Replication of Herpes Simplex Virus DNA*

I. R. Lehman§§ and Paul E. Boehmer¶

From the Department of Biochemistry, Beckman Center, Stanford University School of Medicine, Stanford, California 94305-5307 and Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, Miami, Florida 33101

Herpes simplex virus type 1 (HSV-1) is a member of the herpes group of viruses, the Herpesviridae, which includes the important human pathogens herpes simplex virus type 2, cytomegalovirus, varicella zoster virus, Epstein-Barr virus, human herpesviruses 6 and 7, and Kaposi's associated herpesvirus, human herpesvirus 8 (reviewed in Ref. 1). Of these, HSV-1 has been the most extensively studied. The HSV-1 genome is a 153-kilobase linear duplex with 75 open reading frames. It consists of two unique regions, UL (unique long), comprising 82% and the US (unique short) comprising 18%, flanked by the repeated regions a, b, and c (Fig. 1). As a consequence of recombination through the α sequences, the HSV-1 genome actually exists as an equimolar mixture of four isomeric forms generated by inversion of the UL and US sequences relative to each other.

An interesting but poorly understood aspect of the HSV-1 life cycle, in fact, the life cycle of all the herpesviruses, is the ability to enter latency, a state in which no infectious progeny are produced and only limited transcription of viral genes occurs. Upon presentation of the appropriate stimulus, the latently infected cells enter a lytic phase with the resultant production of infectious viral particles (1).

In the lytic phase, following penetration of the nucleus, a temporally regulated program of viral gene expression mediated by RNA polymerase II begins. A virion protein, VP16, acts in trans together with the cellular octamer DNA-binding protein, Oct-1, to induce the α or immediate early genes, the first set of HSV-1 genes to be expressed. These serve as transactivators of β or early gene expression. Their synthesis peaks at 2–4 h, postinfection. Functional α proteins are required for the synthesis of all the virally encoded proteins. Synthesis of the β class of proteins reaches peak rates 5–7 h and is detectable as early as 3 h postinfection. The β proteins include the enzymes that are required for replication of the viral genome: a DNA polymerase, a single-strand DNA-binding protein, a primosome or helicase-primase, an origin-binding protein, and a set of enzymes involved in DNA repair and in deoxyribonucleotide metabolism. Homologues of these proteins are found in virtually all herpesviruses (1, 2).

Viral DNA synthesis begins shortly after the appearance of the β proteins and is detectable as early as 3 h and continues up to 15 h postinfection, with the bulk of the DNA synthesized during the latter part of the phase. The temporal program of viral gene expression ends with the appearance of the γ or late proteins, which constitute the structural proteins of the virus (1).

Structure of Replicating DNA

As early as 0.5 h postinfection and in the absence of protein synthesis, viral DNA accumulates in the nucleus and adopts an "endless" configuration whose structure is consistent with circularization (1). Although the mechanism of genome circularization has not been established, it may involve a recombination event through the terminal repeat sequences. One clear advantage of this intermediate is that it eliminates the need for a specialized mechanism to replicate the termini.

Although there is no firm proof for the existence of theta replication intermediates that initiate at one or more of the HSV-1 origins in vivo, genome circularization is a prerequisite for viral replication, and replication requires either oriS or oriL, and the action of a viral protein that possesses all the properties of an initiator protein (UL9 protein). These observations strongly suggest that HSV-1 DNA replication initiates by a theta mechanism and make it unlikely that other mechanisms, such as inter- or intramolecular recombination or transcription, all of which could potentially occur on linear DNA templates, are responsible for the initiation of HSV-1 DNA replication. At some point and by an as yet undefined mechanism, theta replication switches to the rolling circle mode, the predominant mode of HSV-1 DNA replication.

In the rolling circle mode, DNA replication generates long head-to-tail concatamers similar to the DNA replication products observed during the theta phase of bacteriophage λ replication (3). At late times during viral DNA replication, the products consist of branched networks of concatamers that may arise from recombination and invading DNA replication forks.

