

Nucleoprotein complex formed between herpes simplex virus UL9 protein and the origin of DNA replication: Inter- and intramolecular interactions

(electron microscopy)

SAMUEL D. RABKIN AND BRIAN HANLON

Program in Molecular Biology, Memorial Sloan-Kettering Cancer Center, New York, NY 10021

Communicated by Paul A. Marks, September 3, 1991 (received for review July 23, 1991)

ABSTRACT The UL9 gene of herpes simplex virus type 1 encodes an origin-binding protein. UL9 protein purified from baculovirus vector-infected insect cells forms a stable complex with DNA containing the herpes simplex virus origin of DNA replication, *ori_S*. Contained within *ori_S* are two UL9 protein-binding sites, I and II, bracketing an (A+T)-rich region. UL9 protein, visualized by electron microscopy, binds selectively at the site of the origin and covers ≈ 120 base pairs. Upon formation of the nucleoprotein complex, the apparent contour length of the DNA is shortened, suggesting that this amount of DNA is wrapped or condensed by the protein. A nucleoprotein complex of similar size and structure forms on an inactive origin deleted for binding site II. Multiple intermolecular interactions occur. In particular, UL9 nucleoprotein complexes interact in trans with other UL9 nucleoprotein complexes such that dimer DNA molecules are formed with a junction at the position of protein binding. The DNA molecules in these intermolecular complexes are aligned predominantly in a parallel orientation.

Herpes simplex virus type 1 (HSV-1) contains a linear double-stranded DNA genome of 153 kilobases (kb) that encodes at least 72 proteins (1). Seven of these viral genes are essential for HSV DNA replication in cultured cells (2, 3). One of these gene products binds specifically to the origin (origin-binding protein or UL9 protein) (4, 5). The UL9 protein also has DNA-dependent NTPase and DNA helicase activities (6).

The HSV genome contains three regions that are able to act independently as origins of DNA replication *in vivo* (7–9). One of these, *ori_L*, lies close to the center of the long unique segment (Fig. 1), between the divergent transcripts for ICP8 and pol. The other origin, *ori_S*, is located in the short inverted repeat segment and therefore is present twice in the genome (Fig. 1). *ori_S* has been mapped (12) to a sequence of about 75 base pairs (bp) (13) that includes a near-perfect palindromic sequence with an alternating AT sequence in the middle (Fig. 1). Both these features are also present in *ori_L*.

UL9 protein binds to two sites, I and II, that are the base of the palindromic sequence (Fig. 1) (5, 10) and contain inverted pentanucleotide repeats (11). Site I is a higher affinity binding site for UL9 protein than is site II (14). Mutagenesis data suggest that UL9 protein-binding site I is essential for transient origin activity (15, 16), whereas alterations to site II lead to varying degrees of activity ranging from undetectable to reduced (13, 17). The structure in the AT region is distorted when UL9 protein binds to supercoiled DNA, so that it becomes sensitive to modification by KMnO_4 (18).

A common set of events is involved in the initiation of DNA replication in many of the *in vitro* replication systems that have been studied to date. These include the binding of a sequence-specific initiator protein to the origin region, localized distortion or melting of the duplex DNA, and recruitment of helicase activity to further unwind this region (19, 20). DNA-bound proteins may interact to form specialized nucleoprotein structures that are visible in the electron microscope (21). Phage λ O protein binds as an octamer, condensing the DNA by forming a DNA-wound nucleoprotein complex ("O-some") visible by electron microscopy (22). Simian virus 40 large tumor (T) antigen, in the presence of ATP, binds as a double-hexamer structure, covering about 90 bp of DNA at the core origin (23, 24), but does not condense or wrap the DNA upon binding. Binding of T antigen leads to a structural alteration in the adjacent (A+T)-rich region (sensitive to KMnO_4 modification) and untwisting of the DNA (25).

To better understand the initiation process on the HSV chromosome, we have begun a structural analysis of UL9 protein binding to the origin of DNA replication. We show here that UL9 protein forms a specialized nucleoprotein complex at the origin of DNA replication that is visible in the electron microscope and is involved in multiple intermolecular interactions.

