Published in final edited form as:

J Mol Biol. 2010 April 16; 397(5): 1276–1286. doi:10.1016/j.jmb.2010.02.042.

Conformational rearrangements of SV40 large-T antigen during early replication events

Isabel Cuesta^{1,*}, Rafael Núñez-Ramírez^{1,*§}, Sjors H. W. Scheres¹, Dahai Gai², Xiaojiang S. Chen², Ellen Fanning³, and Jose María Carazo¹

¹ Unidad de Biocomputación, Centro Nacional de Biotecnología (CNB), CSIC. C/Darwin 3, 28049 Madrid, Spain

² Molecular and Computational Biology, University of Southern California at Los Angeles, Los Angeles, California 90089, USA

³ Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37235-1634, USA

Abstract

SV40 large tumor antigen (LTag) functions as the replicative helicase and initiator for viral DNA replication. For SV40 replication, the first essential step is the assembly of LTag double hexamer at the origin DNA that will subsequently melt the origin DNA to initiate fork unwinding. Here we use three-dimensional cryo electron microscopy to visualize early events in the activation of DNA replication in the SV40 model system. We obtained structures of wild type double hexamer complexes of LTag bound to SV40 origin DNA, to which atomic structures have been fitted. Wild type LTag was observed in two distinct conformations. In one conformation, the central module containing the J-domains and the origin binding domains of both hexamers is a compact closed ring. In the other conformation, the central module is an open ring with a gap formed by rearrangement of the N-terminal regions of the two hexamers, potentially allowing for the passage of single stranded DNA generated from the melted origin DNA. Double hexamer complexes containing mutant LTag that lacks the N-terminal J-domain show the central module predominantly in the closed ring state. Analyses of the LTag C-terminal regions reveal that the LTag hexamers bound to the AT- and EP-*ori* elements are structurally distinct. Lastly, visualization of DNA density protruding from the LTag C-terminal domains suggests that oligomerization of the LTag complex takes place on double stranded DNA.

Keywords

Helicase; Large T antigen; DNA Replication; SV40; Electron microscopy

Correspondence to: Jose Maria Carazo, unidad de Biocomputación, Centro Nacional de Biotecnología (CNB), CSIC. C/Darwin 3, 28049 Madrid, Spain, carazo@cnb.csic.es; ph: +34915854543; Fax +3491372 0112.

* Both authors contributed equally to this work

§Present address: Departamento de Física Macromolecular, Instituto de Estructura de la Materia, CSIC Serrano 113 bis, 28006 Madrid, Spain.

Conflict of interest: The authors have declared that no conflict of interest exists.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

Helicases have conserved important molecular features throughout evolution due to their essential role in some of the most basic cellular functions. Among the eukaryotic helicases, the large T-antigen (LTag) of the simian virus SV40 is one of the best characterized. LTag is involved in multiple activities that enable the virus to replicate and modify host cellular regulatory processes¹. For viral DNA replication, LTag serves as the replicative helicase and initiator protein. LTag oligomerizes into a double hexamer that binds specifically to the viral origin of replication to form a pre-initiation complex, which distorts and locally unwinds dsDNA². After formation of this preinitiation complex, LTag promotes bidirectional replication from the origin of replication (*ori*) in coordination with host proteins such as DNA polymerase α -primase, topoisomerase I, and replication protein A. LTag has served as a useful model for structural and functional studies of the more complex eukaryotic and archaeal MCM helicases.

The minimal SV40 *ori* comprises 64 bp and contains three sequence elements³. A central perfect palindrome, consisting of four 5'-GAGGC-3' pentanucleotides, is flanked by a partial palindromic sequence known as the early palindrome (EP) on one site, and by an A/T-rich tract (AT) on the other side. The four pentanucleotides are the specific binding sites for LTag (reviewed in¹). In the presence of ATP or ADP, LTag assembles around the SV40 *ori* to form a double hexamer^{4; 5; 6; 7}, promoting local structural distortions on the *ori* DNA prior to ATP hydrolysis⁸. The EP element becomes partially unwound and the AT element becomes untwisted^{2; 7; 9}. Also, the highly conserved AT region may naturally assume a bent conformation which may be altered into a straight conformation^{10; 11}.

LTag is a protein of 708 amino acids composed of several functional domains^{1; 12; 13; 14}: a DnaJ homology domain, J-domain (aa 1–82), an origin DNA binding domain, OBD (aa 131–250), a helicase domain (aa 251–627) and a distal C-terminal region that participates in host-range determination. The LTag N-terminal region, termed N-TR, is formed by the J-domain and the 83–130 regions. The N-TR is involved in multiple viral activities, including viral *in vivo* DNA replication, transformation, transcriptional activation and virion assembly¹⁷. It may also interact with the OBD or other C-terminal regions of the protein²³, although it does not participate in DNA binding²⁴.

Atomic structures have been determined for three of the LTag domains. The OBD structure has been solved in various states: as a monomer²⁵, as a dimer in complex with DNA²⁶ and in a filament arrangement in crystal lattice²⁷. The filament adopts a spiral structure that contains six monomers per turn, with one monomer per asymmetric unit. The hexameric helicase domain structure has been solved in different nucleotide binding states²⁸, revealing a mechanism that couples ATP hydrolysis to conformational changes that power its helicase activity^{14; 29}. The J-domain structure has been solved in complex with another protein, the Retinoblastoma tumor suppressor³⁰, and contains the four helices characteristic of other DnaJ family members (reviewed by³¹). Electron Microscopy (EM) studies have shown that the LTag double hexamer is a bilobed structure with the two hexamers arranged in a head-to-head orientation^{4; 32} and rotated relative to each other⁶. The structure can be divided into three modules that are connected by low-density regions. The two large distal modules correspond to the C-terminal helicase domains of the two hexamers, while the third central module is formed by the OBDs and the interacting N-TR's of both LTag hexamers^{4; 6; 33}. Although the LTag N-TR does not display a stable quaternary structure in single hexamers, it is well arranged within the double hexamers after specific *ori* recognition³³. The 3D structure of the full-length LTag double hexamer assembled on the origin of replication reveals structural flexibility along the longitudinal axis of the complex, which has so far limited our understanding of the architecture and dynamics of the initiation complex that forms at the origin^{6; 34; 35}.

