Cell Reports
Article

Mechanism of Origin DNA Recognition and Assembly of an Initiator-Helicase Complex by SV40 Large Tumor Antigen

Y. Paul Chang,[1] Meng Xu,[2] Ana Carolina Dantas Machado,[1] Xian Jessica Yu,[1] Remo Rohs,[1,3,4,5,6,*] and Xiaojiang S. Chen[1,2,3,4,*]
[1]Molecular and Computational Biology Program, Department of Biological Sciences
[2]Graduate Program in Genetics, Molecular and Cell Biology
[3]Department of Chemistry
[4]Norris Comprehensive Cancer Center
[5]Department of Physics and Astronomy
[6]Department of Computer Science
University of Southern California, Los Angeles, CA 90089, USA
*Correspondence: rohs@usc.edu (R.R.), xiaojiang.chen@usc.edu (X.S.C.)
http://dx.doi.org/10.1016/j.celrep.2013.03.002

SUMMARY

The DNA tumor virus Simian virus 40 (SV40) is a model system for studying eukaryotic replication. SV40 large tumor antigen (LTag) is the initiator/helicase that is essential for genome replication. LTag recognizes and assembles at the viral replication origin. We determined the structure of two multidomain LTag subunits bound to origin DNA. The structure reveals that the origin binding domains (OBDs) and Zn and AAA+ domains are involved in origin recognition and assembly. Notably, the OBDs recognize the origin in an unexpected manner. The histidine residues of the AAA+ domains insert into a narrow minor groove region with enhanced negative electrostatic potential. Computational analysis indicates that this region is intrinsically narrow, demonstrating the role of DNA shape readout in origin recognition. Our results provide important insights into the assembly of the LTag initiator/helicase at the replication origin and suggest that histidine contacts with the minor groove serve as a mechanism of DNA shape readout.

INTRODUCTION

SV40 large tumor antigen (LTag) transforms eukaryotic cells and is essential for viral DNA replication. SV40 replication involves essential cellular replication proteins, including primase and polymerase proteins (Fanning and Zhao, 2009). To initiate eukaryotic DNA replication, multiple initiator proteins, such as Orc, cdc6, cdt1, and GINS, are required for origin binding and helicase recruitment/activation (Méndez and Stillman, 2003). For SV40 replication, however, LTag alone fulfills the functions of these multiple initiator proteins, i.e., origin recognition, melting, and unwinding (Simmons, 2000). Thus, LTag is an integrated initiator and replicative helicase for DNA replication. LTag has three defined domains for replication: an origin binding domain (OBD), a Zn domain, and an AAA+ domain (Gai et al., 2004; Li et al., 2003; Singleton et al., 2007; Figure 1A).

The SV40 core origin DNA for replication (ori) can be divided into two halves (Figure 1B), with each half containing two of the four 5′-GAGGC pentanucleotides (PEN1–PEN4) and an AT-rich (AT) or early palindrome (EP) region (Deb et al., 1986). Each PEN can be bound by one OBD (Bochkareva et al., 2006; Deb et al., 1987). Each half origin supports the assembly of one LTag hexamer, and the full origin supports double hexamer formation (Mastrangelo et al., 1989; Valle et al., 2006). The assembly of the LTag hexamer/double hexamer at the replication origin is coupled with ori DNA melting and unwinding (Borowiec et al., 1990; Borowiec and Hurwitz, 1988; Gai et al., 2004; Joo et al., 1998; Li et al., 2003; Mastrangelo et al., 1989; Shen et al., 2005; Sreekumar et al., 2000; Valle et al., 2006).

Despite advances in characterizing the LTag helicase domain structure and the structure of individual OBDs interacting with the PEN origin sequence (Bochkareva et al., 2006; Meinke et al., 2007), information is lacking regarding how the OBD, Zn domain, and AAA+ domain (the helicase domain) together recognize each half of the ori during the assembly of an LTag hexamer. Thus, an LTag structure containing OBD, Zn, and AAA+ domains can address the problem of the origin recognition and assembly mechanism in a way that cannot be addressed by studying the separate OBD or AAA+ helicase domains. Here we describe the crystal structure of the EP half origin bound by a dimeric LTag construct that contains OBD, Zn, and AAA+ domains. Our structure reveals several unexpected features in the protein-ori DNA interactions, including the inversion of a domain to contact ori DNA, a previously unidentified ori sequence for OBD recognition, and a particular DNA structural trait that is critical for recruiting the initiator/helicase (i.e., shape readout for DNA-protein recognition). Our results provide detailed mechanistic insights into how LTag initiator/helicase assembles around ori DNA, which should have broad implications for understanding the initiation of replication in other eukaryotic replication systems.

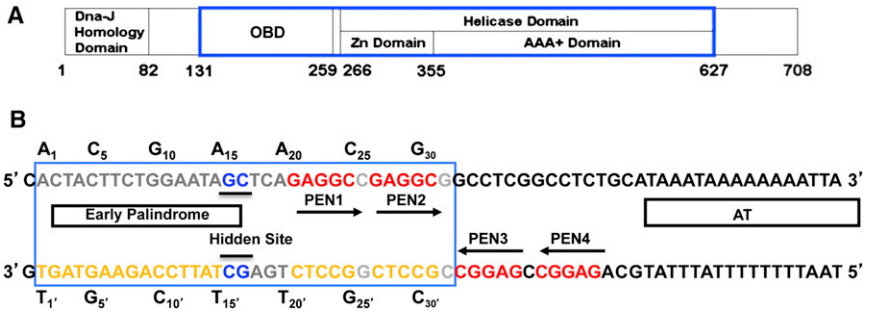

Figure 1. SV40 LTag Domain Structures and Core Replication Ori DNA Sequence

(A) Representation of the LTag domains. The construct used for cocrystallization is boxed in blue (LTag131-627), which contains the OBD, Zn domain, and AAA+ domain.

(B) Core ori DNA sequence of SV40. Each of the four pentanucleotide GAGGC sequences (PEN1–PEN4, labeled in red) is recognized by an OBD, and the EP and AT regions are indicated. The hidden-site sequence (GC, colored in blue) is indicated here and discussed in the text. The EP-half ori DNA (boxed in blue) was used for cocrystallization.

The detailed insights into LTag-DNA binding provided by our structure reveal a critical role of histidine residues in protein-DNA recognition. We observe that the histidine residue of the AAA+ domain interacts with the ori DNA using a mechanism similar to that previously observed for arginine residues (Rohs et al., 2009). On the basis of the analysis of all available crystal structures of protein-DNA complexes, we previously found that arginines can recognize minor groove shape through a shape-dependent electrostatic potential. Here, using cocrystal structures of other protein-DNA complexes, we demonstrate that histidines can play a similar role in DNA shape readout.

RESULTS

Overall Architecture of the Dimeric LTag-dsDNA Complex

The LTag construct (residues 131–627) used for cocrystallization with ori DNA contains three separable domains: the OBD, the Zn domain, and the AAA+ domain (Figure 1A, blue box; Figure S1). This LTag131-627 construct ("LTag" hereafter; Figure S2) is crystallized as a dimer in complex with the 32bp EP-half ori double-stranded (dsDNA; Figure 1B, blue box). The space group is $P2_1$, with each asymmetric unit containing two copies of the dimer-DNA complex, which allowed 2-fold noncrystallographic symmetry (NCS) averaging to yield an excellent electron density map during structural determination.

The electron density corresponding to the entire 32bp EP-ori DNA is sufficiently well featured to allow unambiguous assignment of the DNA nucleotide sequence (Figures 2A and 2B). As a result, the orientation and register of the EP-ori DNA in the complex are well defined. The average rise and helix twist between adjacent base pairs are 3.4 Å and 35.0°, respectively, and thus are close to those reported for standard B-form DNA (3.32 Å and 35.4°; Olson et al., 1998). The EP-ori DNA interacts with all three LTag domains (OBD, Zn, and AAA+; Figures 2E and 2F, blue arrows), with each domain having unique binding features as described in later sections.

