Electron microscopic mapping of proteins bound to herpes simplex virus DNA

Madeline Wu*, Richard W. Hyman** and Norman Davidson*

*Department of Chemistry, California Institute of Technology, Pasadena, CA 91125, and  
**Department of Microbiology and Specialized Cancer Research Center, The Pennsylvania State University College of Medicine, Hershey, PA 17033, USA

Received 4 June 1979

ABSTRACT

Exonuclease digestion experiments have suggested that there is a protein(s) bound close to one or both ends of herpes simplex virus-1 (HSV) DNA. The existence of such bound proteins has been positively demonstrated and their positions on the HSV genome determined by application of a newly developed method for electron microscopic mapping of proteins bound to nucleic acids. Purified HSV DNA was treated with dinitrofluorobenzene under conditions that covalently attach the dinitrophenyl (DNP) group to the proteins in a protein-nucleic acid complex. The HSV DNA-protein-(DNP)m complex was treated with rabbit anti-DNP IgG, and, in some cases, additionally treated with monovalent Fab fragments of goat anti-rabbit IgG, and mounted for examination in the electron microscope. Electron opaque dots representing the protein-(DNP)n-(IgG)m complex were seen on the HSV DNA. Direct measurements of the positions of the protein, as well as partial denaturation mapping, indicate that there are four positions for protein bound to HSV DNA: two near but not at the two ends and two at sites corresponding to the internal inverted repeats of the ends. These results suggest that there is a specific protein binding sequence within the direct terminal repeat of HSV DNA. The previous observation that HSV DNA is more sensitive to digestion by a 3' than by a 5' exonuclease then indicates that the bound protein(s) is more intimately associated with one strand of the specific sequence than with the complementary strand.

INTRODUCTION

Exonuclease digestion experiments have suggested that there is a protein bound close to one or both ends of herpes simplex virus (HSV) DNA, as extracted from infected cells. The protein, which is not removed by the Hirt extraction procedure nor by banding in CsCl, but is removed by pronase digestion (1) and by phenol, SDS extraction (2), preferentially inhibits digestion from the 3' ends of the duplex DNA by exonuclease III as compared to digestion from the 5' ends by λ-exonuclease (1). In order to confirm positively the existence of such bound proteins and to determine their positions on the HSV genome, we have applied a newly developed method of electron microscopic mapping of proteins bound to nucleic acids (3,4).
MATERIALS AND METHODS

HSV DNA preparation. Vero cells in roller bottles were infected with three-times plaque purified HSV-1 (strain Patton) at moi = 1 PFU/cell. Following 1 hr of adsorption at 37°C, the medium was decanted and replaced with 20 ml/bottle of medium 199 supplemented with 5% fetal calf serum. Radiolabel, [¹⁴C] TdR, was added to monitor the purification of the HSV DNA. The HSV DNA was purified from the Hirt supernatant of the HSV-infected cells by sequential glycerol gradient sedimentation and CsCl buoyant density centrifugation (5,1). The HSV DNA was extensively dialyzed at 4°C against 0.01 M Tris, 0.001 M EDTA, pH 8.0, buffer.

Labeling of proteins for electron microscopy. HSV DNA was treated with dinitrofluorobenzene (DNFB), with rabbit anti-DNP IgG and, when specified, with monovalent Fab fragment of goat anti-rabbit IgG, as described previously (3,4). The DNA samples were handled very carefully to minimize mechanical breakage of the long molecules. The condition of the preparation of HSV DNA was constantly monitored by electron microscopy. Samples that had been used to the point that less than 25% of the molecules were full length (or nearly full length) were discarded. In some experiments HSV DNA was treated with 1 mg/ml of Pronase (Calbiochem) at 37°C for 30 min before labeling.