In the present study, we have used new experimental data and an improved image classification algorithm to handle the structural flexibility, resulting in two main conformations of the LTag double hexamer in the presence of ADP. This analysis has allowed us to detect density compatible with dsDNA within the C-terminal domains of the LTag double hexamer. These structures show that, in the presence of ADP, the C-terminal domains could assemble onto dsDNA, providing new insight into a controversial issue in the field^{28; 29; 36; 37}. To further understand the assembly of the pre-replication complex of LTag on the origin DNA, we have analyzed the structure of a mutant LTag, LTag108-627, which is defective for replication in cells but retains the ability to form double hexamers and provoke the initial distortion of the flanking sequences¹⁶. Unlike full length LTag, the mutant complex adopts only a single main conformation. The ensemble of novel structures reported here provides important mechanistic insight into the function of LTag in the initiation of DNA replication.

RESULTS

LTag 108–627 mutant displays reduced conformational flexibility

In an effort to reduce the structural flexibility of wild type (wt) LTag, which has hampered the prior structural characterization^{6; 35}, we have studied the structure of a mutant LTag lacking the N-terminal 107 and the C-terminal 81 residues, which are predicted to be structurally disordered (Fig. 1a). Importantly, this mutant (LTag108-627) is able to distort the EP- and AT-regions of the *ori* DNA¹⁶ and retains helicase activity and partial replication activity for origin-containing DNA in cell extracts (unpublished results, X.S. Chen), although the N-terminal region is required for virus replication *in vivo*^{17; 18; 19}. We recorded a cryo-EM data set of LTag108-627 double hexamers assembled on the SV40 *ori* DNA sequence in the presence of ADP. We put additional dsDNA on both sides of the core *ori* DNA sequence so that one side has a longer dsDNA extension than the other side to have an “asymmetric” dsDNA. For comparison, we also collected cryo-EM images of wt LTag double hexamers, again in the presence of ADP on a similar dsDNA probe. Interestingly, an initial, reference-free 2D analysis of both data sets indicated that the mutant complex displays a reduced flexibility compared to the wt complex (Fig. 1b).

To adequately handle the expected structural flexibility in a 3D analysis, we applied a novel 3D image classification algorithm, called MLF3D³⁸, to both data sets. MLF3D classification of the wt data yielded two different reconstructions, both with a resolution of ~30 Å (Fig. 1c) and suitable angular coverage (supplementary Fig. 1). Apart from a difference in the degree of bending along the longitudinal axis of the complex, which is in agreement with previous observations^{6; 34}, a major difference between these two new reconstructions lies at the central module. The first structure (yellow) shows a compact central module with a parallel conformation of the two main masses. The second structure (blue) contains a central module in which the two adjacent masses are displaced from one another through a swing-like movement. We will refer to the latter as the displaced conformation.

Classification of the LTag108-627 data yielded only a single conformational state, reflecting the reduced flexibility of this mutant compared to wt LTag. The final reconstruction (Fig. 1d) extends to a resolution of ~25 Å, and displays an almost straight conformation along the longitudinal axis of the complex. The two tiers that constitute the central module of this reconstruction are oriented in a parallel fashion. A superimposition of a structural model for the DNA probe on the mutant LTag double hexamer reconstruction shows that the relative positions of the two tiers match the position of the central palindromic pentanucleotide sequences that are responsible for the specific binding of LTag (Fig. 1e; green and yellow), while the flanking sequences (Fig. 1e; red and purple) are covered by the C-terminal domains. Interestingly, the positions of the largest degree of bending of the DNA (as predicted by

bent.it© server) map onto the upper and lower boundaries of the two C-terminal domains (Fig. 1e; arrows).

Architecture of the OBDs and J-domains

The improved resolution of the mutant reconstruction, together with its reduced conformational flexibility, prompted us to perform fittings of crystal structures for several of its domains. A recent crystal structure of the OBD presents six monomers in a left-handed spiral (PDB ID: [2fuf](#))²⁷. Two copies of this spiral structure, after some adjustment, were fitted into the LTag108-627 central module, following the face to face arrangement suggested previously³⁷. To obtain a satisfactory model, the relative positions of the six individual monomers in each spiral had to be adjusted manually, resulting in a small elliptical deformation of the two spiral structures (arrows in Fig. 2a). The high isotropy of the here reported reconstruction of LTag108-627, which is a consequence of the even angular distribution of projection directions, indicates that the elliptical shape of the central module is not an artifact (supplementary Fig. 2). In this fitted arrangement, the N-termini of the six OBD monomers are located near the hexamer-hexamer interface (indicated with red dots in Fig. 2b). Interactions between these termini have been shown to be crucial for double hexamer stability¹⁵. The spiral arrangement gives rise to a gap (indicated with red triangles in Fig. 2b) between the first and the last OBD monomer in each hexamer.

The interior region of the central module in the parallel conformation of the wt complex resembles the density in the mutant reconstruction, suggesting a similar face-to-face arrangement of the OBD hexamers in the wt. The difference map between the mutant and the parallel conformation of the wt central module structures revealed a belt of additional density around the wt complex (Fig. 2c). This suggests that the J-domains (absent in the mutant) are located at the periphery of the central OBDs, which is in accordance with previous observations³⁵. On the other hand, a comparison of the central module of the displaced conformation of the wt reconstruction with the mutant structure suggested important conformational changes in this region.