MATERIALS AND METHODS

Cells and Virus. Recombinant *Autographa californica* nuclear polyhedrosis virus (AcNPV) containing the complete UL9 open reading frame (AcNPV/UL9) (5) was obtained from M. Challberg (National Institutes of Health, Bethesda, MD). Virus was propagated in *Spodoptera frugiperda* (Sf9) cells (26).

Purification of Recombinant UL9. Sf9 cells were infected with AcNPV/UL9 and, after 50 hr at 27°C (5), the nuclei were isolated and proteins were extracted with high salt (1.5 M NaCl) (27). This suspension was dialyzed against PC buffer (20 mM Hepes, pH 7.6/2 mM 2-mercaptoethanol/0.2 mM EDTA/10% glycerol/1 mM phenylmethylsulfonyl fluoride/5 μg of leupeptin per ml/0.7 μg of pepstatin A per ml) containing 0.25 M NaCl. The dialysate was applied to a 10-ml phosphocellulose column (Whatman). The bound protein was eluted with a 100-ml linear gradient of 0.25–0.8 M NaCl in PC buffer, and those fractions (0.35–0.45 M NaCl) containing UL9 protein [as assayed by SDS/polyacrylamide gel electrophoresis, nitrocellulose filter binding (14), and gel mobility shift assays (28)] were pooled. This material was brought to 0.4 M NaCl and 20% glycerol in PC buffer and applied to a 5-ml heparin-Sepharose column (Pharmacia/LKB) that had been equilibrated with the same buffer. The

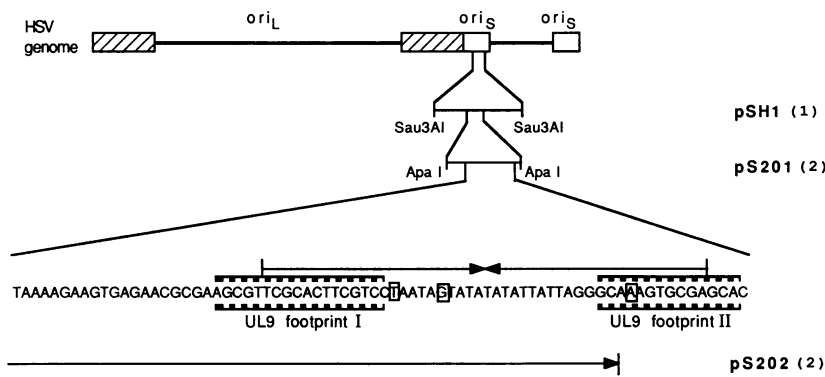


FIG. 1. HSV origin-containing plasmids. Schematic map of the HSV genome indicating the positions of *ori_L* and *ori_S*. The open boxes represent the inverted repeat regions of the S segment and the striped boxes represent those of the L segment. The nucleotide sequence of the core origin from HSV-2 is indicated as well as those bases that are different in HSV-1 (boxed). The sequences contained in the plasmids are indicated by the restriction enzyme-bracketed lines. pS202 contains the sequence shown by the arrow. Numbers in parentheses after plasmid names refer to HSV-1 or HSV-2. The sequences footprinted by UL9 protein are indicated (5, 10, 11). Site I is to the left.

column was developed with a 50-ml linear gradient of 0.4–1.1 M NaCl in PC buffer (20% glycerol). The peak fractions (0.5–0.6 M NaCl) were pooled and then dialyzed against PC buffer with 0.2 M NaCl and 25% glycerol. The dialysate was applied to a 1.5-ml single-stranded DNA-cellulose column (Sigma) that had been equilibrated with the same buffer. The column was eluted with a 15-ml linear gradient of 0.2–0.7 M NaCl in PC buffer (25% glycerol). The peak fractions (0.3–0.4 M NaCl) were pooled and applied directly to a 1-ml hydroxylapatite column (Bio-Rad). The column was washed with 2 ml of HAP buffer (25 mM Tris, pH 7.5/25% glycerol/2 mM 2-mercaptoethanol and the protease inhibitors present in PC buffer) containing 10 mM potassium phosphate (pH 7) and eluted with a 12-ml linear gradient of 10–300 mM potassium phosphate (pH 7). The peak fractions (70–100 mM potassium phosphate) were pooled and used in all experiments. The UL9 protein (160 μg/ml) was >90% pure, as determined by SDS/polyacrylamide gel electrophoresis (silver and Coomassie blue stain).

