Characterization of a Major DNA-Binding Domain in the Herpes Simplex Virus Type 1 DNA-Binding Protein (ICP8)

YISHENG WANG* AND JENNIFER D. HALL

Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona 85721

Received 25 October 1989/Accepted 19 January 1990

We have studied the major DNA-binding protein (ICP8) from herpes simplex virus type 1 to identify its DNA-binding site. Since we obtained our protein from a cell line carrying multiple chromosomally located copies of the ICP8 gene, we first analyzed this protein to assess its similarity to the corresponding viral protein. Our protein resembled the viral protein by molecular weight, response to antibody, preference for binding single-stranded DNA, and ability to lower the melting temperature of poly(dA-dT). To define the DNA-binding domain, we subjected the protein to limited trypsin digestion and separated the peptide products on a sodium dodecyl sulfate-polyacrylamide gel. These fragments were then transferred to a nitrocellulose membrane, renatured in situ, and tested for their ability to bind DNA. From this assay, we identified four fragments which both bound DNA and exhibited the expected binding preference for single-stranded DNA. The sequence of the smallest of these fragments was determined and corresponds to a polypeptide spanning residues 300 to 849 in the intact protein. This peptide contains several regions which may be important for DNA binding based on sequence similarities in single-stranded DNA-binding proteins from other herpesviruses and, in one case, on a conserved sequence found in more distant procaryotic and eucaryotic proteins.

ICP8, the major DNA-binding protein from herpes simplex virus type 1 (HSV-1), belongs to a class of single-stranded DNA-binding proteins. These proteins are required for DNA replication, presumably because they stabilize single-stranded regions at the replication fork. While it has frequently been difficult to show that single-stranded DNA-binding proteins from eukaryotic cells are actually utilized for chromosomal replication, the requirement for ICP8 in HSV replication has been confirmed. This requirement is indicated by the inability of temperature-sensitive ICP8 mutants to replicate at nonpermissive temperatures (12, 53, 59) and by the fact that ICP8 is one of seven viral proteins required to replicate a DNA plasmid carrying a viral replication origin (62). Other examples of such proteins include *Escherichia coli* single-stranded DNA-binding protein or ssb, gene 32 protein (T4 bacteriophage), gene 5 protein (fd bacteriophage), and the adenovirus DNA-binding protein (9, 11, 36).

The most characteristic property of these proteins is their strong preference for single-stranded (as opposed to double-stranded) DNA (11). This preference facilitates their ability to promote melting of certain types of nucleic acids (3, 55). ICP8, for example, enhances the melting of poly(dA-dT) at 40°C (44). In some cases, these proteins bind DNA cooperatively, utilizing protein-protein interactions between monomer subunits (3, 15). Cooperative binding by ICP8 is indicated by the formation of protein-coated DNA filaments in vitro (50) and by the assembly of ICP8 in the absence of DNA into long helical filaments (41). Another consequence of DNA binding by some of these proteins is their ability to regulate gene expression (27, 40). Mutations in the ICP8 gene result in increased accumulations of mRNAs from early and late viral genes (20, 21).

In this study, we attempted to identify the DNA-binding site within the ICP8 protein. Since ICP8 is relatively large (128 kilodaltons [kDa], 1,196 amino acids) (16, 48), its DNA-binding domain may reside in only a portion of the molecule.

Partial proteolysis has been useful for localizing the DNA-binding domains in other single-stranded DNA-binding proteins. For example, proteolytic fragments from the adenovirus DNA-binding protein (C terminus), ssb (N terminus), and gene 32 protein (residues 22 to 253) (4, 11, 23, 26, 29) have all been found to bind DNA. Encouraged by these results, we attempted to localize the binding site in ICP8 by isolating specific trypsin-induced fragments which retain DNA-binding ability.

We identified a 56-kDa peptide which binds DNA and exhibits the expected specificity for single-stranded DNA. This fragment overlaps two sites within the central portion of the molecule which have recently been shown by others to bind DNA (17, 33). Within this peptide, we identified certain regions which may participate directly in DNA binding based on sequence comparisons with other DNA-binding proteins.

MATERIALS AND METHODS

Cells and viruses. The U-35 cell line was obtained from P. Schaffer (Harvard Medical School) (42). It was derived from Vero cells (an established line of African green monkey kidney cells) and contains multiple integrated copies of the ICP8 gene. Cells were used within 15 passages of those sent by Dr. Schaffer. HSV-1 (strain KOS) was propagated in Vero cell monolayers (22).

Purification of ICP8. To obtain an infected cell extract, we infected 100 confluent 100-mm plates of U-35 cells (cultured in Dulbecco modified Eagle medium containing 10% calf serum) at 1 PFU per cell (22). After 24 h, the cells were harvested in the overlying medium, pelleted, and suspended in 30 ml of buffer containing 0.25 M potassium phosphate (pH 7.5), 1 mM EDTA, 20% glycerol, 0.5% Triton X-100, and 2 mM beta-mercaptoethanol. The cell suspension was then sonicated three times for 20 s each, and 5 M NaCl was added to a final concentration of 1.7 M (41). After 30 min on ice, cellular debris was removed by centrifugation at 26 krpm in a 50 Ti rotor at 4°C for 2 h.