Interaction of the LTag C-terminal domains with the AT and EP regions

To analyze the structural differences between the LTag C-terminal domains interacting with the AT and EP regions of *ori* DNA, we exploited the asymmetry of the DNA probe, which contains additional nucleotides on the AT-rich side (Fig. 3a). After an initial reference-free 2D analysis of the wt data set, we selected those side-views that displayed a clear protruding density at one of the C-terminal ends (Fig 3b,c). Interestingly, the protruding DNA was mainly observed in complexes belonging to the straight conformation and not to the displaced one. From each of these side views, we extracted two sub-images: one for each of the two C-terminal domains (supplementary Fig. 3, procedure described in³⁸). Classification of these sub-images into two classes yielded 3D reconstructions for the C-terminal domains with a resolution of ~25Å (Fig. 3d).

As expected, the major difference between the two structures resides in the extra density protruding from the AT-associated C-terminal domains of LTag (Fig. 3d, difference map in red). This difference density extended to well over 3 standard deviations and has dimensions that are in good agreement with the expected density for the overhanging DNA. In addition, the EP-bound hexamer shows a smaller internal cavity than the AT-bound hexamer (Fig. 3d, difference map in green). Central sections through both reconstructions (Fig 3d, inserts) show clear density through the longitudinal channel compatible with the presence of two strands of DNA (Fig. 3d). However, the channel dimension in the ADP C-terminal crystal structure is narrower than the observed in these reconstructions. This size difference can be provoked by

the presence of DNA and the variable nature of the channel size depending on the different states of the complex²⁹.

A similar analysis with the LTag108-627 mutant C-terminal domains yielded two reconstructions with significantly higher resolutions of ~19Å (Fig. 4a). In this case, the difference density for the overhanging dsDNA probe was less strong than in the wt (but still extended to ~3 standard deviations), which may be caused by the shorter probe that was used for the mutant. However, the observation that 80% of the sub-images in each class had their C-terminal counterpart in the other class is an indication that a reliable classification was achieved. As observed for the wt, the AT-bound C-terminal domain shows a larger cavity than the EP-bound C-terminal domain. In both C-terminal domains, this cavity connects with open side channels (Fig. 4a; arrows), albeit at higher thresholds in the EP-associated hexamer. The resolution of these reconstructions allowed a reliable fitting of the crystal structure of the ADP-bound hexameric helicase structure (PDB ID: **1svl**) (Fig. 4b). The fitted structure explains all density except the one attributed to the dsDNA, and its side channels coincide with the channels in our reconstructions. Interestingly, the larger cavity in the AT-bound C-terminal domain is located at the position of the β -hairpin that has been proposed to interact with the DNA^{36; 39}. This hairpin lies within the density of the EP-bound hexamer, while it partially protrudes from the density into the cavity of the AT-bound hexamer (Fig. 4b).

In addition, we compared the structures of the C-terminal domains of wt and LTag108-627. The difference map between these reconstructions yields six density regions located at the periphery of the wt structure (Fig. 4c, red mesh), which could account for the extreme C-terminal region that is missing in the mutant (residues 628–708). Indeed, the residue 627 of the fitted structure locates on the surface close to these densities (Fig. 4c, atoms marked in yellow).

DISCUSSION

The central module of LTag double hexamer adopts two different conformations

Earlier 3D-EM studies of LTag on ori DNA revealed a strong bending flexibility along the longitudinal axis of the double hexamer complex, which limited the structural insights that could be obtained with 3D-EM. However, because this flexibility may be important to understand the structural dynamics of the complex, we have sought to obtain detailed structures in each of these states, and to begin to link these changes to functional activities. We approached this problem in two different ways: by using novel image classification approaches and by analyzing an LTag mutant.

We have applied a novel image classification algorithm, called MLF3D, to analyze LTag double hexamers. MLF3D is a multi-reference refinement algorithm based on maximum likelihood principles with an improved statistical model for the image formation process in the electron microscope³⁸. This classification procedure has allowed a more detailed characterization of the structural flexibility in wt LTag by identifying two structural states (Fig. 1c). The first state displays an almost straight conformation, while the second state is strongly bent along the longitudinal axis of the complex. In particular, compared to previous 3D-EM reconstructions for the double hexamers^{6; 34}, we have obtained improved density for the central module of the complex, which comprises the N-terminal region and the OBDs. Whereas the central module in the first conformation is arranged in two parallel tiers of density, the bent conformation shows two displaced tiers of density, with a reduced area of interaction. The two observed conformations of the N-terminal LTag domain can explain the general “bending pattern” described in our previous studies. The fact that we find a larger conformational change in this region in the wt LTag, which is essential for SV40 replication in cell culture^{17; 18; 19}, prompts us to postulate that this flexibility may be crucial for the initiation of DNA replication.

This hypothesis gains further support from our structural analysis of LTag deletion mutant (LTag108-627), which lacks the N-terminal 107 and the C-terminal 81 residues. This mutant is able to form double hexamers and binds DNA, but cannot proceed further to replicate DNA *in vivo*. In this case, we find that the great majority of complexes adopt a straight conformation, as opposed to the structural mixture of the wild type. Nevertheless, it has also been hypothesized that the J-domain, contained in the N-TR, is needed for interacting with protein factors for viral DNA replication inside the cells¹³. Both hypotheses are complementary and further experimental data is still needed.