Sequence-Specific Recognition of Ori DNA by OBDs of the LTag Dimer

One of the most surprising observations we made in the LTag-ori DNA complex structure was the manner in which the two OBDs of the dimer recognize the PEN sequences. Instead of the anticipated binding of PEN1 and PEN2 by OBD1 and OBD2 of the two subunits, only PEN1 is bound by OBD1, leaving PEN2 untouched (Figures 2D and 2E; OBD1 in pink, PEN1 and PEN2 labeled). Unexpectedly, OBD2 interacts with a GpC dinucleotide ($G_{16}C_{17}/C_{16'}G_{17'}$, termed "hidden site" hereafter; Figures 1B and 2D). Equally surprising was the finding that OBD2 is inverted by almost 180° compared with OBD1 (Figures 2D and 2F), allowing the two OBDs to bind the DNA from opposite faces with orientations inverted relative to each other (Figures 2C and S1) even though the AAA+ domains of the dimer are rotated by only ~60° relative to each other. The flexible linker (residues 258–267) connecting the OBD and Zn domain allows this inversion of OBD2. However, structural modeling indicates that the linker is too short to allow OBD2 in the dimer to reach PEN2, even in its fully extended conformation.

OBD1 interacts with PEN1 in the major groove in a fashion similar to the OBD-PEN interaction reported previously (Bochkareva et al., 2006; Meinke et al., 2007). In particular, the loop region between α1 and β1, and residues just N-terminal of α3 together create a surface that interacts with PEN1 in the major groove (Figures 2E, 3A, and S3A). PEN1-$G_{21}A_{22}G_{23}G_{24}$ is in contact with R154, S152, and N153 (Figure 3A, highlighted). Additionally, the two consecutive cytosines ($C_{24'}C_{23'}$) of the reverse strand of PEN1 interact with the OBD1 protein backbone, whereas the 5′ guanine ($G_{25'}$) of the complementary strand interacts with R204 (Figure 3A). These contacts, especially the bidentate hydrogen bonds between arginine and guanine bases, lead to highly specific interactions with PEN1, which is frequently used for sequence readout through base-specific contacts (Rohs et al., 2010). Overall, 7 bp of PEN1 are recognized by OBD1 through base readout (Figure S1).

The unexpected OBD2 binding to the hidden site involves substantial interactions with the major groove edges of the $C_{16'}G_{17'}$ dinucleotide and its complementary sequence, $G_{16}C_{17}$ (Figure 3B, hidden site $G_{16}C_{17}/C_{16'}G_{17'}$ labeled in Figures 1B and 2D). Therefore, the interactions with the hidden site are sequence specific, but with less specificity than the OBD1-PEN1 interactions. Because of the lack of a full PEN binding site for OBD2, R154 that interacts with the G_{21} base in OBD1 swings away from the DNA in OBD2 and forms a hydrogen bond with N227 (Figure 3B). Similarly, S152 that binds the A_{22} base in OBD1 swings away from the DNA in OBD2 (Figure 3B). The loop region containing A149, V150, and F151, which interact with the DNA backbone in the case of OBD1, reorients away from the DNA backbone in OBD2 (Figures 3B and S3B).

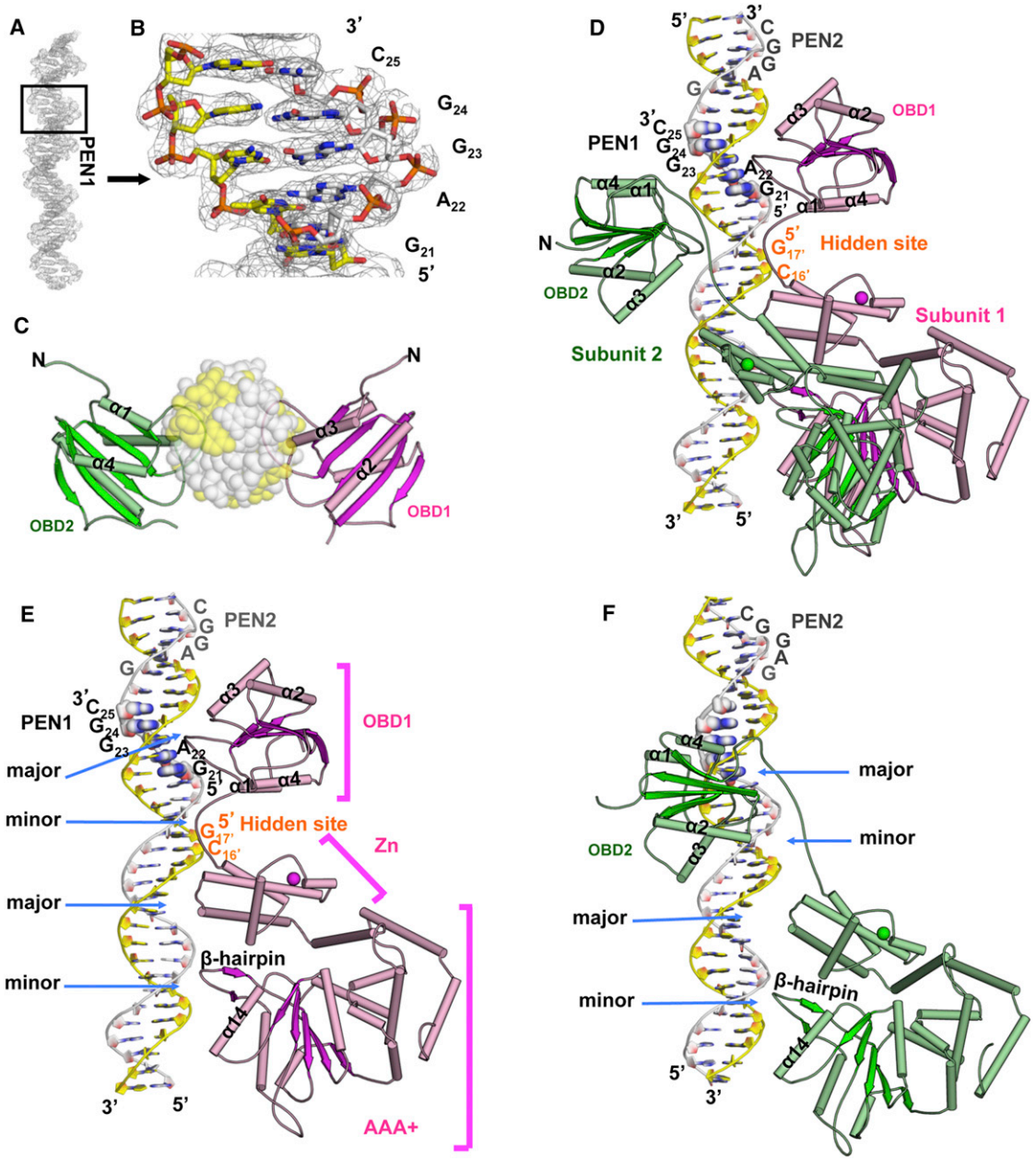

Figure 2. Overall Structure of the LTag Dimer in Complex with EP-Ori

(A) The well-defined electron density map of the EP-ori DNA shows a very clear electron density for the base pairs.

(B) The map section corresponding to PEN1 in (A) is enlarged to show the excellent fitting of the DNA model into the density.

(C) Top view of the two OBDs (drawn in ribbons) on the opposite faces of the ori DNA (drawn in surface representation) around PEN1. The two DNA strands are shown in the same color scheme as in (B), with yellow for one strand and gray for the other.

(D) Overall structure of the dimeric LTag-ori DNA complex. OBD1 binds PEN1 (highlighted), but OBD2 does not bind to PEN2. Instead, OBD2 binds to a hidden site near PEN1 but distal to PEN2. The helicase domains of the two LTag subunits are arranged similarly to two adjacent subunits in LTag hexamers.

(E and F) The interactions between EP-ori and the OBD, linker region, and Zn and AAA+ domains are shown along the dsDNA's consecutive major/minor-groove interfaces (indicated by arrows) in the two subunits (subunit 1 [E] and subunit 2 [F]). Note the difference in the linker conformations and the OBD orientations of the two subunits.

See also Figures S1, S2, and S6.