ICP8 was then purified from the supernatant. First, the

* Corresponding author.

supernatant was dialyzed against buffer A (50 mM Tris hydrochloride [pH 7.5], 20% glycerol, 2 mM beta-mercaptoethanol) containing 0.15 M NaCl. Next, it was loaded onto a 25-ml single-stranded DNA-cellulose column (2) equilibrated with dialysis buffer. The column was washed extensively with the same buffer and then with 2 volumes of buffer A containing 0.3 M NaCl. The protein was eluted with buffer A containing 1 M NaCl and collected in 4-ml fractions. Typically, the protein eluted between 0.30 and 0.36 M NaCl. Samples from the column fractions were run on a sodium dodecyl sulfate (SDS)-7.5% polyacrylamide gel (14 by 14 by 0.15 cm) (31), and the presence of ICP8 protein was detected by Coomassie blue staining as a band with a molecular mass of approximately 128 kDa. Fractions containing ICP8 were pooled, dialyzed against buffer B (20 mM potassium phosphate [pH 7.5], 20% glycerol, 2 mM beta-mercaptoethanol) and loaded onto a DEAE-cellulose column (20 ml). The column was washed extensively with the same buffer, and ICP8 protein was eluted with buffer B containing 400 mM (rather than 20 mM) potassium phosphate. ICP8-containing fractions were pooled, dialyzed against buffer C (50 mM Tris hydrochloride [pH 7.5], 60% glycerol, 2 mM beta-mercaptoethanol), and stored at -20°C . Protein concentration was measured with a Bio-Rad protein assay kit with ovalbumin as a standard.

Immunoprecipitation of ICP8 protein. An infected cell extract (320 μl) which had been dialyzed against buffer A was mixed with 80 μl of $5\times$ buffer D. The recipe for $1\times$ buffer D is 50 mM Tris hydrochloride (pH 8.0), 150 mM NaCl, 2% Triton X-100, and 0.1 mM EDTA. Monoclonal antibody (15 μl) (dissolved in H_2O) was then added, and the mixture was incubated overnight at 4°C . Antibody had been made against the HSV type 2 (HSV-2) equivalent of ICP8 and was provided by K. Powell (Wellcome Research Laboratories, Beckenham, United Kingdom). A 10% suspension of IgG-sorb (The Enzyme Center) (10 μl) was added, and this mixture was incubated on ice for 30 min. The pellet was collected by spinning in a microcentrifuge for 2 min and washing twice with $1\times$ buffer D and twice with 10 mM Tris hydrochloride (pH 8.0). The protein was then suspended in 25 μl of $2\times$ SDS sample buffer (31), boiled, and electrophoresed on an SDS-7.5% gel.

Hyperchromicity of poly(dA-dT) in presence of ICP8. The ability of ICP8 to melt poly(dA-dT) at 40°C was measured by the method of Powell et al. (44). Briefly, 3.8 μg of poly(dA-dT) (Pharmacia, Inc., Piscataway, N.J.) and/or 90 μg of ICP8 were warmed to 40°C in 700 μl of 10 mM Tris hydrochloride (pH 7.6) containing 0.1 mM EDTA. The optical density at 260 nm was read at 5-min intervals in a spectrophotometer maintained at 40°C .

DNA binding studies. DNA binding was measured by the nuclease protection assay described previously (41). pBR322 plasmid DNA was labeled with tritium by nick translation, and unincorporated nucleotides were removed by Sephadex G-50 column chromatography (34). The labeled plasmid (25 ng; 2,000 cpm/ng) was heat denatured, mixed with 1.5 μg of ICP8 plus indicated amounts of competing DNA in 40 μl of buffer E (25 mM Tris hydrochloride [pH 7.6], 10 mM MgCl_2 , 5% glycerol), and incubated at 37°C for 3 min. The competing DNAs were poly(dC), poly(dA), poly(dT) (Pharmacia), M13mp8 single-stranded DNA (prepared according to reference 37), or double-stranded HSV-1 DNA (prepared according to reference 42). A 2- μl portion containing 0.69 mg of DNase I (Sigma Chemical Co., St. Louis, Mo.) per ml and 0.69 mg of venom phosphodiesterase (46 U/mg; Worthington Diagnostics, Freehold, N.J.) per ml in buffer E was added,

and incubation was continued at 37°C for 5 min. Bovine serum albumin (25 μl) at 10 mg/ml and 900 μl of 10% trichloroacetic acid containing 0.1 M sodium PP_i were then added, and the mixture was allowed to precipitate on ice for 30 min. The solutions were transferred onto Whatman GF/C filter papers and washed three times with 1 M HCl containing 0.1 N sodium PP_i and twice with ethanol. The filters were then dried and counted.

Trypsin digestion. To produce specific ICP8 fragments, we mixed 32 μg of ICP8 in buffer C with 14 μl of trypsin (Sigma) (0.2 $\mu\text{g}/\mu\text{l}$ in buffer C) and incubated it at 20°C for 1 h. Trypsin inhibitor (Sigma) (14 μl) (0.8 $\mu\text{g}/\mu\text{l}$ in buffer C) was added to stop proteolysis. Duplicate 50- μl samples were then run on an SDS-12.5% gel (14 by 14 by 0.15 cm) to separate the fragments.