Partial denaturation. The partial denaturation procedure was a modification of the alkaline-formaldehyde method of Inman and Schnös (6,7). Fresh alkaline-formaldehyde stock solution was prepared by mixing 0.68 ml of 1 M Na₂CO₃, with 10 ml of 36-38% formaldehyde (Mallinckrodt) and adjusting to pH 11.0 to 11.5 with 10 M NaOH. For partial denaturation, 5 µl of this alkaline-formaldehyde solution were mixed gently with 9 µl of HSV DNA (16 µg/ml) and 1 µl of form II SV40 DNA (9 µg/ml), both in 0.1 M NaCl, 2 mM Na₃EDTA (pH 8.5). The mixture was incubated at room temperature for time intervals of 3, 7, and 10 min. At each of these times 2 µl samples were withdrawn, mixed with 100 µl of spreading solution and spread over the hypophase immediately. Usually, the 7 or the 10 minute sample showed the desired extent of partial denaturation of 70 to 75%. The spreading solution contained 0.4 M ammonium acetate, 0.1 M sodium acetate (pH 5.5), 0.05 mg/ml cytochrome c, and 0.03 µg/ml of additional form II SV40 DNA. The hypophase was 0.15 M ammonium acetate, 0.05 sodium acetate (pH 5.5).

Preparation of single stranded foldback molecules. 20 µl of HSV DNA (10 µg/ml) in 0.1 M NaCl, 0.01 M Tris, 0.001 M EDTA, pH 8, were mixed with 2.5 µl of 1 M NaOH. After standing at room temperature for 5 min, 20 µl of H₂O and 2.5 µl of 2 M Tris·HCl were added to dilute and neutralize the solu-
tion. The solution was incubated at 65°C for 10, 20, and 30 min at which time samples were withdrawn and mounted for electron microscopy by the formamide-cytochrome c spreading method (8).

Length standards. Circular single or double stranded SV40 and φX 174 DNA's with respective lengths of 5.23 and 5.37 kb, were used as internal length standards.

RESULTS AND DISCUSSION

Labeling the proteins on duplex HSV DNA. The protein labeling technique involves specific covalent attachment of dinitrophenyl (DNP) groups to the proteins in a protein-nucleic acid complex by treatment with dinitrofluorobenzene (DNFB) followed by binding of rabbit anti-DNP IgG. If there are sufficient DNP groups bound to the protein, this primary DNA-antibody complex is sufficiently large so it can be seen in cytochrome c spreads in the electron microscope. The size and ease of visualization of the antibody aggregate around the original protein can be increased by treatment with monovalent Fab fragments of goat anti-rabbit IgG as a second antibody (3,4,).

When HSV DNA was subjected to the single antibody labeling procedure as described in Materials and Methods and mounted for electron microscopy by the formamide-cytochrome c method, a number of long DNA molecules with clearly identifiable dots were seen. These dots had the same general appearance as did those seen in known protein-nucleic acid complexes in previous studies (3,4).

A sample of 100 randomly selected apparently full length molecules (146.4 ± 13.6 kb) was examined. Typical electron micrographs of molecules with dots are shown in Fig. 1. Data on positions of the dots are given in Fig. 2. In the treated sample, 32 molecules had no identifiable dot and 98 dots were present on 68 molecules. 45 molecules had 1 dot each (Fig. 1a, b); 20 had 2 dots (Fig. 1c) and 3 had 3 dots (Fig. 1d). In a sample studied as a control, HSV DNA which was not treated with DNFB, but was incubated with the antibody to DNP was mounted for electron microscopy. In this case, for 100 randomly selected full length molecules, 78 did not have any dots, and 24 dots were observed on 22 molecules.

A histogram for the distribution of dots as a function of distance from the closer end is shown in Fig. 2a for the DNFB treated sample and in Fig. 2b for the control sample. In the latter case the dots are randomly distributed; in the former case, there is evidence for specific sites of labeling as
Figure 1. Electron micrographs of HSV DNA treated with DNFB and rabbit anti-DNP IgG: a) an intact HSV DNA molecule with one internal dot; b) a part of an intact HSV DNA molecule showing a dot close to one end; c) a part of an intact molecule showing a pair of internal dots; d) an intact HSV DNA molecule with dots close to each end and one internal dot.