Previous biochemical work showed that LTag double hexamers prefer to bind to PEN1 for the EP half ori (and PEN3 for the AT half; Joo et al., 1998). For the LTag double-hexamer assembly, PEN1 (on the EP-ori DNA) and PEN3 (on the AT-ori DNA) alone are sufficient, and PEN2 (EP-ori) and PEN3 (AT-ori) are dispensable (Joo et al., 1998). Our structural data are

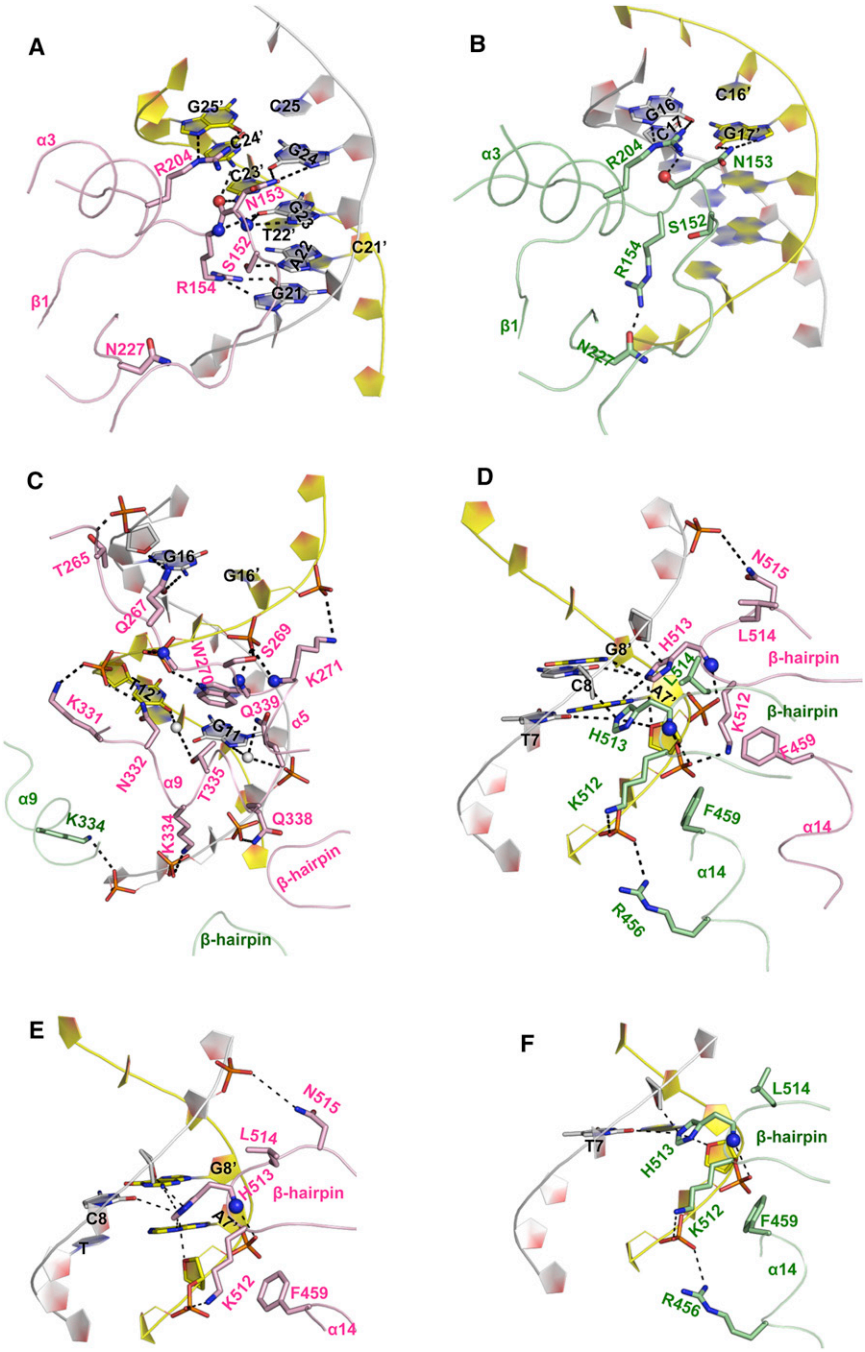

Figure 3. Detailed Interactions between LTag and EP-Ori DNA

In all panels, the leading DNA strand is in gray, the reverse strand is in yellow, subunit 1 is in pink, and subunit 2 is in green. Protein backbone oxygen and nitrogen atoms are represented by red and blue spheres, respectively.

(A) Sequence-specific interactions between OBD1 and PEN1.

(B) Interactions between OBD2 and the hidden-site sequence 5′-$G_{17'}C_{16'}$ and 5′-$G_{16}C_{17}$ (see Figure 1B).

(C) Interactions of the Zn domains of both subunits with EP-ori DNA, which is mediated by charge-charge interactions with the DNA backbones. Water molecules are represented by white spheres.

(D) Interactions between the AAA+ domains of both subunits and the EP region, which are mainly mediated by residues on the β-hairpin (K512/H513). Note that the H513 residues of both subunits insert into the minor groove to interact with both strands, whereas the K512 residues of both subunits track only one strand.

(E and F) The same interactions shown in (D), except that subunit 1 (E) and subunit 2 (F) are shown separately.

See also Figures S2, S3, and S5.

consistent with these results in that the LTag dimer assembly binds to PEN1 and not PEN2. It has also been shown that PEN1, PEN3, and the EP region together constitute a strong assembly unit for two head-to-head hexamers, whereas PEN2, PEN4, and the AT region constitute a weak alternative assembly unit (Sreekumar et al., 2000). We were not able to obtain a crystal of dimeric LTag with the AT-half ori DNA under the tested conditions even after extensive exploration, probably reflecting a different strength of the protein interactions with the AT-half region.

Zn Domain Interactions with Ori DNA

Interactions between the Zn domains and the EP-ori DNA occur mainly through the two DNA backbones via charge-charge interactions in the major groove (Figures 2E and 2F). The Zn domain of subunit 1 (Zn-1) contributes most of the binding interface with DNA, with a buried surface of 980 $Å^2$ (versus 130 $Å^2$ for subunit 2). The Zn domain of subunit 2 (Zn-2) has only one residue that binds the DNA backbone; all other interactions are through Zn-1 (residues in pink for Zn-1 and green for Zn-2 in Figure 3C). A total of 11 residues from Zn-1 contact the DNA backbones in a nonsequence-specific manner (Figure 3C). Such extensive interaction with the DNA backbone through the Zn domain may provide an anchoring point on the DNA that allows the AAA+ domain to bind the AT EP region for assembly.

AAA+ Domain Interactions with DNA

Unlike the interactions of the two OBDs with ori DNA at the major groove, the AAA+ domains of both subunits 1 and 2 interact with the minor groove (Figures 3D–3F). For subunit 1, the four residues at the β-hairpin tip (K512, H513, L514, and N515) and the adjacent helix (F459 on α14) contact DNA (Figure 3E). Together, these residues interact with three phosphate groups, two sugar

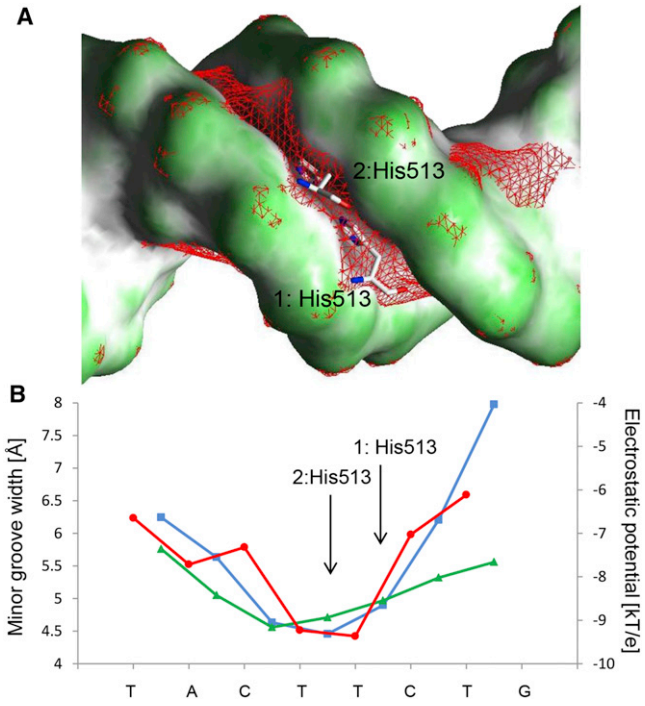

Figure 4. Narrow Minor-Groove Geometry of the 5′-$C_5T_6T_7C_8T_9$ Region of EP-Ori DNA where Both β-Hairpin H513 Residues Bind

(A) The shape of the molecular surface is shown with GRASP2 (concave surfaces in dark gray; convex surfaces in green; Petrey and Honig, 2003). The red mesh represents an isopotential surface at −5 kT/e, calculated with DelPhi at a physiologic ionic strength of 0.145 M (Rocchia et al., 2002). The H513 residues from both subunits intrude into the minor groove in a region with enhanced negative electrostatic potential as a result of narrowing the width of the minor groove to 4.5 Å (with an electrostatic potential of −9.4 kT/e) from the normal width of 5.8 Å (with an electrostatic potential of −7.2 kT/e (see B).

