STRUCTURAL STUDIES OF SV40 LARGE T-ANTIGEN ON THE ORIGIN OF REPLICATION

Maria G. Gomez-Lorenzo[1], Mikel Valle[2], Luis Enrique Donate[1], Claudia Gruss[4], Montserrat Bárcena[1], C. O. S. Sorzano[1], Michael Radermacher[5], Joachim Frank[2,3], Jose Maria Carazo[1]

[1]Centro Nacional de Biotecnología, Campus Univ. Autónoma, 28049 Madrid, Spain, [2]Health Research Inc., Howard Hughes Medical Institute at Wadsworth Center, Empire State Plaza, Albany, NY 12201-0509, and [3]Department of Biomedical Sciences, State University of New York at Albany, [4]University of Konstanz, Department of Biology, 78457 Konstanz, Germany, [5]Max-Planck-Institut fur Biophysik, Frankfurt, Germany

Large T antigen is a multifunctional phosphoprotein which is responsible for both the control of viral infection and the required alterations of cellular processes. Furthermore, it is the helicase that opens up the double stranded DNA during replication of the viral genome. T antigen is the sole viral protein required for SV40 DNA replication, all other factors are provided by the host cell. For this reason studies employing the SV40 in vitro replication system have been used to establish much of what is known about the enzymology of eukaryotic DNA replication.

The initial step of viral DNA replication is the binding of twelve T antigen monomers to the SV40 origin of replication (*ori*) forming a double hexamer. This macromolecular complex is the active helicase in charge of supplying single stranded DNA to the cellular polimerase. By using electron microscopy and negative staining, our laboratory obtained the first images of the double hexamer bound to the *ori*. Further studies using monoclonal antibodies allowed us to map several domains of the protein in the structure, as well as to propose a model of the interaction of the double hexamer with the different areas of the *ori*, model that is supported by biochemical data described in the literature (Valle *et al*., 2000).

The *ori* sequence has three very well defined regions (see figure 1A). The central part, 27 base pairs long, is a perfect inverted repeat that contains two GAGGC sequences in each arm of the palindrome. Biochemical and genetic evidence shows that the GAGGC repeats serve as the recognition site for T antigen (for a review see Borowiec *et al*., 1990). On one side of this region there is a 10 base pairs sequence partially overlapping an imperfect inverted repeat, the early palindrome (EP), and, on the other side, a 17 base pairs region rich in adenines and thymines (AT-tract). These two flanking areas undergo structural transitions during initial stages of SV40 DNA replication.

After the 2D analysis of the double hexamer at the viral *ori* described above (Valle *et al*., 2000), our goal was to obtain a 3D structure, at highest possible resolution, of this macromolecular complex, using three dimensional cryo-electron microscopy of single images. The first step was to get a low resolution volume, based on experimental data, and resolved through the random conical tilt scheme. This low resolution volume was used as a starting point for further improvement in resolution, using Radon transforms to align the experimental projections towards a 3D reference, as described in (Radermacher, 1994). The structure of the double hexamer at SV40 *ori* is shown in figure 1C. We are still working on this volume, adding more images and correcting the

contrast transfer function of the micrographs. This work will allow us to study the structure of the complex with higher degree of detail at a better resolution.

Several important conclusions can be drawn from the structure: (i) the localization of the DNA binding domain and the carboxi terminal domain of the T antigen can be done extrapolating the results with negative staining; (ii) in the junction of the two hexamers some extra masses appear coming out of the structure, probably due to the DNA that the helicase is opening; (iii) the hexamers show a severe rotation among them; (iv) each hexamer has a differentiable entity, which could reflect a distinct binding mode to the AT and EP sequences. Work is in progress with other DNA probe to assess these former findings. Taking into account that the SV40 helicase has traditionally been used to study eukaryotic DNA replication, the analysis of these structures will explain many unknown aspects of this cellular process in higher organisms.

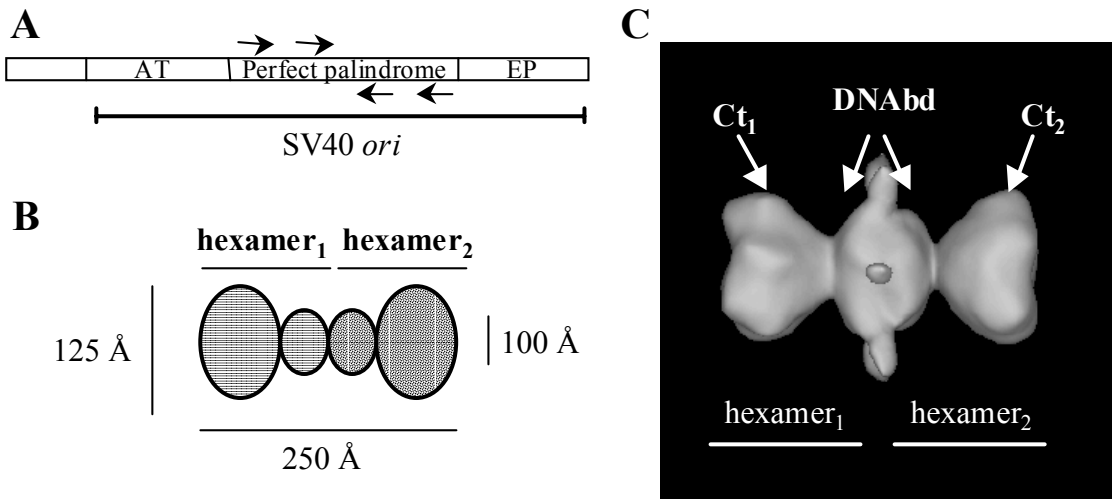

Figure 1. (A) DNA probe used to obtain a double hexamer of T antigen at the SV40 origin of replication (*ori*). The arrows represent the pentanucleotides GAGGC , specific sequence of recognition by the T antigen; AT, rich sequence in A/T ; EP, early palindrome. (B) Diagram of the dimensions of the double hexamer at the *ori*. Each hexamer was marked with a different background. (C) 3D structure of the T antigen double hexamer at SV40 origin of replication. Ct ; carboxi terminal domain ; DNAbd, DNA binding domain.

References

Borowiec, J.A., Dean, F.B., Bullock, P.A., Hurwitz, J. (1990) Cell 60(2), 181.
Radermacher, M. (1994) Ultramicroscopy 53(2), 121.
Valle, M., Gruss, C., Halmer, L., Carazo, J.M., Donate, L.E. (2000) Mol Cell Biol. 20(1), 34.

Acknowledgements

This work was supported by a CAM grant (number 08.1/0042/2000 1). Maria G. Gomez Lorenzo has a post-doctoral scholarship from the Consejería de Educacion of the CAM.

e-mail: carazo@cnb.uam.es