Origins of DNA Replication

The HSV-1 genome contains three cis-acting elements, which function as origins of DNA replication: one copy of oriL, located between the diverging UL29 and UL30 genes within UL1, and two copies of oriS, located within the c sequences (Fig. 2).

Both oriS and oriL contain large palindromes of 45 and 144 bp, respectively, that center around an A + T region of 18 and 20 bp, respectively. Flanking the A + T region are inverted repeats that are binding sites for the HSV-1-encoded origin-binding protein (UL9 protein).

It is unclear why the HSV-1 genome contains multiple origins that appear identical in every respect other than size and location. The existence of two copies of oriS may be rationalized by its location in a repeated c region of the genome. However, neither copy of oriL is essential because deletion of one or both copies of oriS has no effect on viral DNA replication in cultured cells (1).

HSV-1 Gene Products Essential for Origin-specific DNA Replication

The development of a transient DNA replication assay in which origin-containing plasmids are replicated by transfected HSV-1 sequences that supply trans-acting factors led to major advances in the identification of the essential HSV-1 DNA replication genes (4). This analysis permitted identification of seven HSV-1 genes all located within the UL segment of the genome that are necessary and sufficient for origin-specific DNA replication: UL5, UL8, UL9, UL29, UL30, UL42, and UL52 (Fig. 1) (5, 6). These assignments were subsequently confirmed by demonstrating that mutations in these genes abrogate viral replication (7, 8). Although the function of the UL29 (single-strand DNA-binding protein) and UL30 (DNA polymerase) gene products had already been identified, the remaining five genes encoded proteins of unknown function. These genes were subsequently correlated with proteins identified by biochemical analyses. The functions of these gene products are summarized in Table I, which also depicts the conservation of DNA replication functions in a variety of organisms/systems.

Origin-binding Protein (UL9 Protein)—A search for proteins that bind to an HSV-1 origin led to the identification of an HSV-1-induced factor that could recognize sequences within oriS (9). Further purification of this sequence-specific DNA binding activity led to the identification of an 83-kDa protein that bound to two inverted repeats that flank the A + T region, within oriS, designated Boxes I and II (Fig. 2) (10). In oriS, the two Box I sites are flanked by homologous sequences designated Box III, which differ...
The binding of UL9 protein to oriS appears to be a homodimer, only one protein can bind to its recognition sequence as a homodimer, only one negative effect on viral replication, presumably because of its ability to occupy the origin nonproductively (16, 17). Although the UL9 protein can bind to its recognition sequence as a homodimer, only one of the two monomers contacts the DNA (18). The dimerization of UL9 protein is mediated through the N-terminal part of the protein, presumably through a leucine zipper motif encompassing residues 150–171 (19, 20). The binding of UL9 protein to oriS appears to be enhanced by nucleoside triphosphates (21).

In addition to its sequence-specific DNA binding activity, the UL9 protein possesses DNA-stimulated nick-end DNA helicase and nickspecific 3'→5' DNA helicase activities (22–25). ICP8, the HSV-1 single-strand DNA-binding protein, interacts with the extreme C terminus of the UL9 protein, greatly stimulating its DNA-dependent ATPase and DNA helicase activities (25). In the absence of ICP8 the UL9 protein is non-processive. ICP8 appears to stimulate the UL9 protein by preventing its dissociation from the DNA (25, 26).

The UL9 protein, together with ICP8, can unwind specifically from Box I by only 1 residue but act as low affinity recognition sites for the UL9 protein. In oriS there is only one copy of Box I, 3' to Box I. Subsequently, it was demonstrated that the origin-binding protein is the product of the UL9 gene (11). The UL9 gene encodes an 851-amino acid polypeptide (UL9 protein) with a calculated mass of 89,246 Da (12). Amino acid sequence analysis of the UL9 protein led to the identification of conserved ATP binding and DNA helicase motifs (13, 14) that have been shown to be essential for viral replication (15, 16). The sequence-specific DNA binding activity of the UL9 protein resides in the C-terminal 317 amino acids (16). Expression of the 37-kDa C-terminal DNA-binding domain has a dominant negative effect on viral replication, presumably because of its ability to occupy the origin nonproductively (16, 17). Although the UL9 protein can bind to its recognition sequence as a homodimer, only one of the two monomers contacts the DNA (18). The dimerization of UL9 protein is mediated through the N-terminal part of the protein, presumably through a leucine zipper motif encompassing residues 150–171 (19, 20). The binding of UL9 protein to oriS appears to be enhanced by nucleoside triphosphates (21).

