JOURNAL OF VIROLOGY, June 1991, p. 2829–2838
0022-538X/91/062829-10$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 65, No. 6

A 269-Amino-Acid Segment with a Pseudo-Leucine Zipper and a Helix-Turn-Helix Motif Codes for the Sequence-Specific DNA-Binding Domain of Herpes Simplex Virus Type 1 Origin-Binding Protein

SUMITRA DEB* AND SWATI PALIT DEB

Department of Microbiology, University of Texas Health Science Center, 7703 Floyd Curl Drive, San Antonio, Texas 78284-7758

Received 26 November 1990/Accepted 27 February 1991

The UL9 gene of herpes simplex virus (HSV) codes for a DNA-binding protein (OBP) that interacts sequence specifically with the origin of replication. This protein is essential for HSV DNA replication in cultured cells. The UL9 gene was cloned into a plasmid vector downstream of the SP6 RNA polymerase promoter. By using in vitro transcription and translation systems, a full-length OBP was synthesized. This synthetic protein is recognized by an antiserum generated against the C-terminal decapeptide of OBP and is functionally active in binding to Ori_S sequence specifically. The in vitro-synthesized protein has sequence specificity for binding similar to that found for the in vivo-generated OBP. A total of 14 in-frame deletion and insertion mutants of the UL9 gene were generated and expressed in vitro. Using these deletion mutants, we determined that the 269-amino-acid stretch defined by amino acids 564 to 832 localizes the Ori_S-specific DNA-binding domain. The N-terminal boundary is between amino acids 565 and 596, while the C terminus lies between amino acids 833 and 805. This segment contains a helix-turn-helix moiety and a pseudo-leucine zipper, neither of which alone can support DNA binding. The other leucine zipper from amino acids 150 to 173 is not required for the in vitro sequence-specific DNA-binding activity of OBP.

Seven virus-encoded genes have been identified that are required and sufficient for origin-dependent replication of herpes simplex virus (HSV) DNA in cultured cells (3–5, 18, 25, 28, 31). The gene products include a DNA polymerase (UL30), a single-stranded DNA-binding activity (UL29; previously termed ICP8), an abundant protein of 65 kDa (UL42) that binds tightly to double-stranded DNA in a sequence-independent manner and interacts with HSV DNA polymerase as an accessory factor (12, 13, 17), and an origin-binding protein (OBP; UL9) that sequence specifically binds to the origin of replication (10, 14–16, 20, 27, 35). Three additional proteins encoded by the UL5, UL8, and UL52 genes can be isolated from HSV-infected cells as a complex and have both primase and helicase functions (9, 13, 36).

OBP is the only protein known to date that binds sequence specifically to the origin of replication and hence may function as the initiator of DNA replication. Elias et al. (16) initially detected OBP in HSV-infected nuclear extracts. This OBP is the product of the UL9 gene (27). We (10) and others (20, 35) have also identified its presence in HSV-1-infected Vero cell nuclear extracts. OBP has been purified to apparent homogeneity from HSV-infected Vero cells (15) and from insect cells infected with recombinant baculovirus expressing OBP (5a, 20a, 25a). We have shown that the binding of OBP is correlated to the origin function of Ori_S (10). An Ori_S mutant with partial deletion in the sequence-specific binding site can neither replicate nor bind to OBP. Competition experiments with a set of OBP-binding-site mutants performed by us (10) and others (14), a demonstration of homology between origin sequences of different herpes viruses (1, 33), and methylation interference analysis (20) indicate that the sequence 5′-YGYTCGCACT-3′ is crucial for binding. Apart from the origin-binding activity, OBP has helicase and ATPase activities (2, 5a). Correlations among structure and functions of this multifunctional protein are yet to be reported in detail in the literature. Weir et al. (35) cloned approximately two-thirds of the N-terminal portion of the UL9 coding frame in a fusion system. Extracts of *Escherichia coli* transformed with this clone showed sequence-specific binding.

To our knowledge, the data in this report represent the first successful expression of OBP by using in vitro transcription and translation systems. The synthetic gene product binds to the Ori_S sequence with the same sequence specificity as does OBP from HSV-infected Vero cells or that synthesized by recombinant baculovirus. Deletion mutants of the UL9 gene were constructed to define the boundaries of the DNA-binding domain. The N-terminal boundary lies between amino acids 565 and 596, while the C-terminal boundary is between amino acids 833 and 805. These results indicate that a stretch of 269 amino acids near the C terminus is sufficient for sequence-specific Ori_S-binding activity. The C terminus, the N terminus, the presumptive leucine zipper from amino acids 150 to 173, and the presumptive ATP-binding site are not required for in vitro DNA binding. However, the data presented herein show that a pseudo-leucine zipper from amino acids 570 to 591 is required for binding. The 269-amino-acid-long region also encodes a helix-turn-helix motif, but neither this motif nor the pseudo-leucine zipper alone can function as the DNA-binding domain. Insertions of amino acids between these motifs drastically reduce DNA binding. This finding suggests that integrity of the entire segment is necessary for successful DNA binding.

* Corresponding author.

TABLE 1. UL9 mutants

Name	Type of mutation	Restriction site(s) used	No. of amino acids expected in mutant	Ori_S-binding ability[a]
Del 1-364	N-terminal deletion	*Pst*I (22479)	487	+
Del 1-534	N-terminal deletion	*Bam*HI (21655)	317	+
Del 686-851	C-terminal deletion	*Pvu*II (21204)	685	−
Del 536-851	C-terminal deletion	*Bam*HI (21655)	535	−
Del 131-596	Internal deletion	*Nco*I (21470) *Sal*I (22872)	385	−
Del 597-649	Internal deletion	*Nco*I (21470) *Nco*I (21311)	799	−
Del 212-649	Internal deletion	*Nco*I (22628) *Nco*I (21311)	414	−
Del 1-534 833-851	C- and N-terminal deletions	*Bam*HI (21655) *Sac*I (20760)	299	+
Del 1-534 805-851	C- and N-terminal deletions	*Bam*HI (21655) *Mlu*I (20848)	271	−
Del 1-534 753-851	C- and N-terminal deletions	*Bam*HI (21655) *Sma*I (21003)	219	−
Del 1-534 ins 590	N-terminal deletion and 4-amino-acid insertion (ARIR)	*Bam*HI (21655) *Stu*I (21488)	321	−
Del 1-534 ins 685	N-terminal deletion and 4-amino-acid insertion (RGSA)	*Bam*HI (21655) *Pvu*II (21204)	321	±
Del 1-534 ins 804	N-terminal deletion and 4-amino-acid insertion (DARL)	*Bam*HI (21655) *Mlu*I (20848)	321	−
Del 1-534 rep 542-564	N-terminal deletion and 4-amino-acid (ADPR) replacement of 542–564	*Bam*HI (21655) *Nae*I (21568) *Nae*I (21637)	299	+

[a] +, binds to Ori_S (90 to 100% of wild-type binding); −, does not bind to Ori_S (<1% of wild-type binding); ±, binds to some extent (about 10% of wild-type binding).

MATERIALS AND METHODS

Cloning of the UL9 gene in the polylinker of the pGEM3 vector. The entire UL9 gene from *Eco*RV (20461) to *Nar*I (23539) sites was cloned into the *Sma*I site of pGEM3 (Promega) after filling in of the *Nar*I site with the help of the Klenow fragment of *E. coli* DNA polymerase I. The fragment encoding UL9 has been isolated from plasmid clone pKG (a kind gift from S. Weller), which contains the HSV-1 DNA segment *Kpn*I (17793)-*Bgl*II (25149) cloned between *Kpn*I and *Bam*HI sites in pUC19. The UL9 gene in the pGEM3-UL9 clone can be transcribed in vitro by either SP6 or T7 RNA polymerase, but the SP6 RNA polymerase-driven transcript is the UL9 mRNA.