Each dot is marked by an arrow. The bars represent 1 kb.
Figure 2. Distributions of protein dots as a function of distance from the closer end: a) dots on 100 randomly selected full length HSV DNA molecules after treatment with DNFB and rabbit anti-DNP IgG; b) dots on 100 HSV DNA molecules which were treated with rabbit anti-DNP IgG alone; c) a histogram on an expanded scale showing the distribution of those dots close to an end, in the regions enclosed by the arrow in Fig. 2a.

discussed in greater detail below. As another control, HSV DNA digested with Pronase was studied. 26 dots were observed among 100 long HSV DNA molecules and they appear to be randomly distributed along the DNA (data not shown).

The experiments were repeated using the Fab fragment of goat anti-rabbit IgG as a second antibody. Micrographs are presented in Fig. 3. The dots are much more prominent by this method. These observations show that the dots observed in the experiments of Fig. 1 and 2 are indeed due to the binding of rabbit IgG to the DNA. The distribution curve for the position of the dots was similar for the double antibody experiments as for the experiments of Figs. 1 and 2 (data not shown).

The data in Fig. 2 indicate that there are two regions for preferential binding of proteins to the HSV DNA. One is in a region within 1.5 kb of the end of the molecule, the other is in the interval from 21 to 29 kb from an end. Among the 100 molecules observed 70 had no dots and 30 had one dot within 1.5 kb from an end. There were no molecules with two dots close to one end and 2 molecules with one dot close to each of the two ends. These data are consistent with, but in view of the small sample, do not rigorously
prove, our interpretation that there are two binding sites: one at each end. Further evidence in support of this interpretation is given later.

A histogram on an expanded scale for the positions of dots close to an end is shown in Fig. 2c. There is an indication of two preferred positions for binding, one at about 520 nucleotides from the end and one at about 1200 nucleotides.

The data in Fig. 2 indicate that there is considerable variability in the positions of the dots. This variability may be due, entirely or in part, to the microheterogeneity in sequence which frequently occurs at one of the ends of HSV-1 DNA (9). It may be partly due to the variable occurrence of breaks and/or single strand gaps between the protein binding site and the end. Alternatively, there may be variability in the binding site between different molecules, so that one can only determine the most probable binding positions.

As stated above, the data in Fig. 2 indicate that there is a binding site or sites between 21 and 29 kb from an end. There is a small subsidiary peak from 33 to 40 kb in Fig. 2. Whether this is due to an additional weak

Figure 3. Electron micrographs of HSV DNA treated with DNFB, rabbit anti-DNP IgG, and the Fab fragment of goat anti-rabbit IgG. a), b), c) each shows a dot close to an end; d) shows two internal dots. Each dot is identified by an arrow. The bars represent 1 kb.

(a)  
(b)  
(c)  
(d)  

3432
binding site or to nonspecific background is not known. Among the 27 dots in the interval from 21 to 29 kb in Fig. 2, 6 were observed as pairs on 3 DNA molecules. These pairs were separated by 1.25 ± 0.3 kb and their midpoint was 23.6 ± 1.6 kb from an end (Fig. 1c, 3d). No molecules were seen with 2 internal dots at opposite internal positions. Therefore, we believe that the two internal binding sites are about 1.3 kb apart, and about 24 kb from an end.

If there are four specific binding sites, the observed probability of labeling per site (after subtracting the control value) is about 20%. This low efficiency may be due to one or several of the following reasons: a) not all sites are occupied by proteins simultaneously in vivo; b) some proteins are removed from the DNA during the extraction or labeling procedure; c) the efficiency of reaction of the protein with DNFB is low. Hypothesis (b) is supported by the observation that the proteins can be removed by phenol-SDS extraction (2).