In addition to its sequence-specific DNA binding activity, the UL9 protein possesses DNA-stimulated nick-end DNA helicase and nickspecific 3'→5' DNA helicase activities (22–25). ICP8, the HSV-1 single-strand DNA-binding protein, interacts with the extreme C terminus of the UL9 protein, greatly stimulating its DNA-dependent ATPase and DNA helicase activities (25). In the absence of ICP8 the UL9 protein is non-processive. ICP8 appears to stimulate the UL9 protein by preventing its dissociation from the DNA (25, 26).

The UL9 protein, together with ICP8, can unwind specifically from Box I by only 1 residue but act as low affinity recognition sites for the UL9 protein. In oriS there is only one copy of Box I, 3' to Box I. Subsequently, it was demonstrated that the origin-binding protein is the product of the UL9 gene (11). The UL9 gene encodes an 851-amino acid polypeptide (UL9 protein) with a calculated mass of 89,246 Da (12). Amino acid sequence analysis of the UL9 protein led to the identification of conserved ATP binding and DNA helicase motifs (13, 14) that have been shown to be essential for viral replication (15, 16). The sequence-specific DNA binding activity of the UL9 protein resides in the C-terminal 317 amino acids (16). Expression of the 37-kDa C-terminal DNA-binding domain has a dominant negative effect on viral replication, presumably because of its ability to occupy the origin nonproductively (16, 17). Although the UL9 protein can bind to its recognition sequence as a homodimer, only one of the two monomers contacts the DNA (18). The dimerization of UL9 protein is mediated through the N-terminal part of the protein, presumably through a leucine zipper motif encompassing residues 150–171 (19, 20). The binding of UL9 protein to oriS appears to be enhanced by nucleoside triphosphates (21).

in the range of one ICP8 to 12–22 nucleotides (34–36), values that are in agreement with a site size of 15–18 nucleotides estimated by electron microscopy (33).

**Primosome (DNA Helicase-Primase)**—The HSV-1-encoded primosome consists of three subunits with molecular masses of 98,710, 79,291, and 114,416 Da, the products of the UL5, UL8, and UL52 genes, respectively.

The holoenzyme consists of a 1:1:1 association of the UL5, UL8, and UL52 gene products, with a native molecular mass of ~270 kDa (37, 38). A subassembly that consists of the UL5 and UL52 subunits retains DNA-dependent ATPase, helicase, and primase activities and therefore constitutes the core enzyme (39, 40). In contrast, the UL8 protein lacks detectable enzymatic or DNA binding activities (40, 41). Under the appropriate conditions of Mg2+ concentration and ionic strength the helicase activity of the helicase-primase can unwind 60–65 bps, a value consistent with the rate of 50 bps estimated for the rate of fork movement in vivo during replication of pseudorabies virus, another herpesvirus (42).

Residues 610–636 of the UL52 protein contain a proposed divalent metal binding motif that is conserved in DNA polymerases and primases (43, 44). Site-directed mutagenesis of this motif inactivates the primase activity of the mutant holoenzyme in vitro and abolishes origin-specific DNA replication in vivo, indicating that it comprises the primase active site. The products of primase action are oligoribonucleotides that are 6–13 bases in length (45, 46). With single-stranded dX74 DNA as template, the primase exhibits a sequence specificity. The preferred template sequence was mapped to 3'–AGCCCTCCCA, directing the synthesis of a 10-mer oligoribonucleotide, starting at the underlined C residue (48).