(B) The minor-groove width of bound ori DNA in our crystal structure (blue) and unbound ori DNA predicted in MC simulations (green), and the electrostatic potential in the center of the minor groove calculated with DelPhi (red) illustrate that the H513 residues bind a region with an intrinsically narrow minor groove. The enhanced negative electrostatic potential in the narrower groove region attracts H513 residues through favorable electrostatic interactions, a mechanism known as shape readout.

See also Figure S4.

moieties, and three bases (C_8, $G_{8'}$, and $A_{7'}$) through a combination of hydrogen bonding and electrostatic and hydrophobic interactions. For subunit 2, the hydrophobic residues (F459 and L514) pack against the DNA backbone (Figure 3F), and charged residues K512, H513, and R456 interact with the phosphate backbones.

Electrostatic Interaction Anchors a Histidine Pair in the Minor Groove

Of particular interest is how the H513 residues of the two subunits interact with EP-ori DNA. It is evident that, for the most part, the two subunits contact the ori DNA differently (compare Figures 2E and 2F). However, the H513 residues on the β-hairpin from both subunits bind to the minor groove side by side in a nearly identical fashion with both imidazole rings lying in approximately the same plane, following the helical path of the minor groove. The N-N distance between the two H513 imidazole groups is ∼2.7 Å, a distance indicative of the formation of a hydrogen bond (Figure 4A), which may further stabilize the interactions within the ternary complex of the LTag dimer and ori DNA.

Importantly, the H513 residues anchor at the position of the ori DNA that was shown to be melted upon hexamer assembly (Borowiec and Hurwitz, 1988). The minor groove region bound by the two H513 side chains is narrower than its adjacent regions (Figure 4B, blue line), with a minimum width of 4.5 Å (versus 5.8 Å for the minor groove width of standard B-DNA). This narrower width of the minor groove could be induced by protein binding, or it could be an intrinsic structural feature of the ori DNA sequence. To distinguish between these two possibilities, we carried out Monte Carlo (MC) simulations (see Experimental Procedures for details) of the DNA structure using the origin sequence. The result indicates that this DNA region of the H513 contacts is characterized by an intrinsically narrow groove in the absence of protein binding (Figure 4B, green line).

The negative electrostatic potential in the center of the minor groove is enhanced as the groove width decreases (Rohs et al., 2009). The narrowed minor-groove region where the H513 residues bind has an electrostatic potential that is ∼2 kT/e more negative than the potential in the wider minor groove of adjacent regions (Figures 4A, red mesh, and 4B, red line). Thus, the binding of H513 residues to the EP-ori sequence is characterized by a shape readout mechanism whereby positively charged protein residues bind to intrinsically narrow regions of the minor groove with enhanced negative electrostatic potential (Rohs et al., 2009, 2010).

After observing the origin recognition mode of the LTag His513 residue in this structure, we sought to determine whether the observation that histidine residues recognize narrow minor groove widths and enhanced negative electrostatic potential is of a more general nature. We analyzed the minor groove width and electrostatic potential for various structures that are part of the IFN-β enhanceosome (Escalante et al., 2007; Panne et al., 2004, 2007). This analysis revealed that conserved histidine residues from IRF-3 (His40) and IRF-7 (His46) that intrude into the minor groove consistently bind regions of narrow minor groove and enhanced negative electrostatic potential (Figure S4).

Conformation of LTag Subunits in the Dimer Structure

Our LTag dimer-DNA cocrystal structure reveals how the multidomain of LTag that contains OBD, Zn, and AAA+ domains in a single polypeptide is arranged when it binds to DNA. The structure reveals that, besides the orientation difference between the two OBDs, the two subunits also show different relative orientations between the Zn and AAA+ domains (Figures 5A–5D). In particular, α10 of subunit 1 (magenta in Figure 5A), a long helix connecting the Zn and AAA+ domain, shifts away from the Zn domain by 6 Å compared with the conformations in nucleotide-free, ADP-bound, and ATP-bound states (Figure 5A), whereas α10 of subunit 2 is rotated toward the Zn domain by 13° compared with other LTag states (Gai et al., 2004). These displacements of α10 alter the interactions between the Zn and AAA+ domains (Figures 5B and 5C),

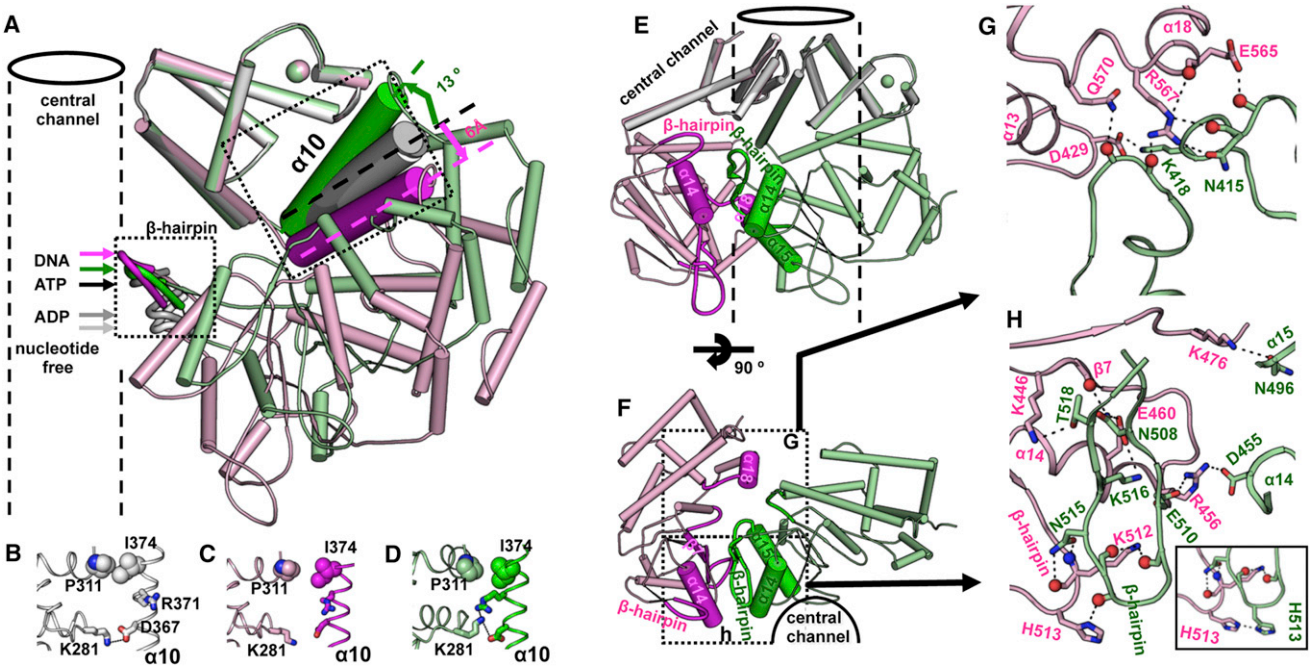

Figure 5. Helicase Domain Conformations and Dimer Interface

(A) Superimposition of helicase domains in various nucleotide-bound states (white) and the two subunits in the DNA-bound state from this structure (pink and green), which reveals shifting of the AAA+ domains and β-hairpins, illustrating the ability of multiple conformational switches that are critical for a motor protein such as LTag to melt origin and unwind DNA. Note that α10 adopts different orientations for subunits 1 (green) and 2 (pink), which generates a rotation of the AAA+ domains to bring their β-hairpins (and H513 residues) into close proximity to the DNA minor groove (see the box in H here, and Figure 3D).

(B–D) Change of contacts between Zn (K281/P311) and AAA+ domains (D367/R371/I374 on α10) in the absence (B: close contact) or presence (C: detached; D: altered contact) of DNA.

(E and F) Two views of the protein dimer interface show the tight interface between the two subunits around the β-hairpin regions in the AAA+ domain.

(G and H) Close-up views of the extensive interactions between the two AAA+ domains. The close proximity of the two H513 residues is shown in the inset (H).

resulting in different degrees of rotation of the AAA+ domain, which brings the β-hairpins of both subunits into close proximity to insert into the minor groove of the EP-region of DNA (Figures 5A and 5H, boxed region). Also, the positions of the β-hairpins move upward toward the Zn domain even more so when compared with those in various nucleotide states (Figure 5A, boxed), which suggests that ori DNA binding induces conformational changes even greater than those caused by nucleotide binding in the LTag hexamer.