Orientation of the protein binding sites relative to the partial denaturation map. Fig. 4a presents a map of the single copy and repeated sequences present on HSV DNA as deduced from studies by several investigators and summarized by Roizman (10). At the two ends, there is a direct terminal repeat, denoted a, of total length estimated between 0.3 and 1.6 kb (1, 10, 11).

Figure 4. a) A diagram of the HSV genome, adapted from that of Roizman (10), showing the relative positions of the repeated and unique sequences. b) The direct terminal repeat a is arbitrarily depicted as being composed of the segments a1, a2, a3, and a4. We propose that the protein binds to a specific sequence a2 and is more intimately associated with one strand than the other.
The sequence ab of total length about 10 kb at the left end of the molecule (in the conventional representation in Fig. 4a) is present as an inverted repeat b'ab at a distance of about 70% of the total length of HSV DNA from the left end. The sequence ca of length approximately 7 kb at the right end is present as an inverted repeat a'c', about 12 kb towards the interior of the molecule. The two inverted repeat sequences b'ab and a'c' are contiguous. Each of the two unique sequences UL and Us which lie between the inverted repeats occurs with equal probability in a direct or an inverted order, so that HSV DNA usually consists of an equimolar mixture of 4 isomers. We wish to assign the protein positions relative to this map.

One method to approach this problem is based on partial denaturation mapping as explained below. HSV DNA was subjected to denaturation conditions as described in Materials and Methods, so that, for the average molecule, about 70 to 75% partial denaturation was achieved. Representative electron micrographs are shown in Fig. 5. Under the conditions used, the G + C rich terminal segment ca remains largely duplex (Fig. 5a,b,c), with two small denaturation bubbles sometimes seen. The left terminus, ab, has a characteristic partial denaturation pattern with several duplex and several single stranded regions (Fig. 5d,e).

The same patterns are seen in the internal inverted repeats. The Us segment is almost fully and the Ul segment extensively denatured (Fig. 5f). Our observations agree in most respects with those of previous investigators (7,13,14).

In these experiments the aqueous spreading method was used since the frequency of picking up denatured molecules on the cytochrome c protein films

Figure 5. Electron micrographs of partially denatured HSV DNA molecules. a) Part of a long DNA molecule showing the characteristic denaturation pattern of the G + C rich right terminus (ac), the fully denatured Us region and the joint region (c'a'a'b'). A dot close to the end is marked by an arrow. b) Part of a long DNA molecule showing a similar denaturation profile as that of Fig. 5a and an internal dot (marked by an arrow). c) The several small denaturation bubbles in the direct repeat region a at the right end are marked by arrows. d) Parts of long DNA molecules showing the characteristic denaturation pattern of the left terminus ab and adjacent regions. A dot close to this end is marked by an arrow in d). e) An intact HSV DNA molecule after partial denaturation treatment. The characteristic denaturation patterns of the right terminus and a dot close to the left terminus (marked by an arrow) were observed. g) A single stranded (ss) SV40 DNA circle resulting from the alkaline formaldehyde denaturation incubation of form II SV40 DNA and an adjacent double stranded (ds) SV40 circle as prepared for EM observation by the aqueous cytochrome c spreading method as described in Materials and Methods. Bars in all micrographs represent 1 kb.
was higher and more reproducible than for the formamide-cytochrome c method. We observed that under these conditions of denaturation and mounting, single stranded DNA was extended and traceable, but was considerably shorter than the identical DNA segment in the double stranded form. Accordingly, double stranded, form II SV40 DNA molecules were mixed with the HSV DNA before denaturation by alkaline formaldehyde. When the HSV DNA molecules were denatured; the resulting single stranded SV40 DNA circles were then used as internal length standards for single stranded HSV DNA regions. The average length of these denatured DNA circles was $0.64 \pm 0.5$ (n = 20) of the length of the adjacent double stranded SV40 DNA circles which were added after the denaturation incubation (Fig. 5g). This calibration factor was used for length measurements of single stranded HSV DNA segments.