Electron Microscopy. Approximately 7 ng of DNA was mixed with 20–40 ng of UL9 protein, unless otherwise noted, in 50 mM Hepes, pH 7.5/5 mM MgCl₂/150 mM NaCl/0.025% Tween 20 in 10 μl for 10 min at room temperature. An aliquot (5 μl) of this reaction mixture was added to a 30-μl drop of 2

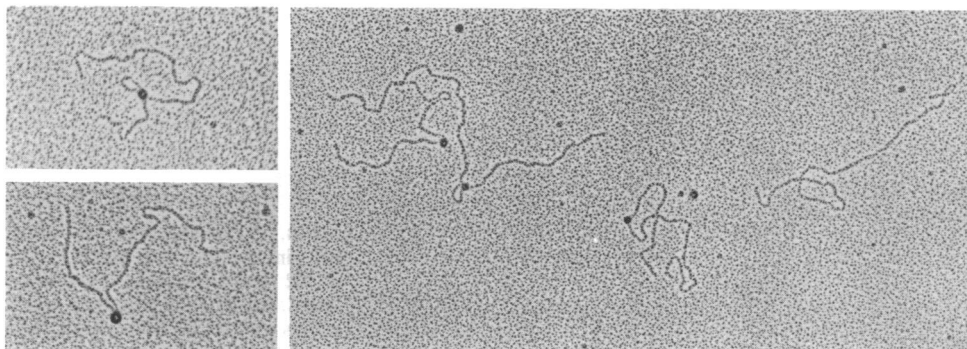
mM spermidine/150 mM NaCl (29) on clean Parafilm. After 5 min at room temperature, a freshly glow-discharged Formvar/carbon-coated grid was touched to the drop. Excess liquid was removed from the grid by blotting with filter paper. The grid was then washed in H₂O, stained with 5% uranyl acetate in H₂O, and again washed in H₂O. Rotary shadowing was with tungsten wire (0.051-cm diameter). Samples were examined in a JEOL 1200 electron microscope, and the lengths of projected molecules were measured on a Numonics (Lansdale, PA) 2400 digitizing board.

RESULTS

Electron Microscopic Visualization of UL9 Nucleoprotein Complexes. The sequence present at the origin of HSV DNA replication is illustrated in Fig. 1. Also indicated are those nucleotides protected from DNase I digestion by UL9 protein binding (UL9 footprint, Fig. 1). In the studies described here, we have used plasmids pS201 (13), containing a 150-bp insert from HSV-2 *ori_S*, pS202 (13), an identical plasmid containing a 30-bp deletion that removes UL9 protein-binding site II (Fig. 1), and pSH1 (30), containing HSV-1 *ori_S*.

HSV UL9 protein has been purified from recombinant baculovirus AcNPV/UL9-infected insect cells by a modifi-

pS201



pS202

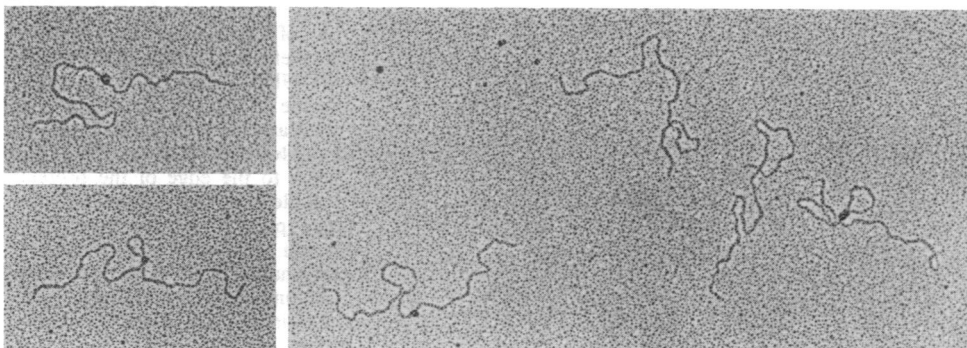


FIG. 2. Electron micrograph of UL9 protein bound to plasmids pS201 and pS202 linearized with *Xmn* I. UL9 protein complexes are seen as dense spheres. The distance from the DNA end to the origin on the short arm should be 1250–1300 bp. Magnification is such that 1 mm corresponds to 75 nm.