Probably as a consequence of the reduced structural flexibility, the mutant reconstruction extended to significantly higher resolution than the wt structure. This allowed us to perform a more detailed analysis of the architecture of the OBD domains in the central lobe. The density map of the central lobe supports the spiral model for the OBD hexamers²⁷ rather than a flat ring arrangement. The crystal structure of six OBDs in a spiral conformation²⁷ required only a minor adjustment of the individual monomers in the form of an elliptical distortion to provide a satisfying fit to our reconstructed density. If we would attribute the density inside the hexameric OBD to dsDNA, this distortion would position OBD subunits **b-e** and **c-f** close enough to interact simultaneously with the dsDNA (Fig. 2a). However, the OBD DNA binding regions in the spiral structure are rotated from the DNA axis and therefore they are not able to interact in a DNA sequence-specific manner²⁶. In addition, our observation that these two spirals arrange within the central module in a manner that positions their gaps on the same side of the complex is in excellent agreement with the model proposed by Reese et al³⁷. In the wild type structure, the parallel conformation of the central module would resemble the face-to-face arrangement of the OBD hexamers in the mutant reconstruction, while the displaced conformation shows important conformational deviations from this arrangement, leading to a less compact structure. The N-terminal domains of the two hexamers, located around the central module of wt LTag, are crucial in determining the occurrence of both the straight and displaced conformations.

Structural difference between the hexamer bound AT- and EP-regions

The studies presented here also provide important new structural insight into the C-terminal modules and their interaction with DNA. The asymmetric design of our DNA probe allowed us to identify structural differences between the C-terminal modules that interact with the AT and EP regions of the SV40 *ori*. Classification of sub-images comprising only the C-terminal modules was driven by density for additional nucleotides on the AT-rich region of the probe. The resulting reconstructions for the AT- and EP- associated hexamers indicate that the C-terminal modules, at least for the most part, are bound to dsDNA. These results suggest that extensive DNA unwinding of the *ori*-flanking regions is not a prerequisite for the formation of the double hexamer-DNA complexes and support the interpretation that the *ori* distortion is generated by LTag oligomerization onto dsDNA^{14; 28} rather than ssDNA^{2; 36}

Permanganate footprinting established that in the presence of ADP, the EP region is more distorted than the A/T-rich region^{2; 7; 9}. These observations correlate very well with our evidence that different conformations of the C-terminal domains interact with the AT and EP regions (Fig. 5a) and suggest that the differences in LTag conformation may generate the local DNA distortion to different extents. However, we can not rule out that the differences observed could be produced by the difference in the dimension of the protruding dsDNA. For both the wt and the mutant structures, we observed a larger internal cavity in the AT C-terminal domain than in its EP counterpart. Interestingly, the fitting of the crystal structure of the ADP-bound helicase hexamer²⁹ reveals that this cavity coincides with the location of the β -hairpin that interacts with the DNA^{36; 39}. This hairpin has been observed to undergo large conformational changes in the presence of different nucleotides²⁹. Apparently, the conformation of the DNA

interaction motifs of the helicase domain differs in the EP- and AT-bound C-terminal domains. We postulate that this conformational change causes the previously observed patterns of dsDNA distortion in the *ori* flanking sequences.

The extreme C-terminal region of LTag (residues 628–708) appears to be partially structured in the LTag double hexamer complex, which is consistent with previous observations^{14; 40}. A direct role of this region in the interaction with ssDNA seems unlikely, since it is dispensable for the helicase activity of LTag⁴¹. On the other hand, it is highly exposed on the surface of the C-terminal domain at the regions of interaction with other cellular proteins, such as p53⁴² and topoisomerase I⁴³. Thereby, this localization could facilitate these interactions, explaining its role in other important processes of the viral infection, such as the host range^{41; 44; 45}.

A working model for LTag-driven origin opening

Building on previous work by our groups and others, the new structures presented in this study lead us to propose a working model for the mechanism of *ori* melting by LTag (Fig. 5). We suggest that the straight and the displaced conformations of the double hexamer would exist in equilibrium. Whereas the straight conformation would be involved in DNA association, the displaced conformation might represent a situation in which dsDNA has been melted and ssDNA is being exposed through the gap between the two distorted spiral OBD hexamers (Fig. 5c). The transition from the parallel to the displaced conformation of the central module would provoke large distortions in the bound DNA. Interestingly, we observe the DNA protruding mainly from the C-terminal module of the straight complexes, and not of the displaced ones. This DNA density was not visualized elsewhere in the displaced conformation, further suggest that the displaced conformation involves additional changes to the DNA structure, e.g. enhanced flexibility that would preclude its visualization. The transition between these two states could be the mechanism for the exit of the newly generated ssDNA strands from the complex through the OBD hexamers without disassembling them and without global disruption of N-TR contacts between the two hexamers in the central module. The adoption of an open ring structure to permit the ssDNA to exit has been proposed for other helicases^{46; 47; 48}.

Due to the complexity of these conformational changes, rearrangement of the central module could be a rate-limiting step in the unwinding mechanism. The fact that this transition is only observed in wt and not in the mutant may imply that N-TR plays a crucial role in this equilibrium. Following this reasoning, LTag mutants lacking the N-TR could undergo the same transition but less efficiently, leading to relatively few complexes in the bent conformation that would be difficult to detect in our dataset. Thereby, these mutants could retain some replication activity under the forced conditions characteristic of *in vitro* systems.

The proposed equilibrium between the parallel and displaced conformations is further supported by the disassembly of LTag double hexamers into single hexamers after treatment of the complex with single strand-specific nuclease⁶. After the digestion of the ssDNA strands extruded from the central domain in the displaced conformation, the few remaining contact points between the two OBD hexamers would not be sufficient to hold the two double hexamers together, displacing the equilibrium towards this conformation and leading to massive double hexamer dissociation. Moreover, the impairment of some of the few contact points between the two OBD hexamers would have the same effect, provoking defects in *ori* unwinding and replication *in vitro*. Thus, LTag mutants in the hexamer-hexamer interaction region (T124A and other class 4 mutants) were dominant-negative for unwinding and replication activities and unable to replicate in cell culture²¹ when a low mutant to wt LTag concentration ratio was used¹⁵.