Construction of mutants of the UL9 gene cloned in the pGEM3 vector. Deletion mutants were generated from the UL9 gene cloned in pGEM3. Restriction sites on the UL9 gene used for constructing the mutants were *Bam*HI (21655), *Mlu*I (20848), *Nae*I (21568 and 21637), *Nco*I (21311 and 21470), *Pst*I (22479), *Pvu*II (21204), *Sac*I (20760), *Sal*I (22872), and *Stu*I (21488). The in-frame deletion mutants generated and the restriction sites used are indicated in Table 1 and Fig. 1. The C-terminal deletion mutants were generated by opening the plasmid at different restriction sites and inserting an *Nhe*I linker (CTAGCTAGCTAG) containing stop codons in all the three reading frames. The *Nco*I sites were used to generate N-terminal deletions that would use start codons in the restriction sites themselves to initiate translation. These mutants were constructed by cutting the pGEM3-UL9 clone partially with *Sal*I and *Nco*I in the UL9 coding sequence, filling in the recessed ends with the Klenow fragment of *E. coli* DNA polymerase I, and self-ligating the linear plasmid by T4 DNA ligase and ATP. One N-terminal mutant with a deletion from the N terminus up to the *Pst*I site at 22479 was used. This construction was made by cleaving the polylinker and the UL9 coding sequence with *Pst*I and self-ligating the linear plasmid after removal of the internal fragment. The *Nco*I sites at positions 21311, 21470, and 22628 were also used to generate two internal deletion mutants by deleting sequences from 21470 to 22628 or from 21311 to 21470. These two deletion mutants preserve the coding frame. An N-terminal deletion mutant was generated by deleting up to the *Bam*HI site at 21655 and inserting an *Nco*I linker to restore the reading frame right at the *Bam*HI site.

Three mutants were constructed, each having N-terminal deletions up to the *Bam*HI site (21655) but with different C-terminal deletions: one up to the *Sma*I site (21003), one up to the *Mlu*I site (20848), and one up to the *Sac*I site (20760). In addition, three insertion mutants and one replacement mutant in the context of the N-terminal deletion mutant del 1-534 (Table 1) were generated. As shown in Fig. 1, four amino acids were inserted at positions 804, 685, and 590 by using restriction sites *Mlu*I at 20848, *Pvu*II at 21204, and *Stu*I at 21488. Synthetic oligonucleotide linkers (CAGATCTG for insertion at 804 and CGCGGATCCGCG for insertions at 685 and 590) were inserted at those sites after repair of ends with either the Klenow fragment of *E. coli* DNA polymerase I or T4 DNA polymerase. *Nae*I sites were used at 21568 and 21637 to generate a replacement mutant, and the coding frame was maintained by substituting a synthetic linker CGCGGATCCGCG for DNA sequence between the two sites. Thus, the protein would have amino acids encoded by the sequence from 542 to 564 replaced by Ala-Glu-Pro-Arg.

In vitro transcription and translation of the UL9 gene cloned in the pGEM3 vector. After linearization of the ethidium bromide-cesium chloride density gradient-purified plasmid DNA by *Eco*RI, the DNA was phenol-chloroform extracted and ethanol precipitated. Capped transcripts corresponding

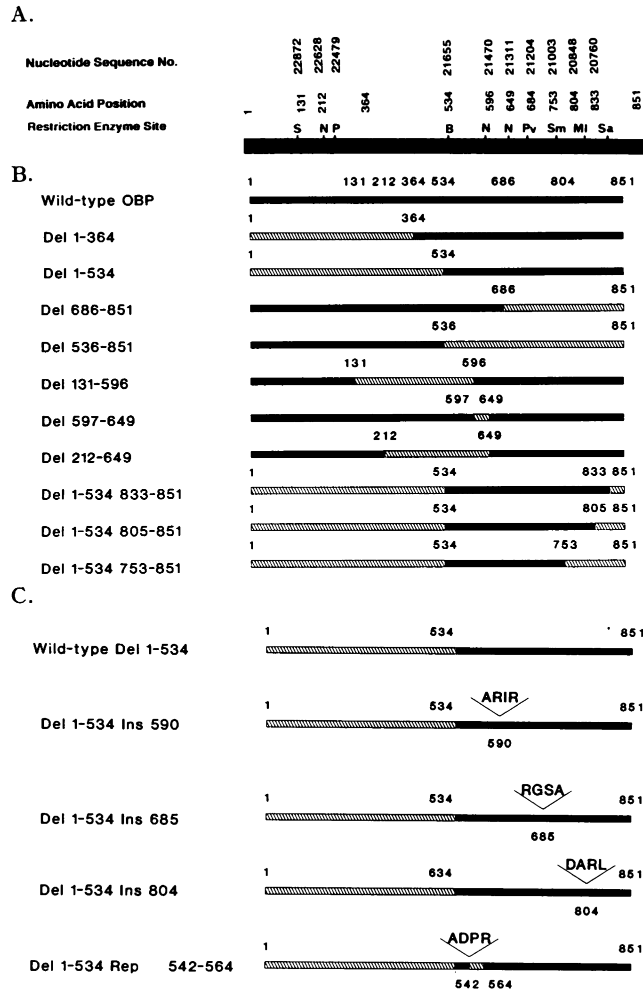


FIG. 1. (A) Schematic representation of the UL9 gene. Landmark amino acid positions and nucleotide positions for a few of the restriction enzyme cleavage sites are shown. S, *Sal*I; N, *Nco*I; P, *Pst*I; B, *Bam*HI; Pv, *Pvu*II; Sm, *Sma*I; Ml, *Mlu*I; Sa, *Sac*I. Sequence numbers are given according to reference 24. (B) Schematic representation of OBP and its deletion mutants. Positions of amino acids are indicated. The solid bar represents amino acid sequences remaining; the hatched bar represents amino acid sequences deleted. (C) Schematic representation of wild-type del 1-534 and its insertion and replacement mutants. Inserted amino acids are indicated by standard one-letter code. The solid bar represents amino acid sequences remaining; the hatched bar represents amino acid sequences deleted.

to the UL9 gene were synthesized by SP6 polymerase as described by Promega. The DNA template was then removed by treatment with RNase-free RQ1 DNase (Promega), phenol-chloroform extracted, and ethanol precipitated. The transcripts were visualized, and the length was compared against RNA molecular weight markers (Bethesda Research Laboratories) by agarose gel electrophoresis. The synthetic RNA was then translated in vitro, using rabbit reticulocyte extracts (Promega) in the presence [^{35}S]methionine to radiolabel the nascent protein as suggested by the manufacturer's protocol.

SDS-polyacrylamide gel electrophoresis and autoradiography. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis of the synthetic proteins and molecular weight markers (Sigma/Pharmacia) was carried out as described by Laemmli (22). After electrophoresis, the gel was stained with Coomassie brilliant blue, destained, and dried. The dried gel was exposed to XAR-5 or X-Omat-RP film (Kodak) with or without the use of En^{3}Hance (New England Nuclear).