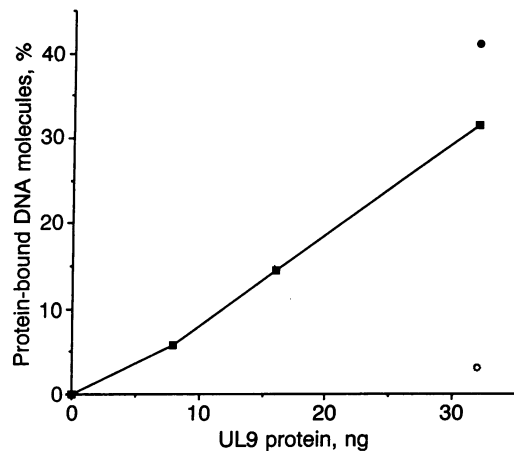


FIG. 3. Effect of increasing concentration of UL9 protein on the number of DNA molecules (pS201/*Xmn* I) containing a UL9 nucleoprotein complex. ■, Values obtained in a single experiment. In a second experiment, UL9 protein either was preincubated alone (●) or with a 33-fold molar excess of an oligonucleotide (5'-GATCTGCGAAGCGTTCGCACTTCGTCCCAATG-3') containing only site I (○).

cation of the published procedures (4, 5). Purified, recombinant UL9 protein was mixed with linearized DNA and spread immediately for electron microscopy, in the absence of any fixative. Any higher order structures will therefore not be due to fortuitous cross-linking by fixative. Tween 20 was an essential component in the reaction mixture to stabilize the complex enough to be visualized.

Large nucleoprotein structures are clearly visible at a unique site on the DNA (Fig. 2). The protein complex is

round, with many complexes containing a clear area (unshadowed) in the center. The size and structure of the UL9 complex on the DNA are virtually identical on plasmids pS201 and pS202 (lacking binding site II) (Fig. 2). Although some of the molecules have a bend at the site of complex binding, this is not true of all molecules. Many of the complexes are present at crossover points in the DNA leading to loops. There is no consistent size to the loops so we do not believe there is a second specific binding site on the DNA, but rather the bound complex may nonspecifically interact with other sites on the DNA. Of 125 measured molecules, 25% had a single crossover at the site of the nucleoprotein complex, whereas only 15% had a crossover not at the site of the nucleoprotein complex.

Measurement of the distance from the end of the DNA molecule to the protein complex demonstrated that the complex lay over the origin of DNA replication (data not shown). Fine mapping of these distances with a short fragment containing the origin will be discussed later. The complex present on pSH1 (*ori*_{S-1}) also lay over the origin region (data not shown). Very few of the complexes mapped to regions other than the origin on these molecules, which contained 3 or 4 kb of non-HSV DNA. Because no fixative was used in these experiments, some of these rare cases could be due to the fortuitous overlapping of free complex and DNA during spreading of the sample. The proportion of protein-bound DNA molecules increases linearly with increasing UL9 protein (Fig. 3). The amount of DNA bound per unit amount of UL9 protein is much less than the amount bound in filter-binding or gel mobility shift assays, suggesting that a large proportion of the complexes dissociate during the spreading protocol. This makes it impossible to estimate the number of UL9 proteins present in any single complex. Complex formation is competitively inhibited by an excess of