In summary, we present here new structures of LTag double hexamer on SV40 *ori* DNA and suggest a model for the complex events that lead from *ori* recognition to unwinding during initiation of replication. The EM structures presented are consistent with the atomic models of the different domains and a large body of biochemical evidence. The use of a novel EM image classification method has allowed us to handle the structural flexibility characteristic of the LTag double hexamer, permitting the linkage of structure and function. However, further experimentation will be necessary to fully elucidate the initiation mechanism.

MATERIALS AND METHODS

EM of wt LTag and LTag108-627 DNA complexes

SV40 wt LTag was purified as reported previously⁴⁹. Using plasmid pOR1 as template⁵⁰ and primers oligo-EcoRI (5'-GAA-TTC-CCG-GGG-ATC-CGG-TCG-AC-3') and oligo-HindIII (5'-AAG-CTT-TCT-CAC-TAC-TTC-TGG-AAT-AGC-3') a 107 bp asymmetric double stranded DNA probe comprising the origin of SV40 replication was obtained by PCR. This probe was designed to extend 23 base pairs at the AT-rich region binding side of the LTag dodecamers. LTag108-627 deletion mutant was obtained by standard cloning method, and the protein was expressed and purified as previously described (Gai et al., 2004a). DNA complexes were formed with a shorter asymmetric DNA probe obtained by PCR which was designed to extend just 16 base pairs at the AT-rich region binding side in order to reduce the heterogeneity. Dodecameric complexes of mutant and wt LTag with DNA were prepared as described previously³⁵. The sample was dialyzed against the same buffer to eliminate residual glycerol. The sample was vitrified on glow-discharged Quantifoil® grids. Micrographs were collected on a Philips Tecnai F20 at 200 kV at a magnification of 50,000 under low-dose conditions.

Image processing

Unless mentioned otherwise, all image processing operations were performed using the Xmipp package⁵¹. Micrographs were scanned on a Zeiss Photoscan TD scanner and down-sampling by linear interpolation yielded a final pixel size of 5.2 and 4.2 Å for the wt LTag and LTag108-627 images correspondingly. The Contrast Transfer Function for each micrograph was estimated using CTFFIND⁵² and corrected by flipping the phases of the Fourier components. 13,321 and 9,808 Side view projections of wt and mutant LTag dodecamers were manually selected, extracted as images of 80×80 pixels and normalized. MLF2D algorithm³⁸ was initially used to evaluate the heterogeneity of the images. 3D classification and reconstructions were performed by MLF3D algorithm³⁸. A very simplistic initial model, obtained by imposing c99d2 symmetry to a reference-free 2D side view average, was used in order to erase any imposed structural information that could bias the analysis. All MLF3D classifications were performed with three classes and all calculations were repeated several times with different random seeds to check classification reproducibility. Wt LTag classifications reproducibly generated two highly populated classes of about 6,000 particles each and a single minority class of about 2,500 particles. Classification of the LTag108-627 data reproducibly produced only a single highly populated class of about 7,000 particles and two minority classes with the remainder of the particles. For the C-terminal domain classification, double hexamer images with protruding DNA were detected by KerDenSOM⁵³ of each MLF2D class. Sub-images of 40×40 pixels were extracted from each of the double hexamer image detected. An initial reference was constructed by low-pass filtering at 56 Å the atomic structure of the hexameric ADP bound LTag helicase domain (PDB ID: 1svl)²⁹, enforcing C6 symmetry. For the MLF classifications, the data were divided in 7 defocus groups, with average defocus values ranging from -3.8 µm to -6.8 µm. Resolution was assessed by the Fourier Shell Correlation criterion⁵⁴.

Available atomic coordinates for the hexameric helicase domain in presence of ADP (PDB ID: **1svl**)²⁹ and for the helical OBD domain (PDB ID: **2fuf**)²⁷ were used for docking within the obtained 3D density maps. A first automatic docking was performed using SITUS package^{55; 56}. The final location of the atomic coordinates was manually refined to fit the map density according with known biochemical data. Reconstructed maps and fittings were visualized using Chimera^{57; 58}.

Sequence analysis

The sequence of LTag (**P03070**)⁵⁹ was obtained from the ExPASy Molecular Biology Server⁶⁰. Prediction of the intrinsic ordered nature of LTag sequence were performed using web server FoldIndex©⁶¹. The web server bent.it©⁶² was used to analyze the bending pattern of the SV40 *ori* DNA sequence.

Electron Microscopy Data Bank accession numbers

The maps have been deposited at the EMBL-EBI Electron Microscopy Data Bank (EMDB). The identifier codes are 1681, 1682, 1683, 1684 for the wt LTag structures and 1685, 1686, 1687 for the LTag108-627 mutant structures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Dr. Kun Zhao for protein preparation. This work was supported by NIH grant (HL70472), MEC grant (BIO2007-67150-C03-01), CAM grant (S-GEN-0166/2006) to J.M.C, by NIH grant R01AI055926 to XSC, by a NIH (GM52948) grant to EF and by a FIS-Instituto de Salud Carlos III Postdoctoral grant (CD07/00131) to I.C.

