JOURNAL OF VIROLOGY, Mar. 1991, p. 1149–1159
0022-538X/91/031149-11$02.00/0

Vol. 65, No. 3

The Conserved DNA-Binding Domains Encoded by the Herpes Simplex Virus Type 1 ICP4, Pseudorabies Virus IE180, and Varicella-Zoster Virus ORF62 Genes Recognize Similar Sites in the Corresponding Promoters

CHIN-LEE WU AND KENT W. WILCOX*

Department of Microbiology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226

Received 10 September 1990/Accepted 30 November 1990

Herpes simplex virus types 1 and 2 (HSV-1 and HSV-2), pseudorabies virus (PRV), varicella-zoster virus (VZV), and equine herpesvirus 1 (EHV-1) are all classified as *Alphaherpesvirinae*. Each of these five viruses encodes an essential immediate-early (IE) regulatory protein referred to as HSV-1 ICP4, HSV-2 ICP4, PRV IE180, VZV ORF62 protein, and EHV-1 IE1, respectively. These five proteins share extensive homology with each other in domains referred to as regions 2 and 4. The HSV-1 ICP4 region 2 domain contains residues that are required for the DNA-binding capability of ICP4. In this report, we describe the expression of region 2 domains from the ICP4, IE180, and ORF62 genes as fusion proteins in *Escherichia coli*. DNA-binding assays revealed that each of these region 2 fusion proteins binds to a sequence that overlaps the transcription start site in the promoter for the gene encoding the corresponding protein. Each of the sites with high affinity for one or more of these fusion proteins contains the sequence 5′-ATCGT-3′. This sequence spans the mRNA cap site in the HSV-2 ICP4 gene promoter and is immediately upstream from the transcription start site in the EHV-1 IE1 gene. These results suggest that formation of a specific complex between an IE protein and its own gene promoter may be a common mechanism used by *Alphaherpesvirinae* to autoregulate transcription of an essential IE gene.

The family *Herpesviridae* is subdivided into three subfamilies called the *Alpha-*, *Beta-*, and *Gammaherpesvirinae*. Assignment of a virus to a given subfamily is determined by properties of the virus particle and features of the viral reproductive process, including aspects such as host range and site of latency (50). Included among the *Alphaherpesvirinae* are herpes simplex virus types 1 and 2 (HSV-1 and -2), varicella-zoster virus (VZV), pseudorabies virus (PRV), and equine herpesvirus 1 (EHV-1). Analyses of DNA sequence data have revealed that the genomes from these five *Alphaherpesvirinae* are essentially colinear with respect to genetic information (7, 8, 32, 34, 37). In addition to the similarities among these five viruses at the genomic level, evidence is accumulating which indicates that a similar pattern of gene expression occurs during productive infection of cells by each of these viruses (2, 5, 19).

Productive infection of cells by members of the *Alphaherpesvirinae* subfamily is typified by a short (18 to 30 h) reproductive cycle that can be subdivided into three temporal phases (designated immediate-early [IE], early, and late) that were originally delineated by significant changes in the pattern of viral protein synthesis (16, 25). Much of our basic understanding of the mechanisms which control these temporal shifts in alphaherpesvirus gene expression has been derived from experiments with HSV. Several different approaches have demonstrated that the shift from the IE phase to the early phase of HSV infection is mediated in part by a virus-encoded IE protein designated ICP4 (α4, IE175, Vmw175) which induces transcription of early and late viral genes and represses transcription of IE genes (9, 14, 21, 40, 46, 49, 57). Two other viral IE proteins, designated ICP0 (α0, IE110, Vmw110) and ICP27 (α27, IE63, Vmw63), also play important roles in the regulatory processes which result in differential viral gene expression during the shifts to the early and late phases of productive infection (15, 23, 31, 47).