Filter binding and gel retardation assays for the determination of sequence-specific DNA binding by the synthetic OBP and mutant derivatives to the Ori$_S$ sequence. DNA binding of the synthetic OBP and its derivatives was analyzed as described by Deb and Deb (10). Ori$_S$ (wild type or mutant) is cloned within *Hin*dIII and *Nco*I sites into the pOR vector (11). Thus, radioactive end labeling can be performed at either the *Hin*dIII or *Nco*I site. Sequences of individual origin derivatives, including the wild type, are shown in the figures. The in vitro translation mixture was incubated with a ^{32}P-labeled wild-type or mutant Ori$_S$ probe on ice for 30 min in a binding buffer (50 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; adjusted to pH 7.5 with NaOH], 0.1 mM EDTA, 0.5 mM dithiothreitol, 10% [wt/vol] glycerol, 100 mM NaCl) with or without competitor DNA. For these assays, different volumes of mutant OBP translation mixtures were used for binding so that comparable amounts of synthetic proteins (as judged by polyacrylamide gel electrophoresis) could be used for the binding assay.

In the filter binding assay, after incubation, the mixtures were passed through nitrocellulose filter papers and washed with the buffer mentioned above. The bound DNA was counted by Cerenkov counting and then eluted from the filter with 100 μl of 10 mM Tris-borate (pH 8.3)–0.2% SDS–10% glycerol–1 mM EDTA for 4 h at 45°C and loaded onto a 5% native polyacrylamide gel. DNA was visualized by autoradiography.

For gel retention analysis, loading dye was added to a 10% volume of the incubated sample and the mixture was loaded onto a native 6% polyacrylamide gel as described by Preston et al. (30), using a 0.5× Tris-borate-EDTA gel and performing electrophoresis at room temperature and at 30 mA. The gel was then dried and autoradiographed.

Production of antisera against the C-terminal decapeptide of the UL9 coding sequence. Antisera were prepared as described by Olivo et al. (27, 28). The C-terminal decapeptide CQGAVNFSTL was synthesized and conjugated to keyhole limpet hemocyanin by Multiple Peptide Systems. This conjugated peptide was used to raise antibody in New Zealand rabbits. The reactivity of the antiserum against the UL9 gene product was tested by Western immunoblot analysis, using purified UL9 (a generous gift of M. Challberg; 25a) and a Western blot kit from Promega according to the manufacturer's protocol. Western blot analysis revealed a band in the lane containing pure OBP at the same position where pure OBP could be seen by staining the acrylamide gel with Coomassie brilliant blue (not shown).

Immunoprecipitation of OBP and its derivatives. Immunoprecipitations of synthetic OBP and its derivatives were carried out by using the rabbit antiserum developed against the C-terminal decapeptide of the UL9 coding frame and *Staphylococcus aureus* (Enzyme Center) slurry (prepared according to the manufacturer's protocol). A protocol provided by Kathy Partin of Duke University, Durham, N.C., was used as follows. After in vitro translation, the rabbit reticulocyte lysates were diluted 10-fold in 10 mM Tris HCl (pH 7.4)–50 mM NaCl–0.1% SDS–10 mM dithiothreitol, boiled for 2 min, and again diluted 10 times in a buffer containing 10 mM Tris HCl (pH 7.4), 50 mM NaCl, 0.5% Triton X-100 (Sigma Chemical Co.), and 0.5% sodium deoxycholate. The samples were then incubated with *S. aureus*

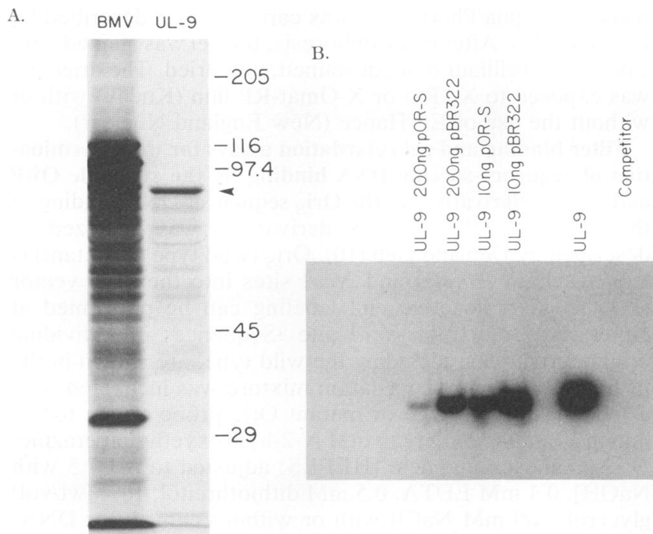


FIG. 2. (A) In vitro synthesis of OBP (^{35}S labeled) (see Materials and Methods). The autoradiogram of an SDS-polyacrylamide gel demonstrates synthesis of a protein of a size identical to that of pure OBP. The arrowhead indicates the position of pure OBP obtained by recombinant baculovirus expression. A control reaction was carried out with brome mosaic virus RNA (BMV); UL-9, product from synthetic UL9 mRNA. Sizes are indicated in kilodaltons on the right. (B) Filter binding competition experiments showing Ori_S-specific binding by the in vitro-synthesized UL9 gene product. The in vitro translation mixture was incubated with ^{32}P-labeled Ori_S probe as described in Materials and Methods. After incubation, the mixtures were passed through nitrocellulose filter papers and washed. The bound DNA was counted by Cerenkov counting and then eluted from the filter and loaded onto a 5% native polyacrylamide gel. DNA was visualized by autoradiography. Each lane contained 5 μl of in vitro-translated product, labeled Ori_S, and the indicated amount of competitor DNA.

slurry and centrifuged to remove the bacteria. Antisera were added, the mixture was incubated for 5 min at 37°C, and *S. aureus* slurry was again added. The suspension was incubated for 30 min at 4°C and centrifuged. The pellet was washed successively with (i) 10 mM Tris HCl (pH 7.4)–50 mM NaCl–0.5% Triton X-100–0.5% sodium deoxycholate, (ii) 10 mM Tris HCl (pH 7.4)–1 M NaCl, and (iii) 10 mM Tris HCl (pH 7.4)–50 mM NaCl, suspended in Laemmli loading buffer (22), boiled for 2 min, and loaded onto a 10% polyacrylamide gel containing SDS.

RESULTS

In vitro transcription and translation of the UL9 gene. An *Eco*RV (20461)-to-*Nar*I (23539) fragment containing the entire UL9 coding sequence was cloned into the *Sma*I site of the pGEM3 vector (Promega). SP6-driven transcript from this plasmid should produce a UL9 mRNA sense product. After linearization of the plasmid with *Eco*RI on the 3′ end of the insert, an in vitro transcript corresponding to the full-length insert was synthesized. This transcript migrated as a single band as assessed by 1% agarose gel electrophoresis (data not shown). This transcript was used to synthesize OBP in vitro, using commercial rabbit reticulocyte lysates (Promega) in the presence of [^{35}S]methionine for labeling the nascent protein. Figure 2A shows that the most prominent protein (lane UL-9) detected by SDS-polyacrylamide gel electrophoresis in the rabbit reticulocyte extract after translation corresponds to a molecular mass of about 82 to 83 kDa. The migration of this protein band matches exactly with the size of the purified OBP from a recombinant baculovirus expression system (27). The migration of OBP (arrowhead in Fig. 2A) from the recombinant baculovirus was visualized by staining the polyacrylamide gel with Coomassie blue. Although the predicted molecular mass from the coding frame of UL9 should be about 94 kDa, both the synthetic OBP and the purified form from insect cells infected with an OBP-expressing baculovirus show apparent molecular masses of about 82 to 83 kDa. The reason for the discrepancy between predicted and observed molecular masses is unknown. Figure 2A also shows the analysis of a control reaction in which in vitro translation was carried out with brome mosaic virus RNA. This resulted in numerous prominent bands as opposed to one major band in the UL9 lane.