Protein blotting. The gel containing the trypsin-digested fragments described above was cut in half to separate the duplicate samples. One-half was stained with Coomassie blue, and one-half was blotted onto a 0.2- μm -pore-size nitrocellulose membrane (BA83; Schleicher & Schuell, Inc., Keene, N.H.) by a modification of the procedure of Towbin et al. (57). Briefly, the gel was soaked in 500 ml of buffer F (25 mM Tris hydrochloride [pH 8.3], 192 mM glycine) for 1 h. The membrane was treated according to the instructions of the manufacturer and soaked in buffer F. Transfer was performed at 4°C in buffer F at 150 mA for about 5 h (smaller trypsin-generated fragments) or 16 h (intact ICP8 protein). The membrane was then soaked in 200 ml of buffer G (25 mM NaCl, 10 mM Tris hydrochloride [pH 7.5], 1 mM EDTA, 0.02% bovine serum albumin, 0.02% Ficoll 400, 0.02% polyvinylpyrrolidone, 2 mM dithiothreitol) for 20 min at 4°C (7) and placed in a Seal-A-Meal bag containing 15 ml of heat-denatured or undenatured ^{32}P -labeled plasmid DNA (10^6 cpm/ml) in buffer G. The plasmid DNA was a pBR322 vector containing the *Bam*HI Q fragment from HSV-1 DNA and was labeled by nick translation. The bag was shaken at room temperature for 1 h. The membrane was then washed five times with 200 ml of buffer G, dried, and autoradiographed for 2 to 3 days.

Protein sequence determination. ICP8 (400 μg) was digested with trypsin and loaded onto an SDS-12.5% polyacrylamide gel (14 by 14 by 0.3 cm). A slice of the gel from the middle was stained with Coomassie blue to locate the 56-kDa band. A strip of the gel containing the 56-kDa fragment was then cut out from the unstained part of the gel, and the fragment was eluted with an Elutrap (Schleicher & Schuell) for 16 h with gel-running buffer. The fragment was concentrated with a Centracon-30 concentrator (Amicon Corp., Lexington, Mass.) to a final volume of 100 μl in 10 mM NH_4CO_3 -0.1% SDS-0.2 mM beta-mercaptoethanol. The N-terminal sequence of the fragment was determined (1) with a protein sequencer (Applied Biosystems, Inc.).

RESULTS

Purification of ICP8. To obtain large quantities of ICP8 for analysis, we used U-35 cells which had been infected with HSV-1 as a source of protein. U-35 cells contain multiple copies of the ICP8 gene integrated into their genomes (42). Upon superinfection, ICP8 protein is overproduced. The presence of ICP8 in extracts from these cells was detected by precipitation with antibody made against the ICP8-equivalent protein from the closely related virus HSV-2. The results of electrophoresis of the precipitated protein in an SDS-polyacrylamide gel are shown in Fig. 1 (lane 3) for an extract harvested at 24 h postinfection. This time was

FIG. 1. Identification of ICP8 protein by SDS-polyacrylamide gel electrophoresis. Samples of cellular extracts from virus-infected U-35 cells at various stages of purification were electrophoresed as described in the Materials and Methods, and the protein bands were detected by Coomassie blue staining. Numbers indicate molecular weight (MW) markers ($\times 10^3$). Lanes: 1, pooled ICP8-containing fractions after DEAE-cellulose chromatography; 2, pooled ICP8-containing fractions after single-stranded DNA-cellulose chromatography; 3, immunoprecipitation of a crude extract harvested 24 h postinfection by antiserum made against HSV-2 ICP8 protein. The arrow indicates the expected position of ICP8 protein based on its amino acid sequence.

maximal for production of ICP8. The calculated molecular mass of the presumed ICP8 band (indicated by the arrow in the figure) is 121 kDa and is within 6% of the value expected from the amino acid sequence (128 kDa).

The procedure for large-scale purification of ICP8 protein by single-stranded DNA-cellulose and DEAE-cellulose column chromatography is described in the Materials and Methods. From these purification steps, pooled fractions containing ICP8 were evaluated for protein purity by SDS-gel electrophoresis (Fig. 1). After single-stranded DNA-cellulose chromatography, ICP8 was the major protein species present (lane 2). The DEAE-cellulose column (lane 1) removed some of the minor species of contaminating proteins and, presumably, any DNA which eluted from the DNA-cellulose column. These preparations typically yielded about 10 mg of ICP8 protein.

ICP8 from U-35 cells has the same DNA-binding properties as the viral protein. Since our protein was purified from a cell line containing many uncharacterized copies of the ICP8 gene (42), we wished to confirm that our protein exhibited the same biochemical properties as those shown by the viral protein. The results described below indicate that our protein has DNA-binding properties similar to those reported for viral ICP8.

First, we confirmed that our ICP8 protein can promote melting of poly(dA-dT) at 40°C, as demonstrated previously for the viral protein (44). While absorbance (at 260 nm) of the homopolymer alone was unchanged by incubation, addition of ICP8 produced a 24% increase in optical density, charac-

FIG. 2. Hyperchromicity of poly(dA-dT). The optical density at 260 nm (OD_{260}) for the homopolymer poly(dA-dT) was determined at 40°C in the presence or absence of ICP8 protein as described in the Materials and Methods. In the presence of ICP8, the optical density contributed by the protein alone was subtracted from the total absorbance, using the value obtained for an equivalent amount of protein in the same buffer without poly(dA-dT).

teristic of melting (Fig. 2). ICP8 was unable to melt linearized pBR322 double-stranded DNA under the same conditions (data not shown).