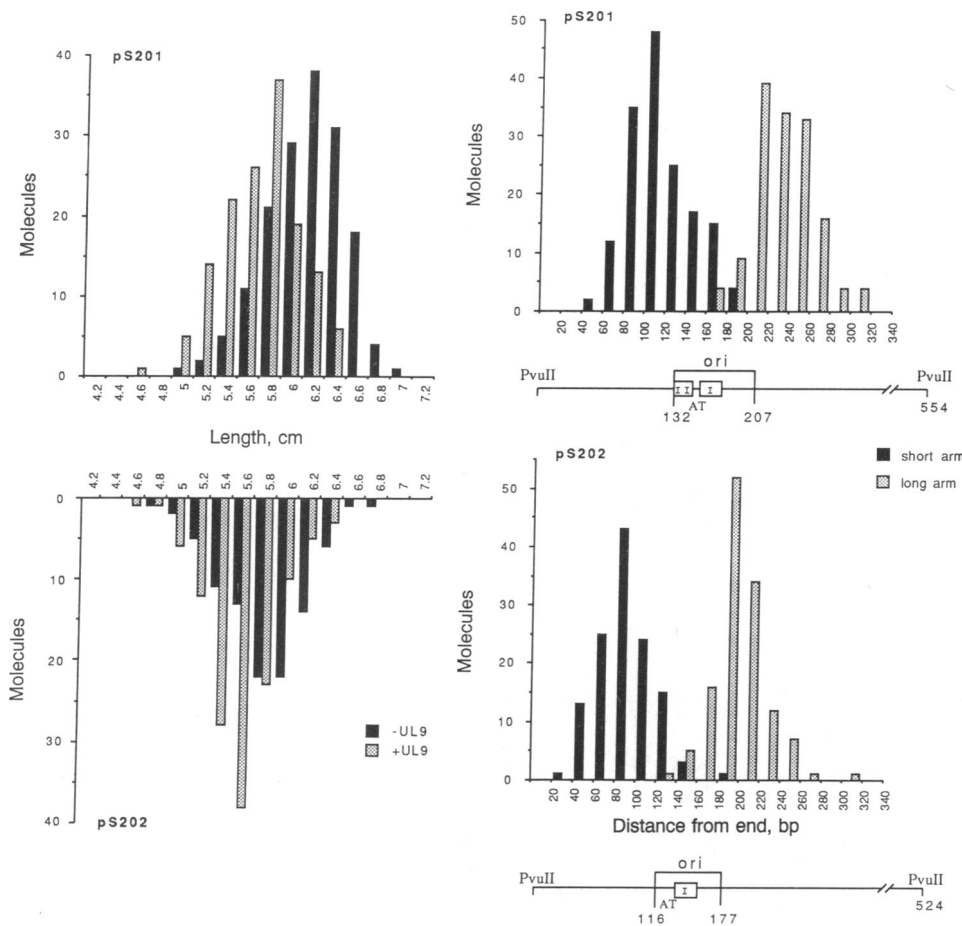


FIG. 4. Quantitation of UL9 protein binding to a pS201 and pS202 *Pvu* II fragment containing *ori*_{S-2}. (Left) Total length of molecules without UL9 protein bound (solid bars) versus total length of molecules with UL9 nucleoprotein complex (stippled bars). The length measured is from a projection of the negative and therefore is given in cm. (Right) Position of the UL9 nucleoprotein complex on the DNA. The distance of the short arm (solid bars) and long arm (stippled bars) from the end of the DNA to the edge of the protein complex has been converted into base pairs using the mean total length of the DNA fragment. A schematized map of the fragment showing the position of the origin and binding sites I and II is illustrated below.

Table 1. Comparison of the lengths of free DNA and DNA complexed to UL9 protein

	Mean length		Molecules measured
	cm	bp*	
pS201/ <i>Pvu</i> II	6.1	554 ± 33	127
+ UL9	5.7	516 ± 32	140
Δ (wrapped DNA)		38	
Distance of UL9 from end			
Short arm	1.2	112 ± 26	144
Long arm	2.6	231 ± 29	143
Δ (UL9 covering)		119	
pS202/ <i>Pvu</i> II	5.8	524 ± 32	98
+ UL9	5.5	502 ± 29	125
Δ (wrapped DNA)		22	
Distance of UL9 from end			
Short arm	1.0	93 ± 28	125
Long arm	2.2	200 ± 27	127
Δ (UL9 covering)		107	

Data represent an analysis of the measurements shown in Fig. 4. *Mean ± SD.

oligonucleotide containing only binding site I, further supporting the sequence specificity of the interaction.

Mapping of UL9 Nucleoprotein Complexes at the HSV Origin of DNA Replication. To precisely map the location of UL9 binding at the origin and determine if any structural alterations were introduced, we repeated these experiments using a short *Pvu* II fragment containing the HSV origin region from plasmids pS201 and pS202 (ori_{S-2}). On the same electron microscope grids, we measured the total length of the DNA fragment with and without bound protein (Fig. 4 *Left*). These experiments show that upon binding of the UL9 protein complex, the DNA fragment is shortened in length (Table 1). This shortening or condensation of the DNA (38 bp for pS201 and 22 bp for pS202) is probably due to the wrapping of the DNA about the protein complex in some fashion. It is interesting to note that the fragment lacking

binding site II was shortened to a lesser extent than the complete origin fragment.