Ori_S-binding activity of the synthetic OBP. To determine the specificity of binding of the in vitro-translated products to the HSV Ori_S region, the DNA-binding ability of synthetic OBP in the presence and absence of competitor DNAs was analyzed by filter binding experiments. After incubation of an extract with radiolabeled Ori_S, the mixture was filtered through nitrocellulose filter paper, the radioactivity was determined, and the labeled Ori_S was eluted from the filters and electrophoresed into a native polyacrylamide gel. The presence of a labeled fragment indicated binding of Ori_S to a factor(s) in the extract (Fig. 2B). To determine the sequence specificity of binding, competition experiments were performed in which incubation of the extract with the radiolabeled Ori_S probe was carried out in the presence of cold pBR322 (nonspecific competitor) or pOR-S (11) (self-competitor). Clearly, the translated mixture binds specifically to Ori_S because this binding was inhibited by cold plasmid containing Ori_S more strongly (4- to 10-fold) than by pBR322 DNA (0- to 2-fold), as determined by comparing counts retained on the filters when the competitor was present at 50-fold molar excess over the probe. The extract itself, which did not have synthetic OBP, did not bind the probe (see Fig. 5A). These data show that the synthetic OBP is functionally active in binding to Ori_S sequence specifically. Later it will be shown that the synthetic protein is recognized by an antiserum generated against the C-terminal decapeptide OBP.

OBP binding domain on Ori_S recognized by pure baculovirus-expressed OBP and in vitro-synthesized OBP. To determine whether synthetic OBP (in vitro-synthesized UL9) is similar in nucleotide requirements to the protein synthesized in vivo, binding-site mutant constructs described previously (10) were analyzed. Six of the binding-site mutants were substitution mutants in which three consecutive base pairs were substituted. In these mutants, T's were substituted by G's, A's were substituted with C's, and vice versa. Two other mutants, del-31 and bs-1′, whose sequences are shown in Fig. 3, were also examined. To determine the relative importance of the nucleotides in binding, the end-labeled origin insert of pOR-S1 was incubated with either the rabbit reticulocyte extract containing the synthetic OBP or the purified baculovirus-expressed OBP (25a) in the presence of individual cold plasmid DNAs mutated at the binding site in 50-fold molar excess over the probe. Mutants bs-9, bs-10, bs-11, and bs-12 were the weakest competitors of the end-labeled origin insert of pOR-S1 for complex formation with the synthetic (Fig. 3A) as well as the pure (Fig. 3B) protein. This finding suggests that bases altered in these mutants are

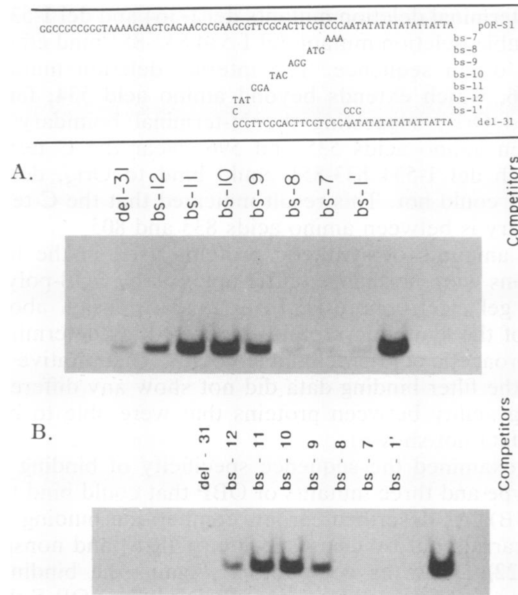


FIG. 3. Ori_S-binding activity of in vitro-synthesized OBP (A) and purified baculovirus-expressed OBP (B) in the presence of different Ori_S mutants that have mutations in the OBP-binding site. The procedure for this experiment was similar to that described in the legend to Fig. 2 except that 200 ng of each of cold competitor DNA with a modified OBP-binding site was added during incubation with radioactive Ori_S. For the synthetic protein, each system had about 0.4 pmol of protein; for pure OBP, each system had about 0.5 μg of protein. The OBP-binding site is overlined on the wild-type sequence. Sequences of the binding site mutants used for the DNA-binding competition experiments are shown at the top. Del 31 is deleted in sequences that are not depicted.

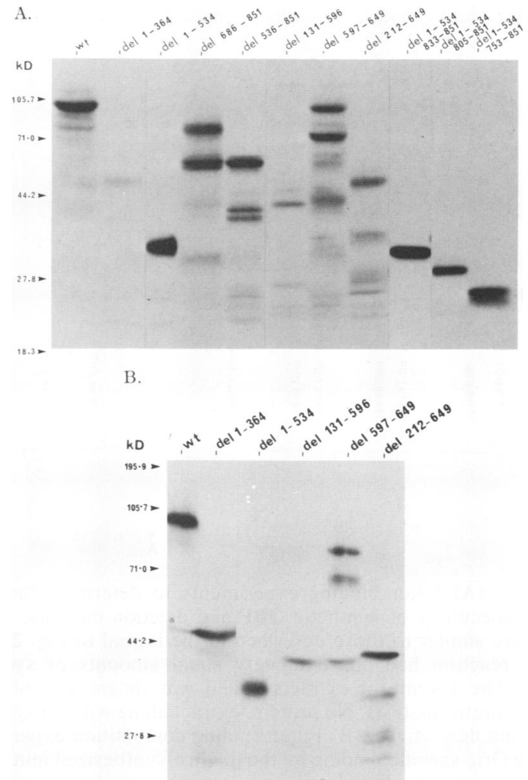


FIG. 4. (A) SDS-polyacrylamide gel analysis of in vitro-synthesized OBP and its deletion mutants. The procedure for production of UL9 and its derivatives is described in the text. Each lane contained 5 μl of a sample. (B) Immunoprecipitation of in vitro-synthesized OBP and its deletion mutants carried out as described in Materials and Methods. Only those mutants with an intact C terminus are shown. wt, wild type.

crucial for binding. These results indicate that the in vitro-synthesized protein is similar in sequence specificity for binding to pure OBP or crude OBP in HSV-1-infected Vero cell extracts (10, 20) and that the sequence 5'-GTTCG CACTT-3' is essential for efficient binding. Thus, the in vitro-synthesized OBP defines the same binding site as does the in vivo-synthesized OBP.

Construction of deletion mutants of the UL9 gene product. To determine the DNA-binding domain of OBP, deletion mutants of the UL9 gene were constructed by using convenient restriction sites in and around the UL9 coding sequence. The mutants are shown diagrammatically in Fig. 1B, and their constructions are explained in Table 1. Two C-terminal deletion mutants, two N-terminal deletion mutants, three N- as well as C-terminal deletion mutants, and three internal deletion mutants were generated. The production of proteins was measured by the appearance of ^{35}S-labeled proteins in SDS-polyacrylamide gels. Figure 4A shows protein bands detected by SDS-polyacrylamide gel electrophoresis and autoradiography. The major bands are slightly smaller in molecular weight than predicted from amino acid numbers shown in Table 1. This was expected because of the observed faster mobility of OBP in an SDS-polyacrylamide gel (27). In the case of mutant del 1-364, a deletion from the N terminus up to the *Pst*I site at 22479 was made. This deletion generated two reading frames that can start either at the Met in the original coding frame at nucleotide position 22169 (codon 365) or at a Met in an alternate reading frame six nucleotides earlier. The corresponding band in Fig. 4A is slightly diffused, probably because of protein synthesis being initiated at both sites.