Second, we confirmed that our protein exhibits preferential binding to single-stranded DNA (32, 51), using a DNase I protection assay (41). In this assay, binding of ICP8 to 3H -labeled, heat-denatured pBR322 DNA protects the DNA from DNase I digestion. Addition of unlabeled competing DNA allowed us to evaluate the relative affinities of our protein for the labeled and unlabeled DNA species. Protection of the [3H]DNA was dramatically reduced upon the addition of small amounts of unlabeled single-stranded M13mp8 DNA (Fig. 3). In contrast, large amounts of double-stranded DNA had to be added to observe any reduction in protection.

ICP8 preferentially binds some homopolymers. The high G+C content of HSV-1 DNA (68.3% G+C) (35) suggested that ICP8 has a preference for G+C-rich DNA. To test this possibility, we compared homopolymers with M13mp8 single-stranded DNA as competing DNAs in the nuclease protection assay. Poly(dA) was a better competitor than M13mp8 DNA at all nucleotide concentrations tested (Fig. 3B). Poly(dT) gave the same results as poly(dA) (data not shown). In contrast, competition by poly(dC) was not different from that by M13 DNA at concentrations less than 30 ng. Hence, ICP8 may actually have a preference for A+T-rich DNA. This possibility is consistent with a previous observation that poly(dT) is a better competitor than single-stranded DNA in an immunoassay (32). A preference for poly(dT) is also known for gene 32 protein (T4 bacteriophage) (28).

Analysis of trypsin-digested fragments of ICP8 protein. To localize the DNA-binding domain within the ICP8 protein, we wished to identify a proteolytic fragment of ICP8 which retains the ability to bind DNA. Consequently, ICP8 was partially digested with trypsin as described in the Materials and Methods. The digestion products shown in the SDS-polyacrylamide gel in Fig. 4 (lane 1) include fragments with masses of 30, 32, 36, 56, 57, 64, and 82 kDa. A small amount of intact protein was also observed at 121 kDa.

Initially, we attempted to separate these fragments by single-stranded DNA-cellulose column chromatography. Analysis of the column fractions indicated that all the fragments bound the column at 0.15 M NaCl and were eluted

FIG. 3. DNA binding by ICP8. As described in the Materials and Methods, 25 ng of heat-denatured ^3H -labeled DNA was mixed with ICP8 in the presence of unlabeled competing DNA. The compositions of these DNAs are indicated in the figure. After 3 min, DNase I was added to digest any DNA not bound by ICP8 protein, and the amount of protein-protected ^3H -DNA was determined. These values were normalized to the value obtained in the absence of competing DNA. ds, Double stranded; ss, single stranded.

between 0.20 and 0.27 M NaCl (data not shown). These concentrations are only slightly lower than the concentrations at which the intact protein was eluted (0.30 to 0.36 M). This result suggests either that all the fragments bound to the DNA column with similar affinity or that only some fragments bound directly and other fragments were retained through protein-protein interactions. Protein-protein interactions are presumably responsible for the cooperative nature of DNA binding by ICP8 (41, 50).

A 56-kDa fragment from trypsin-digested ICP8 is able to bind DNA. To distinguish between these above possibilities, we decided to separate the trypsin fragments by SDS-polyacrylamide gel electrophoresis, allow them to renature in situ, and then determine which fragments possess the ability to bind DNA. Toward this end, the polypeptides shown in the gel in Fig. 4 (lane 1) were transferred to a nitrocellulose membrane and renatured as described in the Materials and Methods. The membrane was then probed with heat-denatured, ^{32}P -labeled DNA to determine which fragments could bind label. As shown (Fig. 4, lane 2), four fragments (56, 57, 64, and 82 kDa) were able to bind the probe and, hence, possess DNA-binding abilities. In addition, we found that intact ICP8 protein binds probe in this assay as shown in lane 2 by the small amount of ICP8 remaining after the trypsin digestion and in lane 3 by undigested ICP8 protein. No binding was detected with the smaller fragments, suggesting that they failed to bind DNA

FIG. 4. Gel electrophoresis of trypsin digestion products of ICP8 protein and measurement of the DNA-binding capacity of these products by a protein blotting assay. ICP8 was digested with trypsin, and the resultant peptides were separated on an SDS-polyacrylamide gel as described in the Materials and Methods. Half of the gel was stained with Coomassie blue (lane 1) to locate the bands, and the other half was blotted to a nitrocellulose membrane (lane 2). In a separate experiment, intact (undigested) protein was also blotted to a membrane (lane 3). These membranes were probed with heat-denatured ^{32}P -labeled DNA to detect DNA binding. Numbers indicate molecular weight (MW) markers ($\times 10^3$). The upper arrow indicates the position of intact protein (121 kDa), and the lower arrow indicates the 56-kDa fragment described in the text.

in the absence of protein-protein interactions with the larger fragments.

DNA binding to the above assay was observed at 0 (data not shown) and 25 mM NaCl but not at 50 mM (data not shown). Hence, binding is observed only within a narrow range of conditions. However, we believe that this binding reflects the major binding activity of ICP8, because none of the peptides could bind double-stranded DNA in the presence of 25 mM NaCl (data not shown). Consequently, the binding observed retains the specificity shown by intact ICP8 for single-stranded DNA. The narrow range of conditions appropriate for this assay may explain why other investigators failed to observe binding in a similar assay (5, 49).