In order to map the protein binding sites, HSV DNA was treated with DNFB, labeled with the first and second antibody as described in Materials and Methods, and then subjected to the partial denaturation procedure. A number of molecules with dots were observed in the electron microscope. The ab and ca termini were readily recognized from the partial denaturation patterns. For each end, 100 molecules were studied. The dot distribution within 55 kb from either terminus was determined and is plotted in Fig. 6a. The

![Histogram of dot distribution](image)

**Figure 6.** The distribution of dots on partially denatured HSV DNA molecules. The histogram is oriented to line up with the diagram shown in (b), except that the scales adjacent to both ends have been expanded. The lower part of the figure shows the partial denaturation profile of the two terminal segments and adjacent regions. Blocks represent denatured loops. Shading of the blocks indicates that the denatured loops were less frequent or less reproducible in dimensions.
data in Fig. 6 show that the internal protein binding sites are located within the inverted repeat region b'a'a'c' (Fig. 4a). Furthermore, these data show conclusively that there is a protein binding site at both the ab and ca ends. Again there is considerable variability, but the average distances of the dots from the ab and ca ends are 0.8 and 1.6 kb, respectively. Thus, these results agree with the evidence presented previously that there are two binding sites at slightly different distances from the ends; the shorter distance is at the left end.

The diagrams in the lower part of Fig. 6 summarize the partial denaturation patterns of the terminal regions as observed by us and others (11,13,14).

Binding of the protein to denatured foldback DNA. In order to obtain further information about the lengths of the several inverted repeat regions and about the positions of the proteins, we have studied the binding of the proteins to alkaline denatured HSV DNA. HSV DNA was denatured and renatured so as to allow the inverted repeat sequences to foldback on single strands as described in Materials and Methods. The sample was then treated with DNFB and both antibodies, and observed by electron microscopy. After renaturation in 0.06 M NaCl at 65° for 10 min, the type of molecule most frequently observed was that due to pairing of the ca sequence with its complement a'c' giving a single strand loop of the Us sequence (Fig. 7). The measured length of the Us loop was 12.0 ± 0.6 kb (n = 20, using φX 174 single stranded DNA as a standard). The measured length of the duplex segment due to the sequence ca was 6.9 ± 0.5 kb (n = 10). These measurement are in good agreement with those of previous investigators (7,15).

Representative micrographs are shown in Fig. 8 and length measurements

![Figure 7. The distribution of dots on the foldback structure of denatured HSV DNA. The histogram has been aligned with a diagram of the foldback structure.](image-url)
Figure 8. Electron micrographs of foldback structures of denatured HSV DNA, each showing the duplex region (ds) and the single stranded loop (US). The protein dots on the duplex region in (a), (c) and (d) are marked by solid arrows. In (a), (b) and (e), there are small loops adjacent to the double strand and single strand junction points. These junction points are marked by open arrows.

are summarized in Fig. 7. A second small loop near the double strand and single strand junction was observed frequently (Fig. 8a,b). This loop may be due to the inverted version of the small single strand loop at the end of HSV DNA as reported by Hyman et al. (16) and Wadsworth et al. (17) or it may be an out-of-register loop due to the pairing of c with c' and of a with the leftward a' of the internal b'a'a'c' sequence.

The average distance between the fork at the large loop-duplex junction and the second small loop was 6.6 ± 0.5 kb (n = 10) and the average length of the small loop was 1.5 ± 0.5 kb (n = 10). Protein dots were observed rather infrequently on these denatured molecules. 16 dots were observed among 100 molecules. Among these, 2 were on the US loop, 11 were on the ca duplex region (Fig. 8a,c,d,), and 3 dots were on the single strand region just adjacent to the duplex. For those dots on the duplex region the peak dot position was 1.0 kb from the single strand:double strand junction and 5.8 kb from
the beginning of the Us loop. The three dots in the single strand region adjacent to the double stranded region were at a distance of 1.8 kb from the peak position for the dots on the duplex region (Fig. 7). The position of the main peak (10 molecules in Fig. 7) generally confirms the assignment of the protein in the ca region as discussed below; the positions of the three dots on the single strand region support our interpretation of a protein bound to the b'a' region. The observation of some dots on denatured single strands supports the hypothesis that the protein can bind to a single strand of DNA, but the low frequency with which dots were observed indicates that considerable dissociation of the protein was caused by the alkaline denaturation treatment.