All of the C-terminal deletions were generated by inserting *Nhe*I linkers with stop codons in all three possible coding frames after cutting with indicated restriction enzymes (Table 1). In the case of C-terminal deletion mutant del 1-534, an *Nco*I linker was used at the restriction site to regenerate the correct coding frame. This mutant DNA was used to generate three more deletion mutants, del 1-534 833-851, del 1-534 804-851, and del 1-534 753-851. These mutants were deleted at both the C and N termini. The three internal deletion mutants were generated by using internal *Nco*I and *Sal*I sites, preserving the coding frame.

Construction of insertion and replacement mutants of del 1-534. To investigate the internal structure of the DNA-binding domain, three insertion mutants were constructed in which four amino acids were inserted after positions 590 (Ala-Arg-Ile-Arg), 685 (Arg-Gly-Ser-Ala), and 804 (Asp-Ala-Arg-Leu). These mutants were generated by cutting the gene of del 1-534 at *Mlu*I (20848), at *Stu*I (21488), and at *Pvu*III (21204), repairing the termini, and inserting synthetic oligonucleotide linkers. One replacement mutant was generated by replacing amino acids 242 to 264 by the four-amino-acid segment Ala-Asp-Pro-Arg, using the *Nae*I sites (21568 and 21637).

Immunoprecipitation of the synthetic OBP and its deletion mutants. To determine whether the synthetic OBP and its

FIG. 5. (A) Filter binding experiments to determine the Ori_S-binding specificity of synthetic OBP and deletion mutants. Conditions were similar to those described in the legend to Fig. 2. Each binding reaction had approximately equal amounts of synthetic protein. The volume of extracts added was determined after gel electrophoretic analysis. No protein, extract alone without synthetic OBP or its derivatives. (B) Filter binding competition experiments showing Ori_S-specific binding by the in vitro-synthesized mutants of OBP. The in vitro translation mixture was incubated with ^{32}P-labeled Ori_S probe with or without indicated competitors and processed as described for Fig. 2B. In each case, the self-competitor inhibited more than did the nonself one. WT, wild type.

derivatives retain the wild-type coding frame, immunoprecipitation with an antiserum reactive against wild-type OBP was performed. An antiserum against the C-terminal decapeptide was generated according to Olivo and coworkers (27, 28). This antiserum can react specifically with purified OBP (25a) in immunodiffusion, immunoprecipitation, and Western blotting assays. It was reacted with the synthetic OBP and its derivatives. The complexes were then precipitated by *S. aureus* slurry as described in Materials and Methods. The immunoprecipitates were analyzed in an SDS-polyacrylamide gel and then subjected to autoradiography. Figure 4B shows bands in all lanes with synthetic proteins having an intact C terminus. C-terminal deletion mutants of OBP were not immunoprecipitated (data not shown). Since an antiserum was raised against the C-terminal protein, synthetic proteins with C-terminal deletions were therefore not expected to be immunoprecipitated. The data presented in Figure 4B demonstrate intactness of the C terminus, suggesting that the correct reading frames for OBP and its derivatives were retained. Because the N terminus was not altered in any way during construction of C-terminal deletion mutants, no change in the reading frame was anticipated. That the N-terminal deletion mutant del 1-364 (Fig. 1B) could be immunoprecipitated (Fig. 4B) showed that the reading frame was intact, and the migration of the band in the SDS-polyacrylamide gel suggested that the translation system utilizes the Met at 365 as the initiation codon.

Ori_S-binding activity of the synthetic OBP deletion mutants as observed by filter binding assay. Results of a filter binding experiment with in vitro-prepared synthetic proteins (Fig. 5A) show that only proteins encoded by the intact UL9 gene, the N-terminal deletion mutants del 1-364 and del 1-534, and the double-deletion mutant del 1-534 833-851 bind efficiently to the origin sequence. The internal deletion mutant del 131-596, which extends beyond amino acid 534, failed to bind to Ori_S, suggesting the N-terminal boundary to be between amino acids 535 and 596. Near the C terminus, although del 1-534 833-851 could bind to Ori_S, del 1-534 804-851 could not. This result indicated that the C-terminal boundary is between amino acids 833 and 805.

The amounts of synthetic proteins used in the binding reactions were normalized after analysis by SDS-polyacrylamide gel electrophoresis. In a standard assay, about 0.4 pmol of the synthetic proteins was used, as determined by trichloroacetic acid-precipitable counts. Quantitative analysis of the filter binding data did not show any difference in binding ability between proteins that were able to bind to Ori_S (data not shown).

We examined the sequence specificity of binding of the wild type and three mutants of OBP that could bind to Ori_S (Fig. 5B). As described earlier, competition binding assays were carried out by using specific (pOR-S) and nonspecific (pBR322) DNAs as competitors. Again, the binding was found to be competed for more efficiently by pOR-S than by pBR322 under a 50-fold molar excess of competitor over the probe. This result suggests that the observed DNA binding is highly sequence specific.

Insertion and replacement mutations in the DNA-binding domain of OBP. To investigate the internal structure of the DNA-binding domain, three insertion mutants (del 1-534 ins 590, del 1-534 ins 685, and del 1-534 ins 804) and one replacement mutant (del 1-534 rep 542-564) were generated (Fig. 6C). Figure 6 shows SDS-polyacrylamide gel analysis of these mutants and their immunoprecipitates. The sizes of the prominent bands correspond to those expected for them. Figure 7 shows a filter binding analysis of the Ori_S-binding ability of these mutant proteins. Only rep 542-564 shows binding ability. This binding is inhibited more by the presence of pGEM3z(−) Ori_S (self-competitor) than by pGEM3z(−). Although not shown in this figure, del 1-534 ins 685 bound to Ori_S; this binding, however, was considerably less than that of the wild type (about 10% of the wild-type level; data not shown). Thus, this analysis suggests that replacement of amino acids 542 to 564 by four foreign amino acids has no significant effect on Ori_S binding.

Gel retardation assay to determine the DNA-binding domain. Ori_S-binding capacities of del 1-534 and its mutants were also determined by a gel retardation analysis (Fig. 8). Figure 8A shows the schematics of mutants tested in this assay, and Fig. 8B depicts an autoradiogram in which the different mutants were analyzed in 6% native polyacrylamide gels after incubation of the in vitro-synthesized protein with wild-type and mutant origins from pOR-S1 (with OBP-binding site I) (11). As shown in the figure, only del 1-534 wild-type, del 1-534 833-851, and del 1-534 rep 542-564 formed a retarded complex, while all other mutants, including insertion mutants del 1-534 ins 590, del 1-534 ins 685, and del 1-534 ins 804, did not show complex formation. Mutant del 1-534 ins 685 formed a complex that was detected after overexposure of the dried gel. Thus, this insertion mutant retained some binding activity even after the four-amino-acid insertion at position 685. Interestingly, although both del 1-534 833-851 and del 1-534 rep 542-564 have 299 amino acids, the complex formed with the former migrates faster than that formed with the latter. Those proteins that could form the retarded complex were also tested for sequence specificity by performing binding reactions with mutant