Interestingly, some residues buried within the interface between neighboring subunits as shown in previous LTag hexamer structures are now exposed on the dimer's outer surface and form contacts with DNA. For example, N515 and K271 of subunit 2 are buried at the dimer interface with subunit 1, and interact with subunit 1 in the hexamer. However, N515 and K271 of subunit 1 are on the exposed dimer surface and interact with DNA (Figures 3C and 3D). Conversely, R456 of subunit 1 is part of the dimer interface that interacts with subunit 2 (Figure 5H), but R456 of subunit 2 is located on the exposed surface and utilized for DNA binding (Figure 3F). Thus, these protein residues that are involved in protein-protein interactions for hexamer formation can also be used for DNA binding during helicase assembly. Such dual binding to either another protein subunit or the DNA of these LTag residues may have implications for the assembly of LTag subunits on ori DNA.

Dimer-DNA Complex: An Assembly Intermediate without Melting Activity

In the LTag dimer-ori DNA complex, the DNA is in B-form conformation. The obviously narrower width of the minor groove where the β-hairpins bind appears to be a pre-existing structural feature of the ori DNA (Figure 4). No obvious deformations of the DNA are induced by LTag binding. This observation is similar to previous findings from the crystal structures of archaeal ORC initiator-dsDNA complexes (Dueber et al., 2007; Gaudier et al., 2007), which revealed no severe deformation or melting of DNA. In contrast, biochemical studies of E1 helicase/initiator, a distant homolog of LTag from papillomavirus, suggested that a trimer assembly intermediate is capable of melting ori DNA (Schuck and Stenlund, 2005). No trimer assembly for LTag has ever been observed. However, given the LTag dimer assembly intermediate observed here, we asked whether the stable dimer intermediate could be captured in vitro and, if so, whether such a dimeric intermediate would be similar to E1 in terms of its ability to melt ori DNA.

To address this question, we designed specific mutations to capture a stable dimer intermediate of LTag in solution, as it was previously shown that wild-type (WT) LTag only exists in either a stable monomeric or hexameric form, and no stable dimer or other intermediate oligomers can be observed in solution (Gai et al., 2004; Li et al., 2003). Two mutants (mut1:

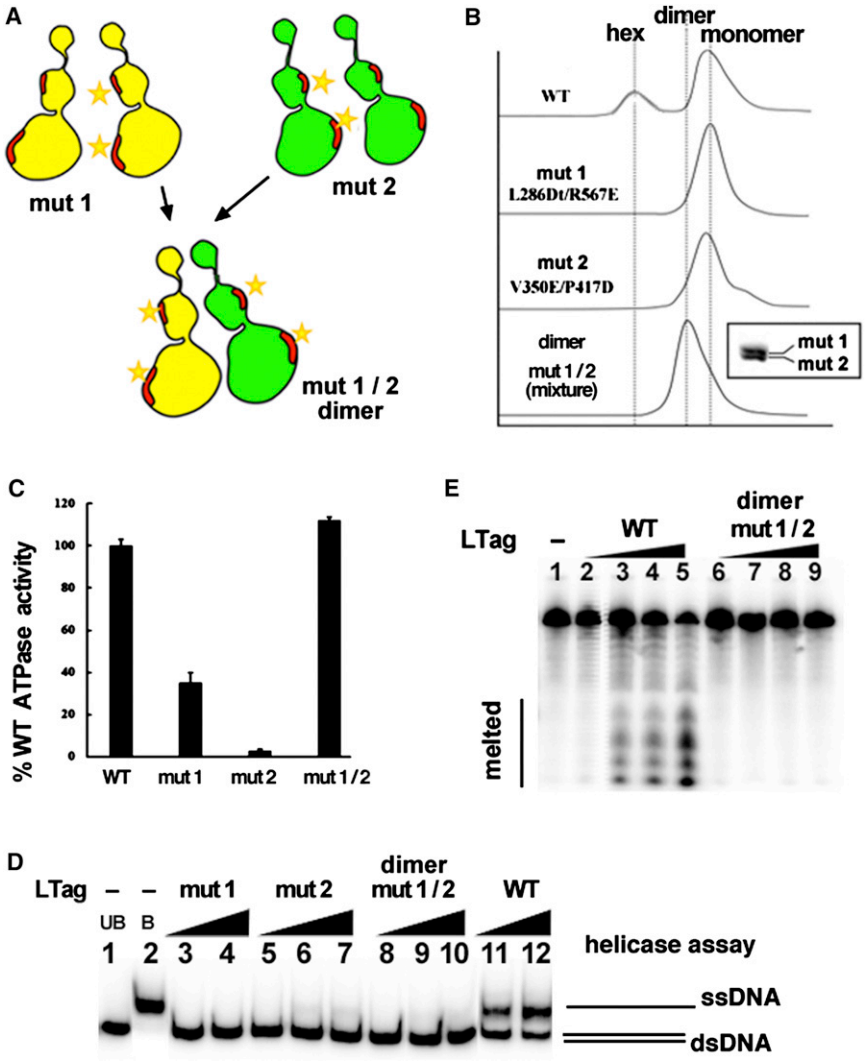

Figure 6. Generation of a Stable LTag Dimer Intermediate and Analysis of its Activities

(A) Strategy to obtain a stable dimer intermediate of LTag through two independent mutations. Mut1 carries V350E/P417D mutations on one side of a subunit, and mut2 has L286Dt/R567E mutations on the opposite surface of LTag131-627. Both mutants are predicted to exist only as monomers when alone, but as a stable dimer when mixed in an equimolar ratio.

(B) Superdex-200 column chromatography, showing that mut1 and mut2 alone exist in monomeric form, but when mixed in equimolar ratios, they form stable dimer intermediates. Inset: the two mutants in the dimer peak were detected in a 1:1 ratio by SDS-PAGE, with mut1 being slightly larger due to a C-terminal extension. Note that WT equilibrates between monomeric and hexameric forms; no dimer or any other intermediate oligomeric form can be detected.

(C) ATPase activity assay, showing a recovery of ATPase activity only when mut1 and mut2 are mixed, indicating a WT ATP pocket at the dimer interface.

(D) Ori DNA unwinding assay, showing that only WT LTag131-627, and not the mutants, can unwind the ori DNA. Lane 1: unboiled blunt-ended ori DNA alone (UB); lane 2: boiled ori DNA (B); lanes 3 and 4: mut1 alone; lanes 5–7: mut2 alone; lanes 8–10: the dimer of mut1 and mut2; lanes 11 and 12: WT LTag131-627.

(E) Ori DNA melting assay by $KMnO_4$ reactivity, showing that WT LTag131-627, but not the mutant dimer, has ori-melting activity. Radiolabeled 92 bp ori containing dsDNA was incubated with increasing quantities of LTag protein (lanes 2–5 and lanes 6–9: 100, 200, 400, and 800 ng for WT and mutant dimer, respectively). Lane 1 has no protein added.

See also Figure S5.

L286D/R567E, and mut2: V350E/P417D) are designed to introduce mutations on only one surface of each mutant within the interface, so that intersubunit interactions are disrupted (Figure 6A). We predicted that, when present alone, these two mutants would be monomeric. When mixed together, however, each mutant contributes one native side to form a dimer. The two outer surfaces of the dimer carry mutations to prevent further oligomerization (Figure 6A), thus forming a stable dimer. Gel filtration chromatography results indicated that each individual mutant was indeed monomeric, and the equimolar mixture of the two mutants formed stable dimers (Figure 6B). As a convenient way to distinguish the two mutant proteins within the stabilized dimer intermediate, we added a 14 amino acid C-terminal tail on mut1 (L286D/R567E) to make the protein ~2 kD larger (slower migration on SDS-PAGE), denoted as L286Dt/R567E. SDS-PAGE analysis of such a dimer revealed a 1:1 stoichiometry of mut1:mut2 based on gel quantification (Figure 6B, inset).

To further validate the intersubunit interface of the dimer, we conducted an ATPase assay to test whether the ATP pocket that forms at the dimer interface can hydrolyze ATP. The results showed clearly that each mutant alone had severely disrupted ATPase activity, as expected for a monomeric form, whereas the mixture of the two mutants had WT-level ATPase activity (Figure 6C), which indicates that the stabilized mutant dimer intermediate reconstitutes a WT ATP pocket at the dimer interface, as in the case of a hexamer of the WT protein. A helicase assay showed that the two LTag131-627 mutants individually or as a mixture no longer unwound the ori DNA (Figure 6D, lanes 3–10), consistent with the fact that only functional hexamers of WT LTag can unwind DNA (lanes 11–12).

With the stabilized dimer intermediate at hand, we assayed the origin-melting activity of the dimer intermediate in the presence or absence of ATP. We used the potassium permanganate ($KMnO_4$) reactivity assay on a 92 bp SV40 ori containing DNA substrate. The results showed that the dimer intermediate had no detectable DNA melting activity (Figure 6E, lanes 2–5), but the WT LTag131-627 showed the expected ori DNA melting activity (Figure 6E, lanes 6–9). Thus, the dimeric state represents an assembly intermediate that can recognize or bind to ori DNA but is not capable of melting it.