By measuring the distance from the end of the DNA molecule to the edge of the complex of the short and long arms, we can determine the amount of DNA covered by the UL9 protein complex (Fig. 4 *Right*). Again, the length of DNA covered on pS201 (119 bp) is longer than on pS202 (107 bp), lacking binding site II. The amount of DNA covered is larger than the minimal origin region and significantly larger than the region that is protected from DNase I digestion (18).

Intermolecular Interactions Between UL9-Bound Complexes. A striking feature of complex formation with the short, origin-containing fragments was the generation of intermolecular complexes (Fig. 5). Very often two UL9 protein complexes (each containing a clear area in the center) are present side by side at the junction of the two DNA molecules. The proportion of dimer molecules ranged from 13% for pS201 to 36% for pS202. Another notable feature is that all of those dimer molecules formed with the pS202 origin fragment (lacking binding site II) are in parallel orientation—that is, both short arms are on the same side of the complex, in some cases lying on top of each other (Fig. 5 *Right, middle*). The orientation is not as clear-cut with the pS201 origin fragment, where about a quarter of the dimer molecules look as if they are in an anti-parallel configuration (Fig. 5 *Left, middle*). This suggests that the UL9 protein complex bound at the origin has a defined orientation with respect to its ability to form an intermolecular complex and this orientation is less fixed when the complex is bound at the complete origin (with two binding sites). The formation of these double adjacent protein complexes is not seen when only a single DNA molecule is bound, suggesting that this interaction requires both protein complexes to be bound specifically to DNA.

DISCUSSION

In many of the DNA replication systems that have been studied *in vitro*, the first step in initiation is the recognition of

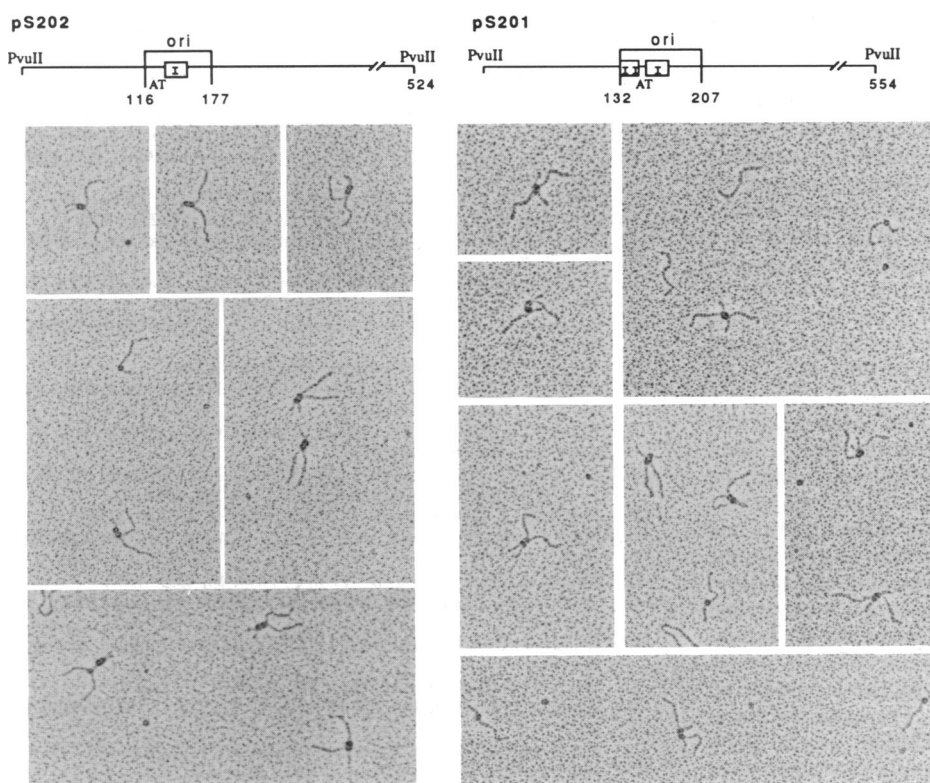


FIG. 5. Electron micrograph of UL9 protein-complexed dimer molecules of the origin-containing *Pvu* II fragment. (*Left*) Samples containing the pS202 fragment. (*Right*) Samples containing the pS201 fragment. A schematized map of the fragment showing the position of the origin and binding sites I and II is illustrated above. Magnification is such that 1 mm corresponds to 40 nm.