Determination of sequence of 56-kDa DNA-binding peptide. The 56-kDa fragment was recovered from the SDS gel shown in Fig. 4, and the sequence of the N-terminal 28 amino acids was determined as described in the Materials and Methods. Examination of the entire ICP8 sequence (16) indicated that the N terminus of the fragment corresponds to residue 300. This assignment is consistent with trypsin cleavage at residue 299 (arginine). The sequence of the 28 amino acids was identical to that of the KOS strain of HSV-1, the strain used for construction of the U-35 cell line, and differed from the sequence of HSV-1 strain 17 (35) at position 306 (P in our sequence and A in strain 17).

We next determined that the C terminus of our fragment occurs between residues 793 and 849. We chose these

GP5 (16-80)	R-X ₄ -R-X ₀₄ -Y-X ₀₇ -Y-X ₀₆ -Y-X ₀₄ -K-X ₂₆ -F-X ₀₆ -R
GP32 (67-123)	K-X ₃ -K-X ₁₂ - <u>Y</u> -X ₁₄ - <u>Y</u> -X ₀₆ - <u>Y</u> -X ₀₃ -K-X ₀₄ - <u>Y</u> -X ₀₇ -K
SSB (43-86)	K-X ₅ -K-X ₀₄ - <u>W</u> -X ₀₅ - <u>F</u> -X ₀₉ -Y-X ₀₂ -K-X ₀₄ -W-X ₀₇ -R
PIKE (16-81)	R-X ₅ -K-X ₀₄ -Y-X ₀₇ -Y-X ₀₆ -Y-X ₁₉ -K-X ₁₂ -F-X ₀₆ -R
HSV-1 (803-849)	K-X ₃ -R-X ₀₉ -F-X ₀₄ -F-X ₀₄ -F-X ₀₃ -K-X ₁₀ -W-X ₀₆ -R
VZV (801-847)	R-X ₃ -R-X ₀₉ -F-X ₀₄ -F-X ₀₄ -F-X ₀₁ -R-X ₁₂ -W-X ₀₆ -R
EBV (741-771)	K-X ₃ -R-X ₁₂ -F-X ₀₃ -Y-X ₀₄ -F-X ₀₃ -K-X ₀₆ -W-X ₀₆ -R
AD-5 (410-499)	R-X ₃ -K-X ₀₃ -F-X ₃₆ -Y-X ₁₃ - <u>F</u> <u>K</u> -X ₁₆ -W-X ₁₁ -R

FIG. 6. Single-stranded DNA-binding sequence motif. Partial protein sequences from various single-stranded DNA-binding proteins were aligned by eye with that portion of the gene 5 protein (fd bacteriophage) involved in DNA binding. This analysis yielded a region in all the sequences which consists of conserved aromatic and basic amino acids (indicated by single-letter abbreviations) separated by variable numbers of unrelated residues. These spacer regions are designated by X with a subscript indicating the number of residues in each case. Underlined residues are discussed in the text. The DNA-binding proteins shown are gene 5 protein of bacteriophage fd (GP5) (13, 38), gene 32 protein of bacteriophage T4 (GP32) (61), ssb of *E. coli* (52), and DNA-binding proteins from bacteriophage ike (PIKE) (43), HSV-1 (35), varicella-zoster virus (VZV) (14), Epstein-Barr virus (EBV) (6), and adenovirus type 5 (AD-5) (30). Numbers indicate the residue numbers for each sequence shown in the alignment.

Epstein-Barr virus. Although the overall homology among these proteins is low, the existence of several conserved sites strongly suggests that these proteins utilize a similar molecular mechanism and that conserved residues are critical for protein function. The most highly conserved regions cluster in the central third of these proteins (from residues 451 to 1010 in the HSV-1 sequence) (Fig. 5). Particularly striking is the high conservation of aromatic and basic amino acids within these sites. These types of residues are directly involved in DNA binding by several single-stranded DNA-binding proteins (see below).

As shown in Fig. 5, three highly conserved regions occur within the DNA-binding domain designated site I above (residues 332 to 564) (17, 33). As previously noted (16), one of these regions contains a zinc finger motif in the HSV-1 (residues 499 to 512) and varicella-zoster virus sequences. Although zinc fingers participate in DNA binding in several proteins (56), including gene 32 protein from bacteriophage T4 (18, 19), the absence of the motif in the Epstein-Barr virus sequence (16) questions whether a zinc finger is involved in the activity of these herpesvirus proteins. However, some feature of this site does appear to be involved in DNA binding, since mutations within this region reduce binding (17).

The second potential DNA-binding domain proposed above (site II; residues 571 to 1160) also contains several conserved regions, most of which are included in our 56-kDa DNA-binding fragment. One of these conserved regions contains a DNA-binding sequence motif (boxed in Fig. 5), which we argue below is involved in DNA binding. The conservation of this motif in all three viral proteins lends support to the idea that this site participates in binding.

The so-called DNA-binding sequence motif is shown in Fig. 6 and was identified from alignments with the gene 5 single-stranded DNA-binding protein from fd bacteriophage. It occurs in binding proteins from both prokaryotic and eukaryotic organisms and consists of a conserved set of basic and aromatic amino acids separated by nonconserved residues. A similar (but not identical) motif has been previously proposed for prokaryotic DNA-binding proteins (45).