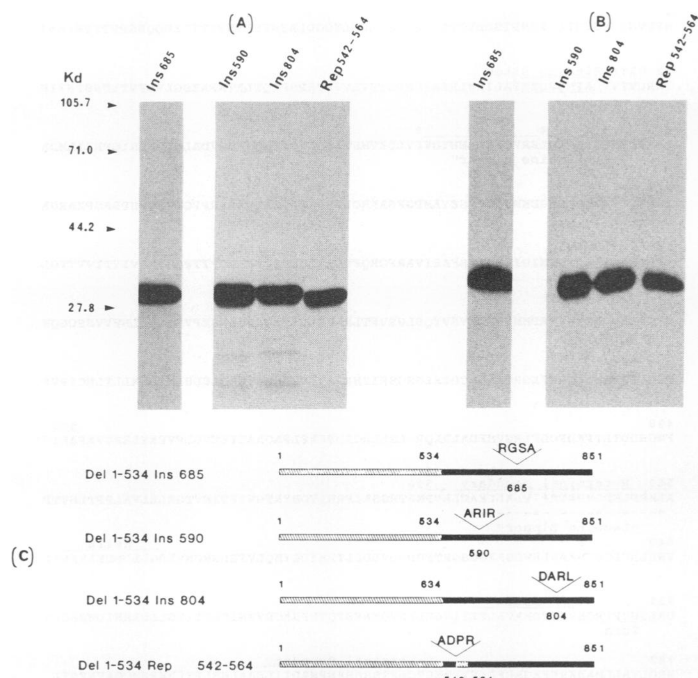


FIG. 6. (A) SDS-polyacrylamide gel analysis of in vitro-synthesized insertion and replacement mutants of del 1-534 and its deletion mutants. The procedure for production of the mutants is described in the text. (B) Immunoprecipitation of in vitro-synthesized products shown in panel A, using an antiserum directed against the C-terminal decapeptide of OBP and *S. aureus* slurry. (C) Schematic representation of wild-type del 1-534 and its insertion and replacement mutants. Inserted amino acids are indicated by standard one-letter code. The solid bar represents amino acid sequences remaining; the hatched bar represents amino acid sequences deleted.

Ori-S1 sequences. None of them could bind to Ori-S1 bs-9, -10 or -11 (10). These mutant origins are mutated in the primary binding site of OBP and failed to bind appreciably to a crude OBP preparation (10). Thus, origin-specific binding

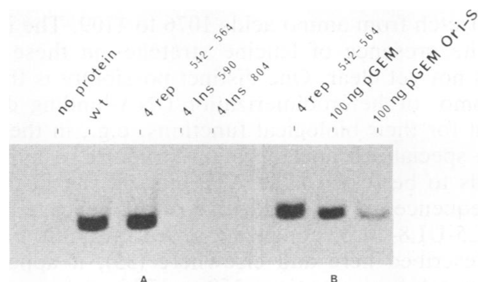
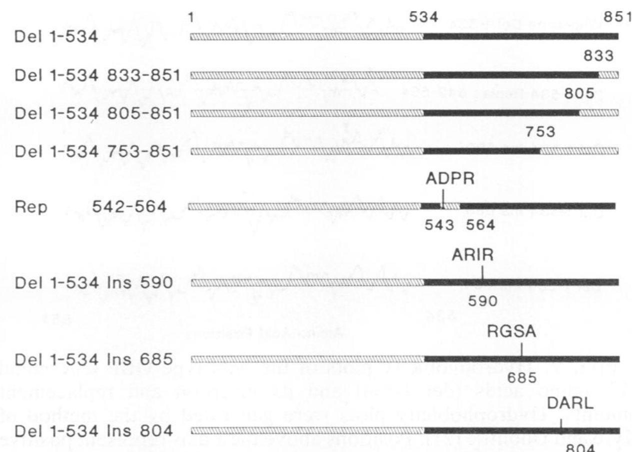


FIG. 7. (A) Filter binding experiments to determine Ori_S-binding specificity of synthetic insertion and replacement mutants of del 1-534. Conditions were similar to those described in the legend to Fig. 5. (B) Filter binding competition experiments showing Ori_S-specific binding by the in vitro-synthesized del 1-534 rep 542-564. The in vitro translation mixture was incubated with ^{32}P-labeled Ori_S probe with or without the indicated competitors and processed as described for Fig. 5B. In this case also, the self-competitor inhibited more than did the nonself one.



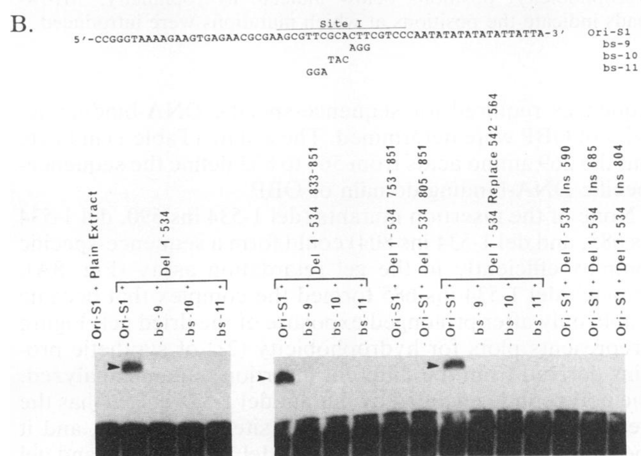


FIG. 8. (A) Schematic representation of wild-type del 1-534 and its deletion, insertion, and replacement mutants. Inserted amino acids are indicated by the standard one-letter code. The solid bar represents amino acid sequences remaining; the hatched bar represents amino acid sequences deleted. (B) Gel retardation assay to determine Ori_S-binding capacity of the C-terminal 317 amino acids of OBP (del 1-534) and their mutants. The procedure for the gel retardation analysis is described in text. The type of origin probe used (either Ori-S1 or its mutant, bs-9, -10, or 11) and the name of the synthetic proteins used are indicated at the top of each lane. Arrowheads point to the retarded bands where protein-bound DNA migrates. Free probes are at the bottom of the gel. Only the portion of autoradiographs showing the free probe and retarded bands are presented. Rabbit reticulocyte extracts alone do not form any significant retarded band. Wild-type del 1-534 and the mutants del 1-534 833-851 and del 1-534 rep 542-564 can bind successfully to wild-type Ori-S1 but not to its mutant derivatives with substitutions in the primary OBP-binding site. All other mutants failed to form a significant protein-DNA complex. Sequences of Ori-S1 and its mutants are depicted at the top.

was observed in those cases of mutant proteins that retained the capacity to bind to Ori_S.

DISCUSSION

To determine the structural motif of the protein that dictates the sequence-specific DNA-binding activity of OBP, advantage was taken of the powerful tools of in vitro transcription and translation. Using synthetic mutant proteins and filter binding and gel retardation assays, amino acid

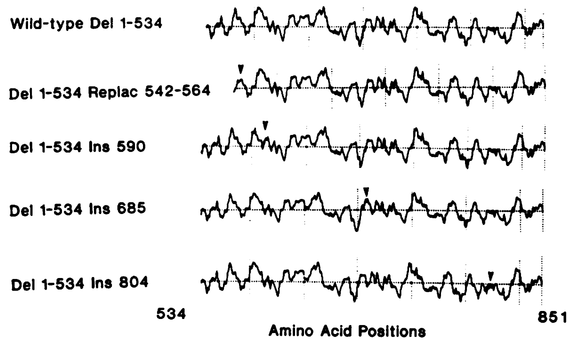


FIG. 9. Hydrophobicity plots of the wild-type OBP C-terminal 317 amino acids (del 1-534) and its insertion and replacement mutants. Hydrophobicity plots were generated by the method of Kyte and Doolittle (21). Positions above the x axis represent positive hydrophobicity; positions below indicate hydrophilicity. Arrowheads indicate the positions at which mutations were introduced.

sequences required for sequence-specific DNA-binding activity of OBP were determined. These data (Table 1) indicate that the 269 amino acids from 564 to 832 define the sequence-specific DNA-binding domain of OBP.