DISCUSSION

In this work we describe the cocrystal structure of a dimeric LTag in complex with the EP-ori DNA, which represents the initial stage of assembly toward a hexameric initiator/helicase complex at the ori DNA for replication initiation. This crystal structure reveals the multiple domain organization of the OBD, Zn domain, and AAA+ domain in one polypeptide of LTag in complex with ori DNA. The detailed molecular interactions of LTag with the EP-ori DNA provide mechanistic insights into origin recognition, recruitment, and assembly of LTag at the initial stage of viral DNA replication in eukaryotic cells.

Importance of the Zn Domain in Assembly of LTag at the Replication Origin

The Zn domain is essential for hexamerization of the helicase domain (Li et al., 2003). In this dimer-DNA structure, the two Zn domains interact with each other in a manner similar to those observed for LTag hexamer structures in various nucleotide states (Gai et al., 2004; Li et al., 2003). Thus, the relative orientation between adjacent Zn domains appears to be independent of the different states in ATP binding, hydrolysis, DNA interaction, and LTag assembly stages.

The LTag dimer-DNA structure reveals that the Zn domain and DNA interface accounts for ~25% of the total protein-DNA interface, suggesting that it plays an important role in the initial assembly of the LTag dimer at the origin. Thus, for LTag assembly around the ori DNA, the Zn domain serves two essential functions: (1) to stabilize the intersubunit interface of the dimer, and (2) to reinforce the OBD-initiated, sequence-specific origin binding through interactions with the DNA backbones.

Role of β-Hairpin in Ori DNA Binding

Even though the interactions of the two LTag subunits with the EP-ori DNA are largely different from each other in all three domains (i.e., the OBD, Zn, and AAA+ domains), the β-hairpin tip residues (K512 and H513 of the two subunits) form almost identical interactions with DNA in the minor groove of the EP sequence region. These interactions may have several implications.

First, the β-hairpins on the two subunits intrude into the minor groove side by side, following the helical path of the groove. This arrangement is stabilized through electrostatic interactions with the DNA and, likely, a hydrogen bond between the H513 residues of both subunits. Even though the details differ, the helical arrangement of the two β-hairpins in this LTag dimer resembles that in the E1 hexamer structures in which the six equivalent β-hairpins form a helical arrangement (Enemark and Joshua-Tor, 2006; Sanders et al., 2007). An intriguing question is whether additional subunits will continue to position their β-hairpins along the helical path of the minor groove during hexamer assembly around the dsDNA. Normally, duplex DNA in B form is characterized by ~10 bp per helical turn (360°). If six LTag subunits bind along six consecutive base pairs and then form a circle, it should distort the duplex DNA during the process and lead to DNA melting.

Second, the two β-hairpin H513 residues insert into a minor groove region with a minimum groove width of 4.5 Å, compared with the standard B-DNA groove width of 5.8 Å (Rohs et al., 2009). The immediate effect of such a narrowed minor groove is that the electronegativity is increased by ~2 kT/e, forming a binding site that attracts the H513 residues more effectively than other regions of the DNA. Our MC simulation results show that the minor groove width of the unbound ori DNA has a minimum groove width of 4.6 Å, matching that in the cocrystal structure. This result indicates that the narrower minor groove is not induced by protein binding but is an inherent structural feature of the particular origin sequence. Thus, we conclude that it is the particular ori DNA structure that is recognized by the β-hairpin H513 residues. This result is consistent with a previous study that showed the importance of the ori DNA sequence by demonstrating that mutation of this DNA region knocks out viral DNA replication (Deb et al., 1987). In those experiments, an A/T base pair was replaced by a G/C base pair, which removed the TpT dinucleotide that supports the narrowing of the minor groove through negative propeller twisting. In addition, histidine residues can better intrude into the minor groove in the vicinity of A/T base pairs because the bulky guanine amino group is absent.

Third, it is known that LTag can translocate along dsDNA and single-stranded DNA (ssDNA), but it is unknown how the direction is determined when LTag hexamers translocate along dsDNA without unwinding. Previous work showed that the β-hairpin is the major lever that shifts back and forth along the central channel upon ATP binding and hydrolysis (Gai et al., 2004), which should be coupled to translocating ssDNA and dsDNA through the central channel. In the complex with ori DNA, the two H513 residues on the β-hairpin insert into the minor groove to interact with both DNA strands. Interestingly enough, the adjacent K512 residues of both subunits interact with only one of the two strands (the strand in yellow in Figure 3D). This tracking on only one strand by two (or possibly more) β-hairpin K512 residues at a given time when binding to dsDNA may provide the directionality needed for translocation not only on ssDNA but also on dsDNA. Consistent with this hypothesis, mutation of K512 completely abolished translocation on dsDNA and ssDNA (Figure S5).

DNA Shape Readout through Histidine Contacts

The electrostatic attraction of histidine to DNA requires its protonation, and histidine residues are frequently protonated in the highly charged environment of DNA (Joshi et al., 2007). In the case of the LTag dimer, the presence of a hydrogen bond is indicative of one histidine being protonated to provide a hydrogen-bond donor, while the second histidine is unprotonated. Thus, the total charge of the histidine pair is likely +1. The H513 pair establishes two crucial interactions within the ternary complex: (1) a hydrogen bond within the LTag dimer and (2) electrostatic attraction of the His513 pair into the minor groove. This readout mechanism of sequence-dependent DNA shape may explain the initial anchoring of the subunits to the DNA, which may be of key importance for LTag assembly.

Histidine was previously observed to bind a narrow minor groove region of the binding site of the Hox protein Sex combs reduced (Scr) but in tandem with an arginine residue (Joshi et al., 2007), forming a hydrogen bond between its guanidinium

group with the histidine. Similar to the situation in the LTag dimer, this hydrogen bond can only form when the histidine is not protonated, assigning the Arg3-His-12 pair a total charge of +1. This conclusion is supported by the fact that a His-12 mutant of Scr had only a minor effect on binding and in vivo activity compared with a mutant of the charged Arg3 (Joshi et al., 2007). In LTag dimer binding to DNA, however, histidine takes on a key role in protein-DNA recognition without the presence of an arginine side chain. The observation that histidine of LTag on its own uses the mechanism of DNA shape readout to bind DNA is also present in other biological systems based on our analysis of cocrystal structures of the IFN-β enhanceosome (Escalante et al., 2007; Panne et al., 2004, 2007; Figure S4). We conclude that histidine in general uses a readout mechanism for protein-dsDNA interactions in a manner similar to that previously described for aginine (Rohs et al., 2009).

The critical role of histidine in the interaction of LTag with DNA is also apparent from its high sequence conservation, as the H513 residue is highly conserved among all LTag proteins in polyomaviruses and within the distantly related E1 helicase of papillomaviruses (see Figure S2C). Published results show that mutation of H513 on LTag to alanine affects DNA unwinding and origin melting (Kumar et al., 2007; Shen et al., 2005). Additionally, mutations of the equivalent H513 to alanine in E1 helicase disrupted ori DNA binding and unwinding (Liu et al., 2007; Schuck and Stenlund, 2005). The structural data presented here, together with these mutational results in the literature, support the observation that the highly conserved histidine residue plays a crucial role in the interaction of LTag with ori DNA to initiate DNA replication.

Interactions between OBD and PEN Sequences for Assembly

It has been a long-standing question how the four PEN repeats (PEN1–PEN4; Figure 1B) of the full ori DNA are utilized in the assembly of a double hexamer to initiate replication (Joo et al., 1998; Sreekumar et al., 2000). Previous studies revealed that each half origin (i.e., the EP-ori carrying PEN1 and PEN2, or the AT-ori with PEN3 and PEN4) can support an efficient hexamer assembly. Here, our dimer-ori DNA structure reveals that only PEN1, and not PEN2, is bound by OBD1. The linker connecting the OBD and Zn domain of subunit 2 in the dimer is not long enough for OBD2 to reach PEN2. However, OBD2 inverts by 180° to bind to a hidden site, yielding a well-positioned OBD2, because the B-factors of OBD2 are comparable to those of OBD1.