This motif appears important for DNA binding based on

information about the gene 5 protein. The three-dimensional structure of this protein complexed to DNA has been solved (8). This structure reveals four aromatic (tyrosines 26, 34, and 41 and phenylalanine 73) and four basic (arginines 16, 21, and 80 and lysine 46) amino acids which form stacking and electrostatic interactions with DNA. Although no other structures of single-stranded DNA-binding proteins have been determined, the presence of the motif in other binding proteins suggests that their binding mechanism is similar to that found for the gene 5 protein.

Spectroscopic studies and site-directed mutagenesis experiments with these other proteins have confirmed that certain amino acids in the sequence motif are critical for DNA binding. These residues are underlined in Fig. 6 and include four tyrosine residues in gene 32 protein (T4 bacteriophage) (46, 54), tryptophan 54 and phenylalanine 60 from ssb (*E. coli*) (10, 24, 25), and phenylalanine 469 and lysine 470 from the adenovirus type 5 DNA-binding protein (39, 47). In the adenovirus, the sequences surrounding the motif are highly conserved in different viral serotypes (58), giving further support to their functional significance.

Based on estimates of the numbers of nucleotides contacted by monomers of these different DNA-binding proteins, the actual size of the binding site appears to vary in each case. For example, a monomer of the gene 5 protein binds 5 nucleotides, while a monomer of the herpesvirus ICP8 contacts at least 12 (41, 50). We suggest, therefore, that the DNA-binding sequence motif constitutes the core of a DNA-binding site and that, where appropriate, additional basic and aromatic residues flanking this site enlarge the site to accommodate a greater number of nucleotides. In ICP8, this expanded binding site might include the upstream DNA-binding site discussed above (included within residues 332 to 564).

In summary, we have identified a 56-kDa fragment of ICP8 which binds to DNA and shows the same binding specificity as the intact protein. This fragment encompasses two domains shown by others to bind DNA and includes several sites which are highly conserved among the binding proteins from herpesviruses. One conserved site encompasses a sequence motif (the DNA-binding sequence motif) which is

also conserved in other unrelated single-stranded DNA-binding proteins. This motif constitutes the entire DNA-binding site in the gene 5 protein (fd bacteriophage) and may serve as a core for the DNA-binding site in ICP8. Upstream sequences also appear to participate in binding, possibly by extending the number of nucleotides contacted by the protein.

ACKNOWLEDGMENTS

We thank P. A. Schaffer (Harvard Medical School) for providing the U-35 cell line and K. Powell (Wellcome Research Laboratories) for the monoclonal antibody made against HSV-2 ICP8 protein. We are also grateful to Adams Szpacenko (University of Arizona) for technical assistance in purifying the 56-kDa protein fragment and to David Teplow (California Institute of Technology) for amino-terminal protein sequence determination.

This research was supported by Public Health Service grant R01 GM32646 from the National Institute of General Medical Sciences.