None of the insertion mutants (del 1-534 ins 590, del 1-534 ins 685, and del 1-534 ins 804) could form a sequence-specific complex efficiently in the gel retardation assay (Fig. 8A), although del 1-534 ins 685 formed the complex that became visible only after prolonged exposure of the dried gel. Figure 9 represents plots for hydrophobicity (21) of synthetic proteins derived from the different insertion mutants analyzed. The polypeptide encoded by mutant del 1-534 ins 590 has the greatest structural alteration at the site of insertion, and it does not bind to the origin. Mutants del 1-534 ins 685 and del 1-534 ins 804 have less effect on hydrophobicity. Of these two, del 1-534 ins 685 can bind to Ori_S to some extent while del 1-534 ins 804 cannot. It is possible that insertion of amino acids in del 1-534 ins 590 caused loss in DNA-binding activity because of structural alteration. Loss of binding capacity may not be related to possible changes in the positions of contact residues. The insertions at 685 and 804 have moderate effects on hydrophobicity but have a significant effect on binding. This might be indicative of more direct roles for the residues near 685 and 804 in DNA binding than for those near amino acid 590. More subtle mutational analysis, including amino acid substitution (both conservative and nonconservative) at these sites, is necessary to resolve this issue.

Sequence analysis by the Chou-Fasman method (7) predicted a helix-turn-helix motif at amino acids 698 to 733 (Fig. 10) in the DNA-binding domain. A helix-turn-helix motif has been implicated in DNA binding and dimerization of a number of sequence-specific DNA-binding proteins (29). However, deletion mutant del 130-596, which does not disturb the motif, cannot bind to Ori_S. Thus, this helix-turn-helix motif, even if required for sequence-specific binding, is not sufficient alone for binding.

OBP is an interesting protein with 851 amino acids, of which only 269 or fewer are required for sequence-specific recognition of the origin of replication. Direct evidence is not yet available to relate function to sequences outside of the DNA-binding domain of 269 amino acids (positions 564 to 832). Several helicase motifs have been predicted for the OBP sequence on the basis of homology (19). Sequence information suggests the existence of a presumptive ATP-

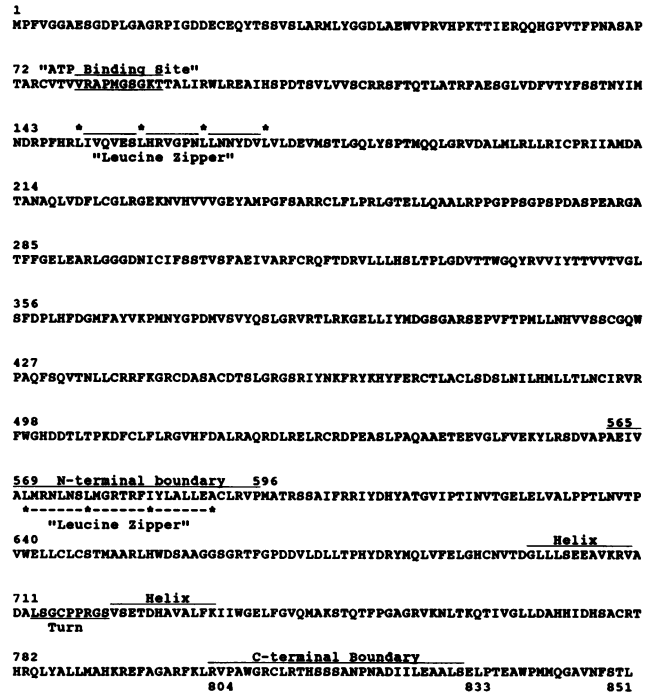


FIG. 10. Amino acid sequence of wild-type OBP. Positions of presumptive leucine zippers, the ATP-binding domain, and the helix-turn-helix motif are indicated. Also indicated are the N- and C-terminal boundaries of the sequence-specific DNA-binding domain identified in this report.

binding site VRAPMGSKRT (25, 32) which may be involved in the ATPase/helicase function. Sequence analysis also reveals a leucine zipper from amino acids 150 to 173 which may be involved in homo- or heterodimer formation (23, 34). Since this leucine zipper is not required for in vitro binding to the origin, it may take part in the formation of a dimer for replication function of the OBP or it could function in heterodimer formation with other replication proteins.

Examination of sequences for other replication protein genes reveals the following: in UL5, a leucine stretch from amino acids 160 to 237 and an ATP-binding site (ITGNAGS GKSTCVQTI); for UL30 (DNA polymerase), a leucine stretch from amino acids 427 to 540; and for UL29 (ICP8), a leucine stretch from amino acids 1076 to 1109. The implication of the presence of leucine stretches in these coding frames is not yet clear. One distinct possibility is that they imply homo- or heterodimerization/DNA-binding domains important for their biological functions, e.g., in the formation of a specialized nucleoprotein structure, a hypothesis that needs to be tested. The ATP-binding site in the UL5 coding sequence may be indicative of the helicase function of the UL5-UL8-UL52 complex (5a, 9, 13, 25, 36). From the results described here and elsewhere (35), it appears the leucine stretch from positions 150 to 179 is not required for DNA binding. Thus, either dimerization is not critical for the sequence-specific DNA-binding activity or the leucine stretch is not involved in this dimerization. Interestingly, however, there is a sequence from amino acids 570 to 591, LMRNLNSLMGRTRFIYLALLEA, which is somewhat similar to the classical leucine zipper. It should be noted that there are reports of leucine zipper motifs that have replacements in some leucine positions (6). Although a deletion

upstream of this segment to amino acid 563 (in del 1-534 rep 542-564) does not affect Ori_S binding significantly, a deletion up to amino acid 596 (in del 131-596) destroys the binding activity. These data signify that sequences between amino acids 563 to 596, a region that includes the pseudo-leucine zipper, are crucial for DNA binding. This result, however, does not directly imply the motif's involvement in dimerization. It is still possible, however, that dimerization is not required for DNA binding or that del 1-534 dimerizes in a manner that is different from that of intact OBP. Nevertheless, the fact that a pseudo-leucine zipper sequence is required for sequence-specific DNA binding is interesting and requires further study. The four-amino-acid insertion (Arg-Ile-Arg-Ala) at the junction of this pseudo-leucine zipper and the rest of the molecule toward the C terminus destroyed DNA binding. This observation strengthens our hypothesis that the pseudo-leucine zipper is crucial for DNA binding. Thus, our results predict that the DNA-binding domain of OBP is composed of two known DNA-binding/dimerization domains, a pseudo-leucine zipper, and a helix-turn-helix motif. Insertions of four amino acids are not tolerated in the DNA-binding domain, suggesting that integrity of the 269-amino-acid stretch is necessary for successful Ori_S binding.

ACKNOWLEDGMENTS

We thank Charles Gauntt and David Martin for critically reading the manuscript. We are indebted to Joel Baseman and Anna Lazzell for their generous help in preparing the antisera. We thank Jane Tate for excellent typing assistance.