a specific DNA sequence at the origin of DNA replication by an origin-binding protein. Upon binding of this protein to the DNA, a higher order nucleoprotein structure is usually formed, consisting of multimers of this protein. Very often this interaction leads to an alteration in the DNA structure involving the wrapping of the DNA about the protein complex, perturbation of sequences adjacent to the binding site, and multiple protein-protein interactions. At the end of this complex assembly, those proteins involved in initiating the replication reaction and assembling the replication fork have been recruited to this site on the DNA and an unwound region has been formed.

Initiation of HSV DNA replication probably follows the same general pathway; however, this process has not yet been reconstituted *in vitro*. HSV DNA contains defined regions that behave as replication origins in transient assays and on defective interfering genomes, called *ori_S* and *ori_L*. HSV encodes an origin-binding protein (UL9) that specifically interacts with sequences present at these origins (5) and is essential for viral DNA replication *in vivo* (31). Use of the electron microscope, as in other systems, allows us to visualize the nucleoprotein complex, characterize the pathway of higher order assembly, and determine the multiple protein-DNA, protein-protein interactions occurring.

Purified, recombinant UL9 protein (from baculovirus vector-infected insect cells) forms a large nucleoprotein structure at the origin of HSV DNA replication that is ≈ 16 –21 nm in diameter with an unshadowed central region, similar in size to that seen for the λ O-some (22) and the λ Int protein-attP DNA complex (32). The approximate size and shape of this complex are the same whether or not binding site II is present at the origin. Because of this, we believe that UL9 protein binds to site I, forming a nucleoprotein complex, and then interacts with site II. The proportion of bound molecules increases linearly with increasing UL9 protein concentration.

Binding of the protein complex to the origin region leads to a shortening or condensation in the total length of the linear DNA, suggesting that the DNA is wrapped by the protein. There is less wrapping of the origin DNA that is deleted for binding site II, suggesting that the regions that are wrapped are those containing the UL9 binding sites. It has been demonstrated that the (A+T)-rich region between binding sites I and II (Fig. 1) becomes susceptible to mung bean nuclease digestion and KMnO₄ modification upon binding of UL9 protein (18), characteristic of DNA that is distorted.

The nucleoprotein complex is able to interact in an intermolecular fashion specifically and nonspecifically. About a quarter of the complexes occur at crossover sites on the DNA, forming loops of variable size, which suggests that the complex is interacting with a second DNA strand in a non-sequence-specific manner. The bound protein complex also interacts in a specific fashion with a second bound complex to form a dimer molecule joined at the site of protein binding. This dimer molecule seems to contain two protein structures that lie adjacent to one another on the DNA. Those dimer molecules formed with origin DNA deleted for binding site II always have the DNA arranged in a parallel fashion, indicating a unique orientation between the bound complexes. This orientation is not completely maintained with the complete origin, where about 20–30% of the dimer molecules are in the antiparallel orientation. These altered intra- and intermolecular interactions of UL9 protein bound to the origin region lacking binding site II may explain why plasmid pS202 is inactive in transient replication assays even though a normal size nucleoprotein complex is formed at the origin.

This intermolecular interaction between multiple nucleoprotein structures resulting in dimer molecules could be involved in a number of viral functions. For example, these structures may play a role in recombination, as the DNA molecules are brought into close proximity at homologous

sites, possibly even synapsing at the complex. Many intertypic recombinant crossovers occur within the region of the origins (33). Intermolecular interactions between origins on the same viral chromosome, in *cis*, could play a role in segregating various regions of the chromosome, possibly affecting transcription. The fact that UL9 can participate in multiple intra- and intermolecular interactions opens the possibility that it may also interact with other proteins. Multimer formation could also lead to an accumulation of initiating genomes at discrete regions within the nucleus.

It is very likely that the generation of this UL9 nucleoprotein complex is just the first step in the process of initiation, as in many other systems, and other proteins will build upon it, forming a more complex structure that will lead to the establishment of a functional replication fork. We have reconstituted DNA synthesis on an artificial replication fork *in vitro* with proteins from HSV-infected cells (27); now the steps between this and the formation of a UL9 nucleoprotein complex at the origin must be filled in.