LITERATURE CITED

1. Aebersold, R. H., D. B. Teplow, L. E. Hood, and S. B. H. Kent. 1986. Electrophoretic transfer onto activated glass. High efficiency preparation of proteins from analytical sodium dodecyl sulfate-polyacrylamide gels for direct sequence analysis. *J. Biol. Chem.* **261**:4229-4238.
2. Alberts, B. M., F. J. Amodio, M. Jenkins, E. D. Gutmann, and F. L. Ferris. 1968. Studies with DNA-cellulose chromatography. I. DNA-binding proteins from *Escherichia coli*. Cold Spring Harbor Symp. Quant. Biol. **33**:289-305.
3. Alberts, B. M., and L. Frey. 1970. T4 bacteriophage gene 32: a structural protein in the replication and recombination of DNA. *Nature (London)* **227**:1313-1318.
4. Ariga, H., H. Klein, A. J. Levine, and M. S. Horwitz. 1980. A cleavage product of the adenovirus DNA binding protein is active in DNA replication *in vitro*. *Virology* **101**:307-310.
5. Arsenakis, M., and B. Roizman. 1984. A post-alpha gene function turns off the capacity of a host protein to bind DNA in cells infected with herpes simplex virus 1. *J. Virol.* **49**:813-818.
6. Baer, R., T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, P. S. Tuffnell, and B. G. Barrell. 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature (London)* **310**:207-211.
7. Bowen, B., J. Steinberg, U. K. Laemmli, and H. Weintraub. 1980. The detection of DNA-binding proteins by protein blotting. *Nucleic Acids Res.* **8**:1-20.
8. Brayer, G. D., and A. McPherson. 1984. Mechanism of DNA binding to the gene 5 protein of bacteriophage fd. *Biochemistry* **23**:340-349.
9. Campbell, J. L. 1986. Eukaryotic DNA replication. *Annu. Rev. Biochem.* **55**:733-771.
10. Casas-Finet, J. R., M. I. Khamis, A. H. Maki, and J. W. Chase. 1987. Tryptophan 54 and phenylalanine 60 are involved synergistically in the binding of *E. coli* ssb protein to single-stranded polynucleotides. *FEBS Lett.* **220**:347-352.
11. Chase, J. W., and K. R. Williams. 1986. Single-stranded DNA binding proteins required for DNA replication. *Annu. Rev. Biochem.* **55**:103-136.
12. Conley, A. J., D. M. Knipe, P. C. Jones, and B. Roizman. 1981. Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by *in vitro* mutagenesis and defective in DNA synthesis and accumulation of gamma polypeptide. *J. Virol.* **37**:191-206.
13. Cuypers, T., F. J. van der Ouderaa, and W. W. de Jong. 1974. The amino acid sequence of gene 5 protein of bacteriophage M13. *Biochem. Biophys. Res. Commun.* **59**:557-563.
14. Davison, A. J., and J. E. Scott. 1986. The complete DNA sequence of varicella-zoster virus. *J. Gen. Virol.* **67**:1759-1816.
15. Delius, H., N. J. Mantell, and B. Alberts. 1972. Characterization by electron microscopy of the complex formed between T4 bacteriophage gene 32-protein and DNA. *J. Mol. Biol.* **67**:341-350.
16. Gao, M., J. Bouchey, K. Curtin, and D. M. Knipe. 1988. Genetic identification of a portion of the herpes simplex virus ICP8 protein required for DNA-binding. *Virology* **163**:319-329.
17. Gao, M., and D. Knipe. 1989. Genetic evidence for multiple nuclear functions of the herpes simplex virus ICP8 DNA-binding protein. *J. Virol.* **63**:5258-5267.
18. Gauss, P., K. B. Krassa, D. S. McPheeters, M. A. Nelson, and L. Gold. 1987. Zinc(II) and the single-stranded DNA binding protein of bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **84**:8515-8519.
19. Giedroc, D. P., K. M. Keating, K. R. Williams, W. H. Konigsberg, and J. E. Coleman. 1986. Gene 32 protein, the single-stranded DNA binding protein from bacteriophage T4, is a zinc metalloprotein. *Proc. Natl. Acad. Sci. USA* **83**:8452-8456.
20. Godowski, P. J., and D. M. Knipe. 1983. Mutations in the major DNA-binding protein gene of herpes simplex virus type 1 result in increased levels of viral gene expression. *J. Virol.* **47**:478-486.
21. Godowski, P. J., and D. M. Knipe. 1985. Identification of a herpes simplex virus function that represses late gene expression from parental viral genomes. *J. Virol.* **55**:357-365.
22. Hall, J. D., D. M. Coen, B. L. Fisher, M. Weisslitz, S. Randall, R. E. Almy, P. T. Gelep, and P. A. Schaffer. 1984. Generation of genetic diversity in herpes simplex virus. An antimutator phenotype maps to the DNA polymerase locus. *Virology* **132**:26-37.
23. Hosoda, J., and H. Moise. 1978. Purification and physicochemical properties of limited proteolysis products of T4 helix destabilizing protein (gene 32 protein). *J. Biol. Chem.* **253**:7547-7555.
24. Khamis, M. I., J. R. Casas-Finet, and A. H. Maki. 1987. Investigation of the role of individual tryptophan residues in the binding of *Escherichia coli* single-stranded DNA-binding protein to single-stranded polynucleotides. *J. Biol. Chem.* **262**:10938-10945.
25. Khamis, M. I., J. R. Casas-Finet, A. H. Maki, J. B. Murphy, and J. W. Chase. 1987. Role of tryptophan 54 in the binding of *E. coli* single-stranded DNA binding protein into single-stranded polynucleotides. *FEBS Lett.* **211**:155-159.
26. Klein, H., W. Maltzman, and A. J. Levine. 1979. Structure-function relationships of the adenovirus DNA-binding protein. *J. Biol. Chem.* **254**:11051-11060.
27. Klessig, D. F., and T. Grodzicker. 1979. Mutations that allow human ad2 and ad5 to express late genes in monkey cells map in the viral gene encoding the 72K DNA binding protein. *Cell* **17**:957-966.
28. Kowalczykowski, S. C., N. Lonberg, J. W. Newport, and P. H. von Hippel. 1981. Interactions of bacteriophage T4-coded gene 32 protein with nucleic acids. *J. Mol. Biol.* **145**:75-104.
29. Krevolin, M. D., and M. S. Horwitz. 1987. Functional interactions of the domains of the adenovirus DNA binding protein. *Virology* **156**:167-170.
30. Kruijer, W., F. M. van Schaik, and J. S. Sussenbach. 1981. Structure and organization of the gene coding for the DNA binding protein of adenovirus type 5. *Nucleic Acids Res.* **9**:4439-4457.
31. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
32. Lee, C. K., and D. M. Knipe. 1985. An immunoassay for the study of DNA-binding activities of herpes simplex virus protein ICP8. *J. Virol.* **54**:731-738.
33. Leinbach, S. S., and L. S. Heath. 1989. Characterization of the single-stranded DNA-binding domain of the herpes simplex virus protein ICP8. *Biochim. Biophys. Acta* **1008**:281-286.
34. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual, p. 464-466. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor. 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* **69**:1531-1574.
36. McPherson, A., and G. D. Brayer. 1984. The gene 5 protein and