FURTHER DISCUSSION

Fig. 4b presents the final interpretation of results. We depict the direct terminal repetition a as being composed of the segments a1, a2, a3, a4. We propose that there is a specific sequence at which the protein binds which we denote as a2. The data in Figs. 2 and 7 indicate that this sequence is approximately 0.52 kb from the left end of the molecules. We then assign the second and smaller peak at a distance of 1.2 kb from the end in the histogram of Fig. 2c as due to binding to the same site in a2 at the right end of the molecule. These data then predict that the total length of the terminal repeat a is 0.5 + 1.2 = 1.7 kb. The two internal binding sites separated by about 1.3 kb at 23.5 kb may then be attributed to binding to the same two sequences a2 in the internal inverted repeats. We believe that our results strongly support this hypothesis for the most probable binding sites. However, we have observed considerable variability in the positions of the dots. This may indicate that there is a somewhat extended region in which a protein binds, and, with the limited data at hand, we can only determine the average or most probable positions. The variability may be related partly to heterogeneity in HSV DNA (9).

The observation by Kudler and Hyman (1) that E. coli exonuclease III digestion to give molecules capable of cyclization is inhibited by the protein to a greater extent than is λ-exonuclease digestion suggests that the protein is more strongly associated with the sequence a2 than with its complement a2, as depicted in Fig. 4b. Suppose the protein represents an absolute block to digestion on the one strand only. Then, as shown in Fig. 4b, exonuclease III digestion of the left end would terminate after 500 nucleo-
tides \((a_1)\), whereas digestion of the right end would not be blocked. \(\lambda\)-exo-nuclease digestion of the right end would be blocked after 1200 nucleotides \((a_3, a_4)\) whereas digestion of the left end would not be blocked. Thus, cyclization would be more probable after \(\lambda\)-exonuclease digestion, in agreement with observation.

The observation that the protein tends to be lost during the alkaline denaturation treatment is consistent with the observation from gel electrophoretic studies that the protein can be dissociated by phenol, SDS treatment as well as by Pronase treatment (2). These observations indicate that the protein is not covalently bound. The strength of the noncovalent binding must be great, however, since at least some of the bound protein resists dissociation due to SDS in the Hirt extraction, and due to the subsequent CsCl banding.

It is interesting that in the case of HSV the bound protein is close to, but not precisely at, the end of the molecules, as it is for adenovirus DNA and for single stranded and double stranded polio virus RNA (4). Heumann (18) has suggested a model for the replication of the ends of linear chromosomes possessing internal repeats of their terminal sequences. He suggested the formation of base paired structures between the terminal sequence and its internal inverted repeat, and hypothesized that a protein bound in this region was probably necessary to stabilize these structures. Jacob et al. (19) have suggested that this model is applicable to the replication of HSV DNA and makes it possible to generate molecules with the several possible sequence isomers of HSV DNA involving the several possible orientations of the \(U_L\) and \(U_S\) sequences by replication from one molecule with a single orientation. It is conceivable that the protein that we have observed bound to the repeated sequences of HSV DNA plays a role in this proposed replication model.

**ACKNOWLEDGMENTS**

M.W. and N.D. acknowledge support by Grant GM 20927 with the United States Public Health Service. We thank Mrs. Linda Kudler for her excellent technical assistance. R.W.H. is supported by Contract #NOI CP 5 3516 with the Virus Cancer Program of the National Cancer Institute, by Grants CA 16498 and CA 18450 awarded by the NCI, and is the recipient of a Faculty Research Award from the American Cancer Society (FRA-158).
REFERENCES