References

1. Simmons DT. SV40 large T antigen functions in DNA replication and transformation. *Adv Virus Res* 2000;55:75–134. [PubMed: 11050941]
2. Borowiec JA, Dean FB, Bullock PA, Hurwitz J. Binding and unwinding--how T antigen engages the SV40 origin of DNA replication. *Cell* 1990;60:181–4. [PubMed: 2153460]
3. Deb S, DeLucia AL, Baur CP, Koff A, Tegtmeyer P. Domain structure of the simian virus 40 core origin of replication. *Mol Cell Biol* 1986;6:1663–70. [PubMed: 3023900]
4. Valle M, Gruss C, Halmer L, Carazo JM, Donate LE. Large T-antigen double hexamers imaged at the simian virus 40 origin of replication. *Mol Cell Biol* 2000;20:34–41. [PubMed: 10594006]
5. Mastrangelo IA, Hough PV, Wall JS, Dodson M, Dean FB, Hurwitz J. ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature* 1989;338:658–62. [PubMed: 2539565]
6. Gomez-Lorenzo MG, Valle M, Frank J, Gruss C, Sorzano CO, Chen XS, Donate LE, Carazo JM. Large T antigen on the simian virus 40 origin of replication: a 3D snapshot prior to DNA replication. *Embo J* 2003;22:6205–13. [PubMed: 14633980]
7. Dean FB, Borowiec JA, Eki T, Hurwitz J. The simian virus 40 T antigen double hexamer assembles around the DNA at the replication origin. *J Biol Chem* 1992;267:14129–37. [PubMed: 1321135]
8. Dean FB, Hurwitz J. Simian virus 40 large T antigen untwists DNA at the origin of DNA replication. *J Biol Chem* 1991;266:5062–71. [PubMed: 1848235]
9. Borowiec JA, Hurwitz J. Localized melting and structural changes in the SV40 origin of replication induced by T-antigen. *Embo J* 1988;7:3149–58. [PubMed: 2846276]

10. Han FX, Hurley LH. A model for the T-antigen-induced structural alteration of the SV40 replication origin based upon experiments with specific probes for bent, straight, and unwound DNA. *Biochemistry* 1996;35:7993–8001. [PubMed: 8672503]
11. Deb S, DeLucia AL, Koff A, Tsui S, Tegtmeyer P. The adenine-thymine domain of the simian virus 40 core origin directs DNA bending and coordinately regulates DNA replication. *Mol Cell Biol* 1986;6:4578–84. [PubMed: 3025662]
12. Fanning E, Knippers R. Structure and function of simian virus 40 large tumor antigen. *Annu Rev Biochem* 1992;61:55–85. [PubMed: 1323237]
13. Sullivan CS, Pipas JM. T antigens of simian virus 40: molecular chaperones for viral replication and tumorigenesis. *Microbiol Mol Biol Rev* 2002;66:179–202. [PubMed: 12040123]
14. Gai D, Li D, Finkelstein CV, Ott RD, Taneja P, Fanning E, Chen XS. Insights into the oligomeric states, conformational changes, and helicase activities of SV40 large tumor antigen. *J Biol Chem* 2004;279:38952–9. [PubMed: 15247252]
15. Weisshart K, Taneja P, Jenne A, Herbig U, Simmons DT, Fanning E. Two regions of simian virus 40 T antigen determine cooperativity of double-hexamer assembly on the viral origin of DNA replication and promote hexamer interactions during bidirectional origin DNA unwinding. *J Virol* 1999;73:2201–11. [PubMed: 9971803]
16. Chen L, Joo WS, Bullock PA, Simmons DT. The N-terminal side of the origin-binding domain of simian virus 40 large T antigen is involved in A/T untwisting. *J Virol* 1997;71:8743–9. [PubMed: 9343233]
17. Pipas JM, Peden KW, Nathans D. Mutational analysis of simian virus 40 T antigen: isolation and characterization of mutants with deletions in the T-antigen gene. *Mol Cell Biol* 1983;3:203–13. [PubMed: 6300656]
18. Peden KW, Pipas JM. Simian virus 40 mutants with amino-acid substitutions near the amino terminus of large T antigen. *Virus Genes* 1992;6:107–18. [PubMed: 1317074]
19. Rutila JE, Imperiale MJ, Brockman WW. Replication and transformation functions of in vitro-generated simian virus 40 large T antigen mutants. *J Virol* 1986;58:526–35. [PubMed: 3009866]
20. Campbell KS, Mullane KP, Aksoy IA, Stubdal H, Zalvide J, Pipas JM, Silver PA, Roberts TM, Schaffhausen BS, DeCaprio JA. DnaJ/hsp40 chaperone domain of SV40 large T antigen promotes efficient viral DNA replication. *Genes Dev* 1997;11:1098–110. [PubMed: 9159391]
21. Schneider J, Fanning E. Mutations in the phosphorylation sites of simian virus 40 (SV40) T antigen alter its origin DNA-binding specificity for sites I or II and affect SV40 DNA replication activity. *J Virol* 1988;62:1598–605. [PubMed: 3357207]
22. Moarefi IF, Small D, Gilbert I, Hopfner M, Randall SK, Schneider C, Russo AA, Ramsperger U, Arthur AK, Stahl H, et al. Mutation of the cyclin-dependent kinase phosphorylation site in simian virus 40 (SV40) large T antigen specifically blocks SV40 origin DNA unwinding. *J Virol* 1993;67:4992–5002. [PubMed: 8392624]
23. Weisshart K, Friedl S, Taneja P, Nasheuer HP, Schlott B, Grosse F, Fanning E. Partial proteolysis of simian virus 40 T antigen reveals intramolecular contacts between domains and conformation changes upon hexamer assembly. *J Biol Chem* 2004;279:38943–51. [PubMed: 15247253]
24. Fradet-Turcotte A, Vincent C, Joubert S, Bullock PA, Archambault J. Quantitative analysis of the binding of simian virus 40 large T antigen to DNA. *J Virol* 2007;81:9162–74. [PubMed: 17596312]
25. Luo X, Sanford DG, Bullock PA, Bachovchin WW. Solution structure of the origin DNA-binding domain of SV40 T-antigen. *Nat Struct Biol* 1996;3:1034–9. [PubMed: 8946857]
26. Meinke G, Phelan P, Moine S, Bochkareva E, Bochkarev A, Bullock PA, Bohm A. The crystal structure of the SV40 T-antigen origin binding domain in complex with DNA. *PLoS Biol* 2007;5:e23. [PubMed: 17253903]
27. Meinke G, Bullock PA, Bohm A. Crystal structure of the simian virus 40 large T-antigen origin-binding domain. *J Virol* 2006;80:4304–12. [PubMed: 16611889]
28. Li D, Zhao R, Liljestrom W, Gai D, Zhang R, DeCaprio JA, Fanning E, Jochimiak A, Szakonyi G, Chen XS. Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature* 2003;423:512–8. [PubMed: 12774115]