The hidden-site sequence of the ori DNA ($G_{16}C_{17}/C_{16'}G_{17'}$; Figure S6) bound by OBD2 revealed in the structure here is consistent with an earlier mutational study by Deb et al. (1986), which showed that $G_{16}C_{17}$ and $C_{16'}G_{17'}$ are important for viral DNA replication. Their data showed that the $G_{16}/C_{16'}$ base pair is more critical than the $C_{17}/G_{17'}$ base pair, which is consistent with the more extensive interactions of OBD2 with the $G_{16}/C_{16'}$ compared with the $C_{17}/G_{17'}$ base pair observed in the structure (see the more extensive hydrogen bonds with the $G_{16}/C_{16'}$ base pair in Figure S6B). Furthermore, the observation that PEN1 alone is bound by OBD is also consistent with prior biochemical studies showing that the assembly of a double-hexamer requires only PEN1 and PEN3 on a full origin, and PEN2 and PEN4 are dispensable (Joo et al., 1998; Sreekumar et al., 2000). However, because all four PEN sequences are required for viral infectivity (Joo et al., 1998; Sreekumar et al., 2000), further investigations will be required to elucidate the exact role of PEN2 and PEN4 in viral DNA replication in vivo.

LTag Residues Interact with Both Protein and DNA during Assembly

At the initial stage of LTag assembly on ori DNA, specific protein-DNA interactions likely are essential for recruiting the first LTag monomer or dimer to the origin. In this dimer structure, several dimer interface residues (such as K271, K512, R456, and N515) are present either on the exposed surfaces of the dimer or in the LTag-DNA interface. Thus, these residues have a dual role in forming contacts at either the protein-protein interface (Figure S1C, middle panel) or the protein-DNA interface (Figure S1C, left panel). The DNA binding by these residues is transient, and contacts will be broken so that these residues can bind to newly arrived protein subunits during assembly (Figure S1C, right panel). Therefore, it is likely that such dual-binding residues are very important for the assembly of LTag around the origin because they can provide stability for the first monomer-ori DNA or dimer-ori DNA intermediate complex and facilitate the subsequent recruitment of new subunits for hexamer assembly (see the model in Figure S1C).

In summary, we have determined the crystal structure of a dimeric LTag in complex with the EP-ori DNA. The structure reveals information about how the OBD in the context of a multidomain LTag polypeptide recognizes the sequences around PEN1 and PEN2 on the EP half origin. We have also illustrated an origin recognition mechanism by the AAA+ β-hairpin, i.e., minor-groove shape readout through histidines that recognize the minor-groove geometry of the EP region. Because the β-hairpin is critical for origin binding and melting (Kumar et al., 2007; Shen et al., 2005), this shape-readout mechanism is critical for the initial assembly of LTag and the melting of origin. The LTag dimer in complex with the EP-ori DNA likely represents a snapshot of an early-stage assembly intermediate for the hexamer/double-hexamer initiator complex, a process that is common to some replicons in eukaryotic cells.

EXPERIMENTAL PROCEDURES

Cloning, Mutations, and Protein Purification

The LTag131–627 construct was cloned in pGEX-6P-1 vector as an N-terminal glutathione S-transferase (GST) fusion. The LTag protein was expressed in *Escherichia coli* BL21 cells at 16–18°C. The protein was purified as follows: After cells were lysed, the protein was purified by glutathione affinity column in a buffer containing 50 mM Tris-HCl (pH 8.0), 250 mM NaCl, and 1 mM dithiothreitol (DTT). After the GST was removed by thrombin protease treatment at 4°C for 16 hr, LTag was purified by Superdex-200 gel filtration chromatography. To generate monomeric mutants for formation of a stable dimer intermediate, V350E/P417D and L286Dt/R567E mutants were made, and mutations were confirmed by sequencing the entire LTag insert.

Cocrystallization and Data Collection

To obtain EF-half ori DNA for cocrystallization, oligonucleotides (ordered from Operon) 5′-actacttctggaatagctcagaggccgaggcg-3′ and 5′-cgcctcggcctct gagctattccagaagtagt-3′ were purified using a Mono Q column. The two

Table 1. Data Collection and Model Refinement Statistics of the LTag-EP Origin Structure

Cell Dimensions (Space Group $P2_1$)	
a, *b*, *c* (Å)	73.63, 128.31, 166.12
α, β, γ (°)	90, 89.911, 90
Resolution (Å)	50–2.8 (2.9–2.8)
Observations	172,018
R_{merge}	8.3 (59.2)
$I/\sigma I$	12.1 (1.2)
Completeness (%)	93.2 (88.4)
Refinement	
Resolution (Å)	50.0–2.8
No. of reflections	65,350
R_{work} /R_{free}	22.69/25.49
B factor (averaged)	
Protein	58.554
DNA	55.396
Root-mean-square deviations	
Bond lengths (Å)	0.008655
Bond angles (°)	1.20198

Highest-resolution shell values are shown in parentheses.

purified oligonucleotides were annealed by heating to 94°C followed by slow cooling to 21°C. The annealed EP-half ori dsDNA was purified by Superdex 75 chromatography. The purified EP-half ori dsDNA and LTag131–627 (10 mg/ml) were mixed in a 1.25:1 molar ratio. Crystals were grown at 4°C by the sitting drop vapor diffusion method using a solution containing 100 mM bis-Tris (pH 6.75) and 20% (v/v) PEG3350. Full-size crystals were harvested after 2 weeks and flash-frozen with liquid nitrogen in the crystallization buffer containing glycerol. Diffraction data were collected and data sets were processed using HKL2000. The crystallization statistics are summarized in Table 1.

Structure Determination and Refinement

The initial phases of both dimeric crystal forms were obtained by molecular replacement, and two copies of each LTag helicase domain (Protein Data Bank [PDB] ID: 1SVM) and the OBD (PDB ID: 2ITL) were determined by PHASER. Very good DNA electron density could already be seen at this stage, and 2-fold NCS averaging improved the density even more, which allowed us to fit the DNA sequences into the density. The registry of the DNA base pairs was determined based on the shape of the electron density corresponding to the DNA bases. Iterative model rebuilding with O and refinement with CNS improved the phase even further. The final refinement statistics and geometry as defined by Procheck were very good and are summarized in Table 1.

Analysis and Prediction of Ori DNA Structure

The crystal structure of bound EP-half ori DNA was analyzed with CURVES (Lavery and Sklenar, 1989). The electrostatic potential was calculated by solving the nonlinear Poisson-Boltzmann equation at physiologic ionic strength of 0.145 M with DelPhi (Honig and Nicholls, 1995) in reference points at the center of the minor groove at midpoints between the O4′ atoms of the closest sugar moieties across the minor groove, using parameters described elsewhere (Rohs et al., 2009). The conformation of unbound DNA was predicted based on all-atom MC simulations using the AMBER force field, an implicit aqueous solvent model, and explicit sodium counterions (Rohs et al., 2005). The structure prediction approach uses an ideal B-form DNA with the origin sequence and identical conformations of all dinucleotides as the starting configuration. The predicted conformation was generated based on three independent 2 million MC cycle simulations, the initial 500,000 MC cycles of which were considered as the equilibration period. The simulation protocol was identical to the one described elsewhere (Joshi et al., 2007; Rohs et al., 2005).

ATPase Assay and Helicase Assay

We performed the ATPase assay for LTag WT or mutants by detecting the phosphate generated by ATP hydrolysis as described previously (Greenleaf et al., 2008). For the helicase DNA substrate, Y fork-shaped DNA with 44 nt ssDNA tails and a 44 nt duplex was made by annealing two oligonucleotides of 5′-$(dT)_{44}$GCTCGTGCAGACGTCGAGGTGAGGACGAGC TCCTCGTGACCACG and 5′-CGTGGTCACGAGGAGCTCGTC CTCACCTCG ACGTCTGCACGAGC$(dT)_{44}$. The helicase assay was performed as previously described (Greenleaf et al., 2008). Briefly, ~10 fmol of [γ-32P]-ATP-labeled substrate DNA was incubated with 300 ng LTag in helicase buffer containing 20 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 5 mM ATP, 1 mM DTT, and 0.1 mg/ml BSA for 45 min at 37°C. The reaction was analyzed on 12% native polyacrylamide gel, and the amount of radioactively labeled oligonucleotide was determined by autoradiography.