Native ICP4 is a sequence-specific DNA-binding phosphoprotein that exists in solution as a homodimer (18, 27, 29, 36, 45). It is found in the nuclei of productively infected cells (28) and in the tegument of mature virions (60). Although denatured ICP4 migrates on sodium dodecyl sulfate (SDS)-polyacrylamide gels with an apparent mass of 175 kDa, conceptual translation of the ICP4 gene sequence predicts a polypeptide with 1,298 amino acid residues and a mass of 133 kDa (33). Comparison of the published sequence data for the VZV, PRV, and EHV-1 genomes revealed that each of these viruses encodes a protein that is similar in size to and shares extensive homology with ICP4. The VZV homolog of ICP4 is encoded by open reading frame 62 (ORF62), which predicts a protein of 1,310 residues (7). The VZV ORF62 gene product is probably identical to an IE phosphoprotein with an apparent molecular mass of 175 to 180 kDa that was reported to be in both the nuclei of VZV-infected cells and the tegument of VZV virions and which apparently binds to both single-stranded and double-stranded DNA (20, 27a, 48, 52). Cells that produce the ORF62 protein are permissive hosts for HSV strains which contain conditional or lethal mutations in the ICP4 gene (20). The PRV homolog of ICP4 is encoded by an open reading frame of 1,446 to 1,460 codons, depending on the viral strain, and is referred to as PRV IEP or IE180 (3, 26, 56). PRV IE180 is a phosphoprotein that binds nonspecifically to single-stranded DNA and to specific sites in double-stranded DNA (4, 6). The EHV-1 homolog of ICP4 is a nuclear phosphoprotein referred to as EHV-1 IE1 which is predicted to contain 1,487 residues (22, 24).

On the basis of the predicted amino acid sequences for the

* Corresponding author.

VZV, PRV, and EHV-1 homologs of HSV-1 ICP4, the primary structures of these proteins have been conceptually divided into five regions that are colinear among the proteins (3, 22, 33, 56). The four proteins share extensive (approximately 50%) amino acid identity within regions 2 and 4 and have little or no homology within regions 1, 3, and 5 except for a serine-rich tract found in region 1 of all four proteins. Despite the differences among these proteins in regions 1, 3, and 5, they apparently perform similar functions during productive infection. It has been reported that ICP4, PRV IE180, and the VZV ORF62 protein transactivate expression of early viral genes and apparently act as negative regulators of their own gene expression (16, 20, 26). The molecular mechanisms by which these IE proteins regulate viral gene expression have not been fully elucidated.

In the case of ICP4, the negative regulation of ICP4 gene expression is apparently mediated at least in part by formation of a complex between ICP4 and a specific site in the ICP4 gene that spans the region from −8 to +20 relative to the transcription initiation site (+1) (18, 38, 39, 44, 49). This complex is presumed to down-regulate transcription either by sterically blocking movement of RNA polymerase II or by displacing an essential transcription factor such as TFIID (12, 18, 30). The regulatory role of specific components of this complex have been investigated by mutational analysis of both the ICP4 promoter and the ICP4 coding region. A mutant ICP4 promoter with greatly reduced affinity for native ICP4 was not subject to ICP4 autoregulation in transient transfection assays (49). Variants of ICP4 with mutations in region 2 that had reduced affinity for DNA also failed to autoregulate transcription from the ICP4 promoter (10, 11, 14, 42–44, 51). Residues outside region 2 must play some role in the autoregulatory process, since a mutation in the carboxy-terminal third of the ICP4 gene yields a protein that binds to DNA but fails to autoregulate (44).

We have previously reported that a 240-amino-acid polypeptide that spans region 2 of ICP4 is capable of binding to the transcription initiation site in the ICP4 gene (59). Given that both the VZV ORF62 protein and PRV IE180 are DNA-binding proteins that autoregulate their own gene expression and share extensive homology with region 2 of ICP4, we hypothesized that the region 2 domains from VZV ORF62 and PRV IE180 would bind specifically to a site at or near the transcription initiation site in the corresponding gene. To test this hypothesis, we expressed the region 2 domain from each of these proteins in *Escherichia coli* and conducted DNA-binding experiments with fragments spanning the putative promoter from the corresponding genes.

