subunit Rpn1 is the one that connects Bag1 with the proteasome. We have prepared a complex between the 26S proteosome and Bag1 (26S:Bag1). After that, the sample was vitrified, and a large dataset was acquired in a FEI Titan Krios cryo-electron microscope.

A total of 14079 movies were recorded, and 1,600,000 particles were extracted (Fig. 6A). Particles were 2D classified and those with the best signal were selected (Fig. 6B). Then, particles were 3D classified into 10 classes generated *de novo* and those classes representing the 26S complex were selected (640000 particles). The 26S particles were subjected to a Homogeneous Refinement that produced a 6.9 Å resolution model (pixel size was 3.3 Å/px at that stage of the processing). Particles were then reextracted with a smaller pixel size (1.6 Å/px), the model was subjected to Relion Postprocess and the map obtained had 3.84 Å resolution (Fig. 6C).

Since the interaction of Bag1 with the proteasome is mediated by Rpn1, the processing was focused on the area of Rpn1, which unfortunately was one of the areas of the model with less details (arrow in Fig. 6C). A circular mask was placed around Rpn1 and a classification without alignment was performed. A total of 6 models were obtained, one of them had a high detailed Rpn1 with an extra mass close to the toroidal part (Fig. 6E, red, dotted circle).

We're now in the process of further refine the volume and in the meantime, the best correlation with our map was found with the PDB:5VFP that represents a conformation that is prepared for the substrate incoming. Regarding the local refinement map, structures of Rpn1 (6MSB) and

the UBL domain of Bag1 (2LWP), the domain involved in proteasome binding, have been docked in the latest volume (not shown).

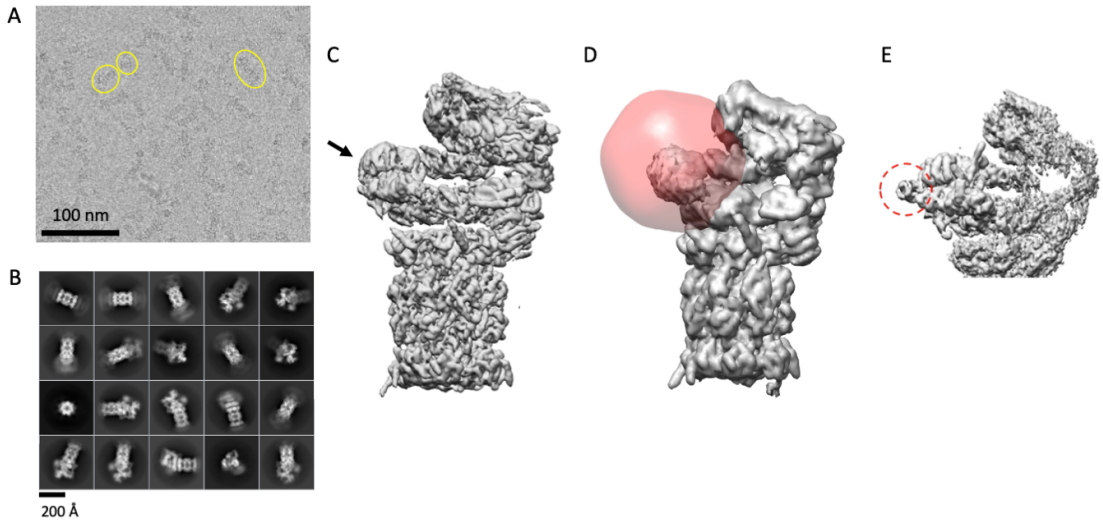

Fig. 6. MX2263-10. (A) cryoEM image of the 26S:Bag1 complex. **(B)** 2D classes. **(C)** A preliminary 3D map of the complex. The arrow points to the area where Bag1 binds. **(D)** Mask applied to that particular area in search of the mass that belongs to the UBL domain of Bag1. **(E)** Partial visualisation of the preliminary 3D reconstruction of the 26S:Bag1 complex in which the UBL domain of Bag1 is clearly observed.