$KMnO_4$ Reactivity Assay

A $KMnO_4$ reactivity assay to test DNA melting activity was performed as follows: A 92 bp dsDNA substrate that contained EP-half ori was generated by annealing two ssDNA. Only the lower strand was 5′-end labeled with [γ-32P] ATP. Then 10–15 fmol DNA was incubated with 100–800 ng LTag protein in buffer containing 20 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 5 mM ATP, 1 mM DTT, and 0.1 mg/ml BSA for 30 min at 37°C. Then $KMnO_4$ was added to a final concentration of 6 mM and reactions were incubated at 37°C for 2 min. To stop the oxidation reactions, a stop solution containing 160 mM β-mercaptoethanol, 0.3% SDS to 0.3%, and 10 mM EDTA was added to the reaction mixture. Modified DNA substrates were deproteinized first by digestion with proteinase K (20 μg/ml, 60 min at 37°C) and then binding to QIAGEN QIAEX II beads. Piperidine was added to 20% (v/v) on dried beads to cleave DNA substrates at modified nucleotides (30 min at 90°C).

ACCESSION NUMBERS

The atomic coordinates of the cocrystal structure have been deposited in the Protein Data Bank under ID code 4GDF.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.03.002.

LICENSING INFORMATION

ACKNOWLEDGMENTS

We thank the staffs of the USC NanoBiophysics core, Argonne National Laboratory synchrotron beamlines 23ID and 19ID, and Berkeley's ALS beamlines 5.0.2, 8.2.1, and 8.3.1 for assistance with data collection. This work was supported by NIH grants GM080338 and AI055926 (to X.S.C.) and American Cancer Society grant IRG-58-007-51 (to R.R.). R.R. is an Alfred P. Sloan Research Fellow. Y.P.C. did the crystallography work, structural analysis, and initial writing of the manuscript; M.X. obtained the stable dimer intermediate of LTag and performed melting and related biochemical studies; A.C.D.M. performed the computational analyses; X.J.Y. made LTag K512 mutations and performed functional assays; R.R. supervised the computational work and wrote parts of the manuscript; and X.S.C. coordinated the project and wrote the manuscript.

Received: October 6, 2012
Revised: January 10, 2013
Accepted: March 1, 2013
Published: March 28, 2013

REFERENCES

Bochkareva, E., Martynowski, D., Seitova, A., and Bochkarev, A. (2006). Structure of the origin-binding domain of simian virus 40 large T antigen bound to DNA. EMBO J. *25*, 5961–5969.

Borowiec, J.A., and Hurwitz, J. (1988). Localized melting and structural changes in the SV40 origin of replication induced by T-antigen. EMBO J. *7*, 3149–3158.

Borowiec, J.A., Dean, F.B., Bullock, P.A., and Hurwitz, J. (1990). Binding and unwinding—how T antigen engages the SV40 origin of DNA replication. Cell *60*, 181–184.

Deb, S., DeLucia, A.L., Baur, C.P., Koff, A., and Tegtmeyer, P. (1986). Domain structure of the simian virus 40 core origin of replication. Mol. Cell. Biol. *6*, 1663–1670.

Deb, S., Tsui, S., Koff, A., DeLucia, A.L., Parsons, R., and Tegtmeyer, P. (1987). The T-antigen-binding domain of the simian virus 40 core origin of replication. J. Virol. *61*, 2143–2149.

Dueber, E.L., Corn, J.E., Bell, S.D., and Berger, J.M. (2007). Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. Science *317*, 1210–1213.

Enemark, E.J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. Nature *442*, 270–275.

Escalante, C.R., Nistal-Villán, E., Shen, L., García-Sastre, A., and Aggarwal, A.K. (2007). Structure of IRF-3 bound to the PRDIII-I regulatory element of the human interferon-beta enhancer. Mol. Cell *26*, 703–716.

Fanning, E., and Zhao, K. (2009). SV40 DNA replication: from the A gene to a nanomachine. Virology *384*, 352–359.

Gai, D., Zhao, R., Li, D., Finkielstein, C.V., and Chen, X.S. (2004). Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. Cell *119*, 47–60.

Gaudier, M., Schuwirth, B.S., Westcott, S.L., and Wigley, D.B. (2007). Structural basis of DNA replication origin recognition by an ORC protein. Science *317*, 1213–1216.

Greenleaf, W.B., Shen, J., Gai, D., and Chen, X.S. (2008). Systematic study of the functions for the residues around the nucleotide pocket in simian virus 40 AAA+ hexameric helicase. J. Virol. *82*, 6017–6023.

Honig, B., and Nicholls, A. (1995). Classical electrostatics in biology and chemistry. Science *268*, 1144–1149.

Joo, W.S., Kim, H.Y., Purviance, J.D., Sreekumar, K.R., and Bullock, P.A. (1998). Assembly of T-antigen double hexamers on the simian virus 40 core origin requires only a subset of the available binding sites. Mol. Cell. Biol. *18*, 2677–2687.

Joshi, R., Passner, J.M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M.A., Jacob, V., Aggarwal, A.K., Honig, B., and Mann, R.S. (2007). Functional specificity of a Hox protein mediated by the recognition of minor groove structure. Cell *131*, 530–543.

Kumar, A., Meinke, G., Reese, D.K., Moine, S., Phelan, P.J., Fradet-Turcotte, A., Archambault, J., Bohm, A., and Bullock, P.A. (2007). 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. *81*, 4808–4818.

Lavery, R., and Sklenar, H. (1989). Defining the structure of irregular nucleic acids: conventions and principles. J. Biomol. Struct. Dyn. *6*, 655–667.

Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E., Jochimiak, A., Szakonyi, G., and Chen, X.S. (2003). Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. Nature *423*, 512–518.

Liu, X., Schuck, S., and Stenlund, A. (2007). Adjacent residues in the E1 initiator beta-hairpin define different roles of the beta-hairpin in Ori melting, helicase loading, and helicase activity. Mol. Cell *25*, 825–837.

Mastrangelo, I.A., Hough, P.V., Wall, J.S., Dodson, M., Dean, F.B., and Hurwitz, J. (1989). ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. Nature *338*, 658–662.

Meinke, G., Phelan, P., Moine, S., Bochkareva, E., Bochkarev, A., Bullock, P.A., and Bohm, A. (2007). The crystal structure of the SV40 T-antigen origin binding domain in complex with DNA. PLoS Biol. *5*, e23.

Méndez, J., and Stillman, B. (2003). Perpetuating the double helix: molecular machines at eukaryotic DNA replication origins. Bioessays *25*, 1158–1167.

Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M., and Zhurkin, V.B. (1998). DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. Proc. Natl. Acad. Sci. USA *95*, 11163–11168.

Panne, D., Maniatis, T., and Harrison, S.C. (2004). Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-beta enhancer. EMBO J. *23*, 4384–4393.

Panne, D., Maniatis, T., and Harrison, S.C. (2007). An atomic model of the interferon-beta enhanceosome. Cell *129*, 1111–1123.

Petrey, D., and Honig, B. (2003). GRASP2: visualization, surface properties, and electrostatics of macromolecular structures and sequences. Methods Enzymol. *374*, 492–509.

Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A., and Honig, B. (2002). Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J. Comput. Chem. *23*, 128–137.

Rohs, R., Sklenar, H., and Shakked, Z. (2005). Structural and energetic origins of sequence-specific DNA bending: Monte Carlo simulations of papillomavirus E2-DNA binding sites. Structure *13*, 1499–1509.

Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. Nature *461*, 1248–1253.

Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B., and Mann, R.S. (2010). Origins of specificity in protein-DNA recognition. Annu. Rev. Biochem. *79*, 233–269.

Sanders, C.M., Kovalevskiy, O.V., Sizov, D., Lebedev, A.A., Isupov, M.N., and Antson, A.A. (2007). Papillomavirus E1 helicase assembly maintains an asymmetric state in the absence of DNA and nucleotide cofactors. Nucleic Acids Res. *35*, 6451–6457.

Schuck, S., and Stenlund, A. (2005). Assembly of a double hexameric helicase. Mol. Cell *20*, 377–389.

Shen, J., Gai, D., Patrick, A., Greenleaf, W.B., and Chen, X.S. (2005). The roles of the residues on the channel beta-hairpin and loop structures of simian virus 40 hexameric helicase. Proc. Natl. Acad. Sci. USA *102*, 11248–11253.

Simmons, D.T. (2000). SV40 large T antigen functions in DNA replication and transformation. Adv. Virus Res. *55*, 75–134.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structure and mechanism of helicases and nucleic acid translocases. Annu. Rev. Biochem. *76*, 23–50.

Sreekumar, K.R., Prack, A.E., Winters, D.R., Barbaro, B.A., and Bullock, P.A. (2000). The simian virus 40 core origin contains two separate sequence modules that support T-antigen double-hexamer assembly. J. Virol. *74*, 8589–8600.

Valle, M., Chen, X.S., Donate, L.E., Fanning, E., and Carazo, J.M. (2006). Structural basis for the cooperative assembly of large T antigen on the origin of replication. J. Mol. Biol. *357*, 1295–1305.