MATERIALS AND METHODS

Construction of recombinant *trpE*-region 2 genes. The pATH expression vectors were generous gifts of Carol Dieckmann (13) and contain a portion of the *E. coli* tryptophan operon extending from several hundred nucleotides upstream of the promoter to codon 323 in the *trpE* gene, followed immediately by derivatives of the pUC12 multiple cloning region (MCR). Plasmids containing the wild-type HSV-1 ICP4, VZV ORF62, and PRV IE180 genes were obtained from Nigel Stow, Jeff Ostrove, and Andrew Cheung, respectively (3, 20, 55). The construction of plasmid pXK449 has been described (59). It encodes a 561-residue fusion protein designated FP449 that contains the first 323 residues from the *trpE* gene, 8 residues (PGDPLEST) from the MCR, and ICP4 residues 262 to 490.

Plasmid pXK617 was constructed by isolating a *Cla*I-to-*Bst*EII fragment that contains VZV ORF62 codons 417 to 647 and inserting it between the *Eco*RI and *Cla*I sites in pATH1. pXK617 encodes a 561-residue fusion protein designated FP617 that contains the first 323 residues from *trpE*, 6 residues (PGRARI) from the MCR, VZV ORF62 residues 417 to 647, and 1 residue (A) from the MCR at the carboxy terminus.

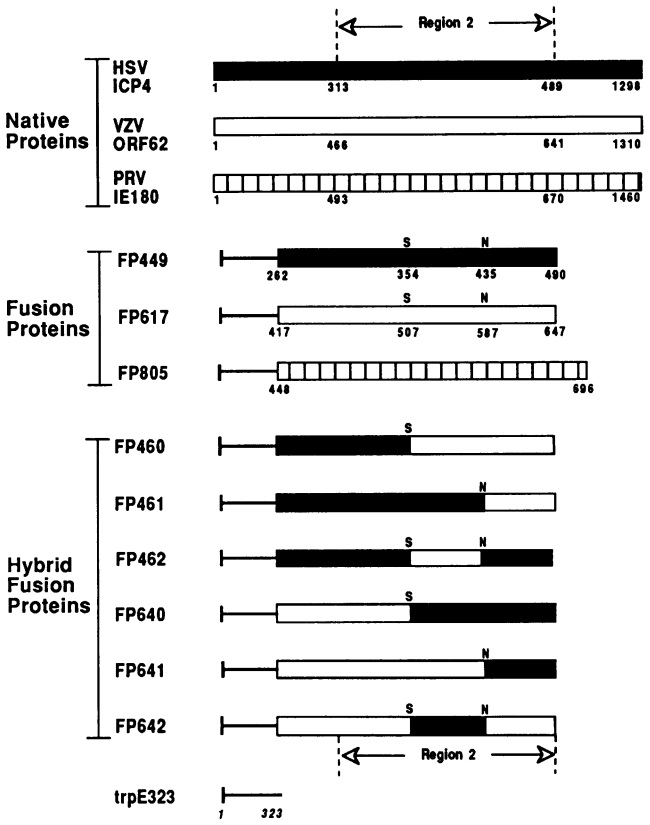

FIG. 1. Schematic diagram of recombinant *trpE*-region 2 fusion proteins. Symbols: ■, amino acid residues encoded by the HSV-1 ICP4 gene; □, residues encoded by the VZV ORF62 gene; ▥, residues encoded by the PRV IE180 gene; —, the *trpE* portion of each fusion protein. Numbers below the boxes indicate the positions of specific residues with respect to the amino terminus of the native viral protein. Portions of each schematic drawing outside region 2 are not drawn to scale. *Stu*I (S) and *Nco*I (N) sites used to construct hybrid genes are indicated.

Plasmid pXK805 was constructed by isolating an *Sfi*I-to-*Alu*I fragment that contains PRV IE180 codons 453 to 696 and linking it to the expression vector via a double-stranded oligonucleotide that restored PRV codons 448 to 452. pXK805 encodes a 579-residue fusion protein designated FP805 that contains the first 323 residues from *trpE*, 6 residues (PPNSGG) from the MCR, IE180 residues 448 to 696, and 1 residue (G) from the MCR at the carboxy terminus.