The UL8 protein stimulates primer synthesis by the UL5/52 sub-assembly (47–49). Thus, although UL8 protein has no effect on the K_m of DNA (15 μM nucleotides) of the subassembly, it increases the rate of primer synthesis approximately 3-fold (47). The UL8 protein is also required for optimal DNA helicase and DNA-dependent ATPase activities of the enzyme in the presence of ICP8-coated DNA templates (48, 49). Surface plasmon resonance measurements did, in fact, demonstrate an interaction between ICP8 and the UL5/52 subassembly or the UL8 protein alone (49).

**DNA Polymerase**—The HSV-1 DNA polymerase has been intensively studied, both as a model eukaryotic DNA polymerase and as a target for antiviral drugs. It is a 1235-amino-acid, 136,413-Da protein that is encoded by the UL30 gene (29, 51–53). The UL30 protein exhibits significant sequence similarity to other viral and cellular DNA polymerases including human DNA polymerase α-primase, Saccharomyces cerevisiae DNA polymerase δ, E. coli DNA polymerase I, and the bacteriophage T4 DNA polymerase (54, 55).

The HSV-1 DNA polymerase purified from virus-infected cells exists as a heterodimer in which the UL30 protein is associated with a 65-kDa protein (56–58). The 65-kDa protein is the product of the UL42 gene which encodes a 488-amino-acid, 51,154-Da phosphoprotein that possesses double-strand DNA binding activity (6, 59). The interaction between the UL30 and UL42 proteins occurs with an association constant of 1 × 10^9 M^-1 (60) and has been shown to be essential for viral replication (61).

The HSV-1 DNA polymerase possesses 3'→5' exo activity that is intrinsic to the UL30 subunit (62, 63). The exo activity is active under conditions that are optimal for deoxyribonucleotide polymerization and can act on a variety of DNA substrates including gapped DNA, DNA hairpin structures, and single-stranded oligodeoxyribonucleotides (10). In the absence of dNTPs, the 3'→5' exo activity is twice as active at removing 3'-unpaired deoxyribonucleotides as paired deoxyribonucleotides. In the presence of all four dNTPs, there is no detectable 3'→5' exo activity on paired deoxyribonucleotides, presumably because of continuous polymerization.

In contrast unpaired deoxyribonucleotides are completely removed (63). Consequently, the 3'→5' exo activity serves as a proofreading activity to ensure high fidelity of DNA replication. In addition to its 3'→5' exo activity, the HSV-1 DNA polymerase has a RNase H activity that seemingly serves to remove RNA primers during the processing of Okazaki fragments (58).

**FIG. 1. Structure of the herpes simplex virus type 1 genome.** A diagrammatic representation of the HSV-1 genome is shown. The positions of the a, b, and c repeats within the terminal repeats (TR1, TR2, and TR3) and internal repeats (IR1, and IR2) and the positions of the DNA replication origins (oriA, and oriS) are indicated by the arrows. The enlarged area shows the composition of a unit length a sequence, consisting of direct repeats (DR) 1, 2, and 3 and unique (U) domains b and c, t, terminal, n, variable number of copies, 1, inverted orientation, L, long, S, short. This drawing is not to scale; see text for details. (Reprinted with permission from Ref. 1.)
pol, polymerase.

ORC, origin recognition complex; MCM, maintenance of minichromosome; PCNA, proliferating cell nuclear antigen; RF, replication factor; FEN, flap endonuclease; RP, replication protein.

polymerization. However, these DNA polymerase clamps lack DNA binding activity, and their assembly on the DNA is mediated by another set of proteins, the clamp loaders. In contrast, the HSV-1 DNA polymerase is tethered to the DNA by the double-strand DNA binding activity of the UL42 protein itself. This model is supported by the finding that stimulation of DNA polymerase activity requires the DNA binding activity of the UL42 protein (64, 65).