29. Gai D, Zhao R, Li D, Finkelstein CV, Chen XS. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* 2004;119:47–60. [PubMed: 15454080]
30. Kim HY, Ahn BY, Cho Y. Structural basis for the inactivation of retinoblastoma tumor suppressor by SV40 large T antigen. *Embo J* 2001;20:295–304. [PubMed: 11226179]
31. Hennessy F, Nicoll WS, Zimmermann R, Cheetham ME, Blatch GL. Not all J domains are created equal: implications for the specificity of Hsp40-Hsp70 interactions. *Protein Sci* 2005;14:1697–709. [PubMed: 15987899]
32. Mastrangelo IA, Hough PV, Wall JS, Dodson M, Dean FB, Hurwitz J. ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *J Nature* 1989;338:658–62.
33. Valle, MCX.; Donate, LE.; Fanning, E.; Carazo, JM. *J Mol Biol.* Vol. 357. 2006. Structural basis for the cooperative assembly of large T antigen on the origin of replication; p. 1295-305.
34. Scheres SH, Gao H, Valle M, Herman GT, Eggermont PP, Frank J, Carazo JM. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods* 2007;4:27–9. [PubMed: 17179934]
35. Valle M, Chen XS, Donate LE, Fanning E, Carazo JM. Structural basis for the cooperative assembly of large T antigen on the origin of replication. *J Mol Biol* 2006;357:1295–305. [PubMed: 16481006]
36. Kumar A, Meinke G, Reese DK, Moine S, Phelan PJ, Fradet-Turcotte A, Archambault J, Bohm A, Bullock PA. Model for T-antigen-dependent melting of the simian virus 40 core origin based on studies of the interaction of the beta-hairpin with DNA. *J Virol* 2007;81:4808–18. [PubMed: 17287270]
37. Reese DK, Meinke G, Kumar A, Moine S, Chen K, Sudmeier JL, Bachovchin W, Bohm A, Bullock PA. Analyses of the interaction between the origin binding domain from simian virus 40 T antigen and single-stranded DNA provide insights into DNA unwinding and initiation of DNA replication. *J Virol* 2006;80:12248–59. [PubMed: 17005644]
38. Scheres SH, Nunez-Ramirez R, Gomez-Llorente Y, San Martin C, Eggermont PP, Carazo JM. Modeling experimental image formation for likelihood-based classification of electron microscopy data. *Structure* 2007;15:1167–77. [PubMed: 17937907]
39. Shen J, Gai D, Patrick A, Greenleaf WB, Chen XS. The roles of the residues on the channel beta-hairpin and loop structures of simian virus 40 hexameric helicase. *Proc Natl Acad Sci U S A* 2005;102:11248–53. [PubMed: 16061814]
40. Weisschart K, Friedl S, Taneja P, Nasheuer HP, Schlott B, Grosse F, Fanning E. Partial proteolysis of simian virus 40 T antigen reveals intramolecular contacts between domains and conformation changes upon hexamer assembly. *J Biol Chem* 2004;279:38943–51.
41. Peden KW, Spence SL, Tack LC, Cartwright CA, Srinivasan A, Pipas JM. A DNA replication-positive mutant of simian virus 40 that is defective for transformation and the production of infectious virions. *J Virol* 1990;64:2912–21. [PubMed: 2159552]
42. Lilyestrom W, Klein MG, Zhang R, Joachimiak A, Chen XS. Crystal structure of SV40 large T-antigen bound to p53: interplay between a viral oncoprotein and a cellular tumor suppressor. *Genes Dev* 2006;20:2373–82. [PubMed: 16951253]
43. Khopde S, Simmons DT. Simian virus 40 DNA replication is dependent on an interaction between topoisomerase I and the C-terminal end of T antigen. *J Virol* 2008;82:1136–45. [PubMed: 18003733]
44. Stacy T, Chamberlain M, Cole CN. Simian virus 40 host range/helper function mutations cause multiple defects in viral late gene expression. *J Virol* 1989;63:5208–15. [PubMed: 2555552]
45. Poulin DL, DeCaprio JA. The carboxyl-terminal domain of large T antigen rescues SV40 host range activity in trans independent of acetylation. *Virology* 2006;349:212–21. [PubMed: 16510165]
46. Skordalakes E, Berger JM. Structural insights into RNA-dependent ring closure and ATPase activation by the Rho termination factor. *Cell* 2006;127:553–64. [PubMed: 17081977]
47. Gomez-Llorente Y, Fletcher RJ, Chen XS, Carazo JM, San Martin C. Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from *Methanobacterium thermoautotrophicum*. *J Biol Chem* 2005;280:40909–15. [PubMed: 16221680]
48. Bochman ML, Bell SP, Schwacha A. Subunit organization of Mcm2–7 and the unequal role of active sites in ATP hydrolysis and viability. *Mol Cell Biol.* 2008