- its molecular complex, p. 323–392. In F. A. Jurnak and A. McPherson (ed.), *Biological macromolecules and assemblies*, vol. 2. Nucleic acids and interactive proteins. John Wiley & Sons, Inc., New York.
37. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* **101**:20–78.
 38. Nakashima, Y., A. K. Dunker, D. A. Marvin, and W. Konigsberg. 1974. The amino acid sequence of a DNA-binding protein, the gene 5 product of fd filamentous bacteriophage. *FEBS Lett.* **43**:125.
 39. Neale, G. A. M., and G. R. Kitchingman. 1989. Biochemical analysis of adenovirus type 5 DNA-binding protein mutants. *J. Biol. Chem.* **264**:3153–3159.
 40. Nevins, J. R., and J. J. Winkler. 1980. Regulation of early adenovirus transcription: a protein product of early region 2 specifically represses region 4 transcription. *Proc. Natl. Acad. Sci. USA* **77**:1893–1897.
 41. O'Donnell, M. E., P. Elias, B. E. Funnell, and I. R. Lehman. 1987. Interaction between the DNA polymerase and single-stranded DNA-binding protein (infected cell protein 8) of herpes simplex virus 1. *J. Biol. Chem.* **262**:4260–4266.
 42. Orberg, P. K., and P. A. Schaffer. 1987. Expression of herpes simplex virus type 1 major DNA-binding protein, ICP8, in transformed cell lines. Complementation of deletion mutants and inhibition of wild-type virus. *J. Virol.* **61**:1136–1146.
 43. Peeters, B. P. H., Y. Nakashima, and W. H. Konigsberg. 1983. Characterization of the DNA binding protein encoded by the N-specific filamentous *Escherichia coli* phage IKe. Binding properties of the protein and nucleotide sequence of the gene. *J. Mol. Biol.* **169**:197–215.
 44. Powell, K. L., E. Littler, and D. J. M. Purifoy. 1981. Nonstructural proteins of herpes simplex virus. II. Major virus-specific DNA-binding protein. *J. Virol.* **39**:894–902.
 45. Prasad, B. V. V., and W. Chiu. 1987. Sequence comparison of single-stranded DNA binding proteins and its structural implications. *J. Mol. Biol.* **193**:579–584.
 46. Prigodich, R. V., Y. Shamoo, K. R. Williams, J. W. Chase, W. H. Konigsberg, and J. E. Coleman. 1986. ¹H NMR (500 MHz) identification of aromatic residues of gene 32 protein involved in DNA binding by use of protein containing perdeuterated aromatic residues and by site-directed mutagenesis. *Biochemistry* **25**:3666–3672.
 47. Quinn, C. O., and G. R. Kitchingman. 1986. Functional analysis of the adenovirus type 5 DNA-binding protein: site-directed mutants which are defective for adeno-associated virus helper activity. *J. Virol.* **60**:653–661.
 48. Quinn, J. P., and D. J. McGeoch. 1985. DNA sequence of the region in the genome of herpes simplex virus type 1 containing the genes for DNA polymerase and the major DNA binding protein. *Nucleic Acids Res.* **13**:8143–8163.
 49. Roberts, C. R., A. Weir, J. Hay, S. E. Straus, and W. T. Ruyechan. 1985. DNA-binding proteins in varicella-zoster virus-infected cells. *J. Virol.* **55**:45–53.
 50. Ruyechan, W. T. 1983. The major herpes simplex virus DNA-binding protein holds single-stranded DNA in an extended configuration. *J. Virol.* **46**:661–666.
 51. Ruyechan, W. T., and A. C. Weir. 1984. Interaction with nucleic acids and stimulation of the viral DNA polymerase by the herpes simplex virus type 1 major DNA-binding protein. *J. Virol.* **52**:727–733.
 52. Sancar, A., K. R. Williams, J. W. Chase, and W. D. Rupp. 1981. Sequences of the ssb gene and protein. *Proc. Natl. Acad. Sci. USA* **78**:4274–4278.
 53. Schaffer, P. A., D. R. Bone, and R. J. Courtney. 1976. DNA-negative temperature-sensitive mutants of herpes simplex virus type 1: patterns of viral DNA synthesis after temperature shift-up. *J. Virol.* **17**:1043–1048.
 54. Shamoo, Y., L. R. Ghosaini, K. M. Keating, K. R. Williams, J. M. Sturtevant, and W. H. Konigsberg. 1989. Site-specific mutagenesis of T4 gene 32: the role of tyrosine residues in protein-nucleic acids interactions. *Biochemistry* **28**:7409–7417.
 55. Sigal, N., H. Delius, T. Kornberg, M. L. Gelter, and B. Alberts. 1972. A DNA-unwinding protein isolated from *Escherichia coli*: its interaction with DNA and with DNA polymerase. *Proc. Natl. Acad. Sci. USA* **69**:3537–3541.
 56. Struhl, K. 1989. Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. *Trends Biochem. Sci.* **14**:137–140.
 57. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. *Proc. Natl. Acad. Sci. USA* **76**:4350–4354.
 58. Vos, H. L., F. M. van der Lee, A. M. C. B. Reest, A. E. van Loon, and J. S. Sussenbach. 1988. The genes encoding the DNA binding protein and the 23K protease of adenovirus type 40 and 41. *Virology* **163**:1–10.
 59. Weller, S. K., K. J. Lee, D. J. Sabourin, and P. A. Schaffer. 1983. Genetic analysis of temperature-sensitive mutants which define the gene for the major herpes simplex virus type 1 DNA-binding protein. *J. Virol.* **45**:354–366.
 60. Wilber, W. J., and D. J. Lipman. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA* **80**:726–730.
 61. Williams, K. R., M. B. LoPresti, and M. Setoguchi. 1981. Primary structure of the bacteriophage T4 DNA helix-destabilizing protein. *J. Biol. Chem.* **256**:1754–1762.
 62. Wu, C. A., N. J. Nelson, D. J. McGeoch, and M. D. Challberg. 1988. Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis. *J. Virol.* **62**:435–443.