Numerous nucleotide analogues (9-(2-hydroxyethoxymethyl)guanosine 5'-triphosphate (acyclovir triphosphate), 9-(β-d-arabinofuranosyl)adenosine 5'-triphosphate (araATP), aphidicolin, (E)-5-(2-bromovinyl)-2'-deoxyuridine 5'-triphosphate) and pyrophosphate analogues (phosphonoacetate, phosphonoformate, oxalate) have been examined for their ability to inhibit viral replication in vivo and to inhibit selectively the HSV-1 DNA polymerase in vitro. This has led to the development of several antiviral drugs, notably gancyclovir and acyclovir, which are of significant therapeutic value (52).

In addition to ICP8, the DNA polymerase-UL42 complex, the UL9 protein, and the DNA helicase-primase, all of which are essential for HSV-1 DNA replication, the HSV-1 genome encodes a set of enzymes whose function is not required for its replication in cultured cells. They undoubtedly serve some function in the viral life cycle. These include an alkaline endo-exonuclease that exhibits 3'-exonuclease activity (UL12 protein) and may thus participate in the completion of lagging-strand DNA synthesis, a uracil DNA glycosylase, a deoxyuridine triphosphatase, a thymidine kinase, and a ribonucleotide reductase (2).

**Attempts to Reconstitute HSV-1 DNA Replication in Vitro**

As described previously, studies of HSV-1 DNA replication in vivo have demonstrated that the linear 153-kilobase pair genome circulates shortly after infection of susceptible host cells and then enters a rolling circle mode of DNA replication generating branched concatameric DNA, which is then cleaved and packaged as unit-length molecules. The existence of specific origins of replication oriS or oriL as well as a protein (the UL9 protein) with helicase activity that binds the origins with high affinity suggests that a theta mode of DNA replication precedes the rolling circle phase (Fig. 3). A bipartite mode of DNA replication of this kind has been observed for bacteriophage λ (3). Although, as noted above, rolling circle replication of the HSV-1 genome as well as of plasmids into which an HSV-1 origin has been inserted has been observed in vivo, there is as yet no compelling evidence in vivo for a theta mode of replication.

Numerous attempts have been made to observe origin (oriS or oriL)-dependent DNA replication in vitro. None has thus far been successful.

In contrast to the inability to reconstitute the theta model of

---

**Table I**

<table>
<thead>
<tr>
<th>E. coli/phage λ</th>
<th>Phage T4</th>
<th>Phage T7</th>
<th>HSV-1</th>
<th>S. cerevisiae</th>
<th>SV40/Homo sapiens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin-binding protein</td>
<td>DnaA/O protein</td>
<td>None</td>
<td>None</td>
<td>UL9 protein</td>
<td>ORC</td>
</tr>
<tr>
<td>Helicase-primase loading protein</td>
<td>DnaB</td>
<td>gp41</td>
<td>gp4</td>
<td>UL5/52 proteins</td>
<td>MCM proteins</td>
</tr>
<tr>
<td>Primase</td>
<td>DnaG</td>
<td>gp61</td>
<td>gp5</td>
<td>UL5/52 proteins</td>
<td>DNA pol α-primase</td>
</tr>
<tr>
<td>DNA polymerase</td>
<td>DNA pol III</td>
<td>gp48</td>
<td>gp5</td>
<td>DNA pol α-primase</td>
<td>DNA pol α-primase</td>
</tr>
<tr>
<td>DNA polymerase clamp</td>
<td>β subunit</td>
<td>gp45</td>
<td>E. coli thioredoxin</td>
<td>UL42 protein</td>
<td>PCNA</td>
</tr>
<tr>
<td>DNA polymerase clamp loader</td>
<td>γ complex</td>
<td>gp44/62</td>
<td>None</td>
<td>None</td>
<td>RF-C</td>
</tr>
<tr>
<td>Single-strand DNA-binding protein</td>
<td>SSB</td>
<td>gp32</td>
<td>gp2.5</td>
<td>ICP8</td>
<td>RP-A</td>
</tr>
<tr>
<td>RNase H</td>
<td>RNase H</td>
<td>gp33.2</td>
<td>gp6</td>
<td>UL30 protein</td>
<td>RNase H</td>
</tr>
<tr>
<td>5' → 3'-Exonuclease</td>
<td>DNA pol I</td>
<td>gp33.2</td>
<td>gp6</td>
<td>UL12 protein</td>
<td>FEN</td>
</tr>
</tbody>
</table>

---

**FIG. 2.** Structure of the herpes simplex virus type 1 origins of DNA replication, oriS and oriL. The DNA sequence of the minimal oriS is shown. Boxed bases indicate sequence differences between oriS and oriL. The arrows depict the relative orientation of the UL9 protein recognition sites, Boxes I, II, and III. The triangle indicates the center of symmetry: A + T, sequence. (Reprinted with permission from Ref. 2.)