We are grateful to Mark Challberg for providing the UL9 recombinant baculovirus and antisera to UL9 protein and Dan Lockshon and Nigel Stow for providing plasmids. We thank Ken Mariani and Stewart Shuman for critical reading of the manuscript. These investigations were supported by grants from the G. H. and L. Y. Mathers Charitable Foundation and the Society of Memorial Sloan-Kettering Cancer Center.

- McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E. & Taylor, P. (1988) *J. Gen. Virol.* **69**, 1531–1574.
- McGeoch, D. J., Preston, V. G., Weller, S. K. & Schaffer, P. A. (1990) *Genetic Maps* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 115–120.
- Challberg, M. D. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9094–9098.
- Elias, P., O'Donnell, M. E., MocarSKI, E. S. & Lehman, I. R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6322–6326.
- Olivo, P. D., Nelson, N. J. & Challberg, M. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5415–5418.
- Bruckner, R. C., Crute, J. J., Dodson, M. S. & Lehman, I. R. (1991) *J. Biol. Chem.* **266**, 2669–2674.
- Vlazny, D. A. & Frenkel, N. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 742–746.
- Spaete, R. R. & Frenkel, N. (1982) *Cell* **30**, 295–304.
- Stow, N. D. (1982) *EMBO J.* **1**, 863–867.
- Elias, P. & Lehman, I. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2959–2963.
- Koff, A. & Tegtmeyer, P. (1988) *J. Virol.* **62**, 4096–4103.
- Stow, N. D. & McMonagle, E. C. (1983) *Virology* **130**, 427–438.
- Lockshon, D. & Galloway, D. A. (1988) *Mol. Cell. Biol.* **8**, 4018–4027.
- Elias, P., Gustafsson, C. M. & Hammarsten, O. (1990) *J. Biol. Chem.* **265**, 17167–17173.
- Weir, H. M. & Stow, N. D. (1990) *J. Gen. Virol.* **71**, 1379–1385.
- Deb, S. & Deb, S. P. (1989) *Nucleic Acids Res.* **17**, 2733–2752.
- Hernandez, T. R., Dutch, R. E., Lehman, I. R., Gustafsson, C. & Elias, P. (1991) *J. Virol.* **65**, 1649–1652.
- Koff, A., Schwedes, J. F. & Tegtmeyer, P. (1991) *J. Virol.* **65**, 3284–3292.
- Bramhill, D. & Kornberg, A. (1988) *Cell* **54**, 915–918.
- Borowiec, J. A., Dean, F. B., Bullock, P. A. & Hurwitz, J. (1990) *Cell* **60**, 181–184.
- Echols, H. (1990) *J. Biol. Chem.* **265**, 14697–14700.
- Dodson, M., Roberts, J., McMacken, R. & Echols, H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4678–4682.
- Dean, F. B., Dodson, M., Echols, H. & Hurwitz, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8981–8995.
- Mastrangelo, I. A., Hough, P. V. C., Wall, J. S., Dodson, M., Dean, F. B. & Hurwitz, J. (1989) *Nature (London)* **338**, 658–662.
- Borowiec, J. A. & Hurwitz, J. (1988) *EMBO J.* **7**, 3149–3158.
- Summers, M. D. & Smith, G. E. (1987) *Tex. Agric. Exp. Stn. Bull.* **1555**, 1–48.
- Rabkin, S. D. & Hanlon, B. (1990) *J. Virol.* **64**, 4957–4967.
- Weir, H. M., Calder, J. M. & Stow, N. D. (1989) *Nucleic Acids Res.* **17**, 1409–1425.
- Griffith, J. D. & Christiansen, G. (1978) *Annu. Rev. Biophys. Bioeng.* **7**, 19–35.
- Stow, N. D. (1985) *J. Gen. Virol.* **66**, 31–42.
- Carmichael, E. P., Kosovsky, M. J. & Weller, S. K. (1988) *J. Virol.* **62**, 91–99.
- Hamilton, D., Yuan, R. & Kikuchi, Y. (1981) *J. Mol. Biol.* **152**, 163–169.
- Halliburton, I. W. (1980) *J. Gen. Virol.* **48**, 1–23.