49. Simanis V, Lane DP. An immunoaffinity purification procedure for SV40 large T antigen. *Virology* 1985;144:88–100. [PubMed: 2998049]
50. DeLucia AL, Deb S, Partin K, Tegtmeyer P. Functional interactions of the simian virus 40 core origin of replication with flanking regulatory sequences. *J Virol* 1986;57:138–44. [PubMed: 3001340]
51. Scheres SH, Nunez-Ramirez R, Sorzano CO, Carazo JM, Marabini R. Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc* 2008;3:977–90. [PubMed: 18536645]
52. Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 2003;142:334–47. [PubMed: 12781660]
53. Pascual-Montano A, Donate LE, Valle M, Barcena M, Pascual-Marqui RD, Carazo JM. A novel neural network technique for analysis and classification of EM single-particle images. *J Struct Biol* 2001;133:233–45. [PubMed: 11472094]
54. Saxton WO, Baumeister W. The correlation averaging of a regularly arranged bacterial cell envelope protein. *J Microsc* 1982;127:127–38. [PubMed: 7120365]
55. Wriggers W, Milligan RA, McCammon JA. Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. *J Struct Biol* 1999;125:185–95. [PubMed: 10222274]
56. Wriggers W, Birmanns S. Using situs for flexible and rigid-body fitting of multiresolution single-molecule data. *J Struct Biol* 2001;133:193–202. [PubMed: 11472090]
57. Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. *J Struct Biol* 2007;157:281–7. [PubMed: 16963278]
58. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–12. [PubMed: 15264254]
59. Reddy VB, Thimmappaya B, Dhar R, Subramanian KN, Zain BS, Pan J, Ghosh PK, Celma ML, Weissman SM. The genome of simian virus 40. *Science* 1978;200:494–502. [PubMed: 205947]
60. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 2003;31:3784–8. [PubMed: 12824418]
61. Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 2005;21:3435–8. [PubMed: 15955783]
62. Munteanu MG, Vlahovicek K, Parthasarathy S, Simon I, Pongor S. Rod models of DNA: sequence-dependent anisotropic elastic modelling of local bending phenomena. *Trends Biochem Sci* 1998;23:341–7. [PubMed: 9787640]

Figure 1.

Cryo-EM analysis of wt LTag and LTag108-627 double hexamers at the SV40 origin of replication. (a) Sequence scheme of wt LTag and LTag108-627. The N-terminal region is indicated with a red line. The plot indicates the intrinsic ordered (green) or disordered (red) nature of LTag sequences predicted by the server FoldIndex©⁶¹. (b) Reference-free 2D averages of wt and mutant LTag Cryo-EM images. Red lines indicate the orientation of the central region of each hexamer (c) Cryo-EM maps of wtLTag-ori complex structures in the parallel and displaced conformations. Annotations indicate each of the two hexamers (Hex 1 and 2) and the C-terminal (C-ter) and central regions. (d) Cryo-EM map of the single class found for LTag108-627 (e) A predicted dsDNA structure of the SV40 *ori* as calculated by the web server bent.it©⁶² is superimposed to LTag108-627 map. The central perfect palindrome is depicted in green (P1, P4) and yellow (P2, P3) and the flanking regions in red (A/T-rich

region) and in purple (EP region). Arrows point towards the edges of the flanking regions, where a marked bend is predicted.

Figure 2.

Structure of the central region of LTag. (a) Top view of the fitting of the atomic structure of two OBD hexamers into LTag108-627 central region. The OBDs are named as in ²⁷. The distances between domains **b-e** and **c-f** (~23-24Å) and between domains **a-d** (~35Å) are represented by arrows. (b) Side view of the same fitting. Red dots indicate the location of each OBD N-terminus. The minor cathetus of the triangles depict the pitch of each OBD hexamer and the hypotenuse depicts the direction of the opening of the spiral, where is located the gap. A red square indicates the localization of the LTag108-627 central region. (c) Comparison of the parallel conformation of the wt LTag central region with the LTag108-627 central region. The left panel displays the wt central region in two colors. The green color represents the mass similar to the mutant structure and the red one, the extra mass in the wt map. The right panel shows the difference map between the central regions of the wt and mutant densities displayed as a red mesh on the mutant central domain structure.

Figure 3.

Wt LTag C-terminal domain structures at the SV40 *ori*. (a) Diagram of LTag double hexamer assembled onto an asymmetric DNA probe, which extends beyond the LTag-occluded region at the AT end (blue). The probe used for wt LTag (dark blue) extends 7 nucleotides more than for LTag108-627 (light blue). The central perfect palindrome is depicted in yellow and the flanking regions in red (A/T-rich region) and in purple (EP region). (b) Raw images of LTag double hexamer complex. The hexamers with the protruding DNA have been oriented to the upper part of the image. (c) Class averages with a clear protruding DNA obtained by classification with a combination of MLF2D³⁸ and KerDenSOM⁵³. (d) Maps of wt LTag C-terminal regions bound to A/T-rich (AT-Cter) and EP (EP-Cter) regions displayed at the same threshold and their difference map. Central slices of each map are shown as insets. In the difference maps, red mesh shows extra density in AT-Cter and green mesh, extra density in EP-Cter. Bars represent 10nm.

Figure 4. LTag108-627 C-terminal domain structures. (a) Maps of LTag108-627 C-terminal regions bound to SV40 *ori*. As in Figure 3, AT-Cter and EP-Cter are indicated. Central slices of each map are shown as insets. In the difference maps, red mesh shows extra density in AT-Cter and green mesh, extra density in EP-Cter. Bar represents 10nm (b) Side cut-away views of the ADP-helicase domain hexamer structure (yellow) fitted into LTag108-627 EP-Cter and AT-Cter maps. The β -hairpins implicated in DNA interaction are shown in cyan. (c) Difference map of wt LTag versus LTag108-627 AT-Cter domains. Difference densities (red mesh) are displayed onto a side view of the LTag108-627 AT-Cter fitting. The residue 627 of each LTag helicase domain is depicted as yellow spheres.

Figure 5.

Model for the initiation mechanism of SV40 DNA replication. LTag double hexamer assembled onto *ori* sequence in the presence of ADP (a) and in the presence of ATP (b). The bubbles indicate the distortion of the DNA structure. The distortion of the A/T- rich region requires the presence of ATP and the presence of the region 121–135¹⁶, but not ATP hydrolysis^{2; 7; 9}. (c) The equilibrium between the straight and displaced conformations could be the mechanism for the ssDNA exit. (d) ATP hydrolysis powers the DNA unwinding.