**FIG. 3.** Model for the bipartite replication of the HSV-1 genome in which an initial transient phase of theta replication is followed by a rolling circle mode, the predominant mode of HSV-1 DNA replication.

The UL42 protein acts to increase the processivity of the UL30 DNA polymerase (36, 64, 65). It is therefore a functional homologue of the DNA polymerase processivity enhancing factors, which include the β subunit of E. coli DNA polymerase III holoenzyme, the eukaryotic proliferating cell nuclear antigen, and bacteriophage T4 gene 45 protein (28, 66) (see Table I). These proteins form a ring that encircles the DNA duplex and interact with their cognate DNA polymerase to prevent dissociation of the enzyme from the primer-template, thereby ensuring high processivity of deoxyribonucleotide
HSV-1 DNA replication, rolling circle replication promoted by extracts of HSV-1-infected cells has been achieved. The first demonstration of rolling circle replication made use of an artificial replication fork consisting of M13 single-stranded DNA to which a complementary oligonucleotide with an unpaired 5' single-stranded tail had been annealed. Rolling circle replication was observed both with crude extracts of HSV-1-infected cells and with a partially purified fraction containing the HSV-1 DNA polymerase and the UL5 and UL52 subunits of the DNA helicase-pimase. ICP8 stimulated the reaction but was not essential (67).

In a second approach, a high molecular mass complex (M, > 10^6) was isolated by gel filtration and ion-exchange chromatography from extracts of insect (Sf9) cells that had been multiply infected with baculovirus recombinant for the seven HSV-1 genes that are required for its replication. This complex, which consisted of the HSV-1 DNA polymerase, ICP8, and DNA helicase-pimase, could promote the rolling circle replication of circular plasmid templates to generate concatameric DNA molecules up to four times unit length as judged by gel electrophoresis and electron microscopy. Neither the UL9 protein nor an HSV-1 origin was required, and in fact, the presence of orf5 within the plasmid inhibited DNA replication in the presence, but not the absence, of the UL9 protein (68).

Whole-cell extracts of HSV-1-infected human cells (293 cells) can also promote the rolling circle replication of plasmid templates in a reaction that closely resembles that observed with the high molecular weight complex isolated from the baculovirus-infected Sf9 cells (67). Purification evidence suggests that a similar enzyme complex is present in the HSV-1-infected 293 cells. Although the rate of form movement in these reactions has not been accurately determined, it appeared to approach the in vitro rate cited above. However, the efficiency of the reaction is very low (1–2% of the template molecules replicated) for both the complex obtained from the baculovirus-infected Sf9 cells and the extract of HSV-1-infected 293 cells. The form of the plasmid DNA that serves as the template for rolling circle replication is not known, and the small proportion of the template molecules replicated may result from the use of an appropriate template that is generated by either the enzyme complex or the crude extract. Possibly a recombinational event analogous to that which appears to be involved in the switch from theta to rolling circle replication in bacteriophage A may be required (3).

In summary, HSV-1 is a large (153 kilobase pairs), double-stranded DNA virus whose genome contains multiple, redundant origins of replication. HSV-1 encodes a repertoire of proteins that, with the exception of a DNA ligase and topoisomerase, should be present in all cells with the HSV-1 DNA replication apparatus. A host DNA polymerase, DNA polymerase α-primase, may function during the theta phase (70) (Fig. 3). Finally, the functions of the HSV-1 DNA replication proteins resemble those found in other organisms, further indicating strong conservation in the mechanisms of DNA replication.

REFERENCES