This work was supported by grants from the American Cancer Society (MV-420) and the Elsa U. Pardee Foundation and by a Basil O'Connor Starter Scholar Research Award from the March of Dimes to Sumitra Deb. This work was also supported by an American Cancer Society institutional grant to Sumitra Deb and an institutional research grant to Swati Palit Deb.

REFERENCES

1. **Baumann, R. P., V. R. R. Yalamanchili, and D. J. O'Callaghan.** 1989. Functional mapping and DNA sequence of an equine herpesvirus 1 origin of replication. J. Virol. **63**:1275–1283.
2. **Bruckner, R. C., J. J. Crute, M. S. Dodson, and I. R. Lehman.** 1991. The herpes simplex virus 1 origin binding protein: a DNA helicase. J. Biol. Chem. **266**:2669–2674.
3. **Carmichael, E. P., M. J. Kosovsky, and S. K. Weller.** 1988. Isolation and characterization of herpes simplex virus type 1 host range mutants defective in viral DNA synthesis. J. Virol. **62**:91–99.
4. **Carmichael, E. P., and S. K. Weller.** 1989. Herpes simplex virus type 1 DNA synthesis requires the product of the UL-8 gene: isolation and characterization of the ICP6::*lacZ* insertion mutation. J. Virol. **63**:591–599.
5. **Challberg, M. D.** 1986. A method for identifying the viral genes required for herpesvirus DNA replication. Proc. Natl. Acad. Sci. USA **83**:9094–9098.

5a. **Challberg, M. D.** Personal communication.

6. **Chang, Y.-N., D. L.-Y., Dong, G. S. Hayward, and S. D. Hayward.** 1990. The Epstein-Barr virus Zta transactivator: a member of the bZIP family with unique DNA-binding specificity and a dimerization domain that lacks the characteristic heptad leucine zipper motif. J. Virol. **64**:3358–3369.
7. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of secondary structure of proteins from their amino acid sequences. Adv. Enzymol. **47**:45–50.
8. **Crute, J. J., and I. R. Lehman.** 1989. Herpes simplex-1 DNA polymerase. Identification of an intrinsic 5'-3' exonuclease with ribonuclease H activity. J. Biol. Chem. **264**:19266–19270.
9. **Crute, J. J., T. Tsurumi, L. Zhu, S. K. Weller, P. D. Olivo, M. D. Challberg, E. S. Mocarski, and I. R. Lehman.** 1989. Herpes simplex virus 1 helicase-primase: a complex of three herpes-encoded gene products. Proc. Natl. Acad. Sci. USA **86**:2186–2189.
10. **Deb, S., and S. P. Deb.** 1989. Analysis of Ori-S sequence of HSV-1: identification of one functional DNA binding domain. Nucleic Acids Res. **17**:2733–2745.
11. **Deb, S., and M. Doelberg.** 1988. A 67-base-pair segment from the Ori-S region of herpes simplex virus type 1 encodes origin function. J. Virol. **62**:2516–2519.
12. **Digard, P., and D. M. Coen.** 1990. A novel functional domain of an alpha-like DNA polymerase. The DNA binding site on the herpes simplex virus polymerase for the viral UL42 protein. J. Biol. Chem. **265**:17393–17396.
13. **Dodson, M. S., J. J. Crute, R. C. Bruckner, and I. R. Lehman.** 1990. Overexpression and assembly of the herpes simplex virus type 1 helicase-primase in insect cells. J. Biol. Chem. **264**: 20835–20838.
14. **Elias, P., C. M. Gustafsson, and O. Hammarsten.** 1990. The origin binding protein of herpes simplex virus 1 binds cooperatively to the viral origin of replication ori_s. J. Biol. Chem. **265**:17167–17173.
15. **Elias, P., and I. R. Lehman.** 1988. Interaction of origin binding protein with an origin of replication of herpes simplex virus 1. Proc. Natl. Acad. Sci. USA **85**:2959–2963.
16. **Elias, P., M. E. O'Donnell, E. S. Mocarski, and I. R. Lehman.** 1986. A DNA binding protein specific for an origin of replication of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **83**:6322–6326.
17. **Gallo, M. L., D. I. Dorsky, C. S. Crumpacker, and D. S. Parris.** 1989. The essential 65-kilodalton DNA-binding protein of herpes simplex virus stimulates the virus encoded DNA polymerase. J. Virol. **63**:5023–5029.
18. **Goldstein, D. J., and S. K. Weller.** 1988. An ICP6::*lacZ* insertional mutagen is used to demonstrate that the UL52 gene of herpes simplex virus type 1 is required for virus growth and DNA synthesis. J. Virol. **62**:2970–2977.
19. **Gorbalenya, A. E., E. V. Konin, A. P. Donchenko, and V. M. Blinov.** 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res. **17**:4713–4730.
20. **Koff, A., and P. Tegtmeyer.** 1988. Characterization of major recognition sequences for a herpes simplex virus type 1 origin-binding protein. J. Virol. **62**:4096–4103.

20a. **Koff, A., and P. Tegtmeyer.** Personal communication.

21. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. **157**:105–132.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **224**:680–685.
23. **Landschulz, W. H., P. F. Johnson, and S. McKnight.** 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Nature (London) **243**:1681–1688.
24. **McGeoch, D. J., M. A. Dalrymple, A. J. Davidson, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69**:1531–1574.
25. **McGeoch, D. J., M. A. Dalrymple, A. Dolan, D. McNab, L. J. Perry, P. Taylor, and M. D. Challberg.** 1988. Structures of the herpes simplex virus type 1 genes required for replication of virus DNA. J. Virol. **62**:444–453.

25a. **Natajaran, M., and S. Deb.** Unpublished data.

26. **Olivo, P. D., and M. D. Challberg.** 1988. Herpes simplex virus DNA replication: identification of the essential genes and their products. Cancer Cells **6**:43–51.
27. **Olivo, P. D., N. J. Nelson, and M. D. Challberg.** 1988. Herpes simplex virus DNA replication: the UL9 gene encodes an origin binding protein. Proc. Natl. Acad. Sci. USA **85**:5414–5418.
28. **Olivo, P. D., N. J. Nelson, and M. D. Challberg.** 1989. Herpes simplex virus type 1 gene products required for DNA replication: identification and overexpression. J. Virol. **63**:196–204.
29. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition.

Annu. Rev. Biochem. **53**:293–321.

30. **Preston, C. M., C. M. Frame, and M. E. M. Campbell.** 1988. A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. Cell **52**:425–434.
31. **Roizman, B., and A. E. Sears.** 1990. Herpes simplex viruses and their replication. *In* B. N. Fields et al. (ed.), Virology, vol. 2. Raven Press, New York.
32. **Sekimizu, K., D. Bramhill, and A. Kornberg.** 1987. ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. Cell **50**:259–265.
33. **Stow, N. D., and A. J. Davison.** 1986. Identification of a varicella-zoster virus origin of DNA replication and its activation by herpes simplex virus type 1 gene products. J. Gen. Virol. **67**:1613–1623.
34. **Turner, R., and R. Tjian.** 1989. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. Nature (London) **243**:1689–1694.
35. **Weir, H. M., J. M. Calder, and N. D. Stow.** 1989. Binding of the herpes simplex virus type 1 UL9 gene product to an origin of viral DNA replication. Nucleic Acids Res. **17**:1409–1425.
36. **Weller, S.** Personal communication.