Chimeric genes were constructed by swapping DNA fragments between pXK449 and pXK617. These fragments were generated by cleaving pXK449 and pXK617 at either the *Stu*I or *Nco*I site (Fig. 1). Because the cleavage sites for both *Stu*I and *Nco*I are located at the same relative position (with respect to region 2) in the ICP4 and ORF62 genes, there are no changes in either codon usage or reading frame at the ligation site.

Preparation of bacterial extracts. *E. coli* cultures were grown and the *trp* promoter was induced as previously described (59). All of the fusion proteins except FP640 were partially purified by ammonium sulfate precipitation of an extract derived from cells disrupted in buffer A (25 mM Tris hydrochloride [pH 8.0], 10 mM EDTA) as previously described (59). FP640 was insoluble in buffer A. For preparation of the FP640 extract, cells from a 200-ml induced culture were harvested by centrifugation, resuspended in 3 ml of cold buffer A, and disrupted by sonication at 4°C. After centrifugation, the pellet was resuspended in 3 ml of buffer A supplemented with 4 M urea. This suspension was centrifuged at 12,000 × *g* for 10 min, and the supernatant was dialyzed at 4°C against 100 ml of buffer A with 1 M urea for 15 h, 100 ml of buffer A with 0.3 M urea for 3 h, and then 500 ml of buffer B (10 mM Tris hydrochloride [pH 8.0], 1 mM EDTA, 10 mM β-mercaptoethanol) containing 0.01% CHAPS {3-[(3-cholamidopropyl)-dimethyl-ammonio]-1-propanesulfonate} for 4 h. The solution was clarified by centrifugation, mixed with an equal volume of glycerol, and stored at −20°C.

DNA-binding assays. Targets for DNA-binding assays were isolated from derivatives of pUC13 and pUC19 that contain portions of the promoters for the genes encoding HSV-1 ICP4, VZV ORF62, and PRV IE180. pXK26 contains the *Eco*RI-to-*Sal*I fragment spanning the region from −107 to +183 relative to the cap site (+1) in the HSV-1 ICP4 gene. pXK626 contains the *Sal*I-to-*Afl*III fragment spanning the region from −409 to −242 in the ORF62 gene. pXK606 contains the *Nde*I-to-*Sca*I fragment spanning the region from −262 to +21 in the ORF62 gene. pXK607 contains the *Xho*I-to-*Eco*47III fragment spanning the region from −130 to +95 in the ORF62 gene. pXK826 contains the *Sph*I-to-*Bam*HI fragment spanning the region from −154 to +51 in the IE180 gene. The appropriate fragments were isolated from gels and end labeled with [γ-^{32}P]ATP to 10,000 to 50,000 cpm/ng (59). DNase I footprinting assays were conducted as described before (59). Briefly, 4 ng of radiolabeled DNA and 15 μl of extract from induced *E. coli* were incubated at 22°C for 30 min in buffer B containing 100 mM NaCl. The reaction was then sequentially incubated for 1 min at 22°C with DNase I (0.25 U; Pharmacia) and for 20 min at 50°C with protease K (10 μg). The DNA was then extracted and electrophoresed on 8% polyacrylamide–8 M urea sequencing gels.

RESULTS

Construction and expression of recombinant *trpE*-region 2 genes. We previously reported that expression in *E. coli* of a recombinant gene which contains 323 codons from the *E. coli trpE* gene fused in frame to codons 262 to 490 from the ICP4 gene resulted in production of a fusion protein designated FP449 which exhibits the same DNA-binding specificity as native ICP4 (59). In contrast, no DNA-binding activity was detected for a smaller fusion protein that contained ICP4 residues 306 to 490 (59). The portion of ICP4 that has been defined as region 2 on the basis of homology with the VZV ORF62 and PRV IE180 proteins extends from amino acid residues 313 to 489. Our previous results suggest that formation of a polypeptide with high affinity for specific sites in DNA also requires approximately 40 residues extending upstream from the amino-proximal border of region 2 which are not conserved in these related IE proteins. Although we do not understand the role of these additional residues, we decided that it would be prudent to include this nonconserved portion when constructing recombinant genes that express the region 2 domains from the VZV ORF62 and PRV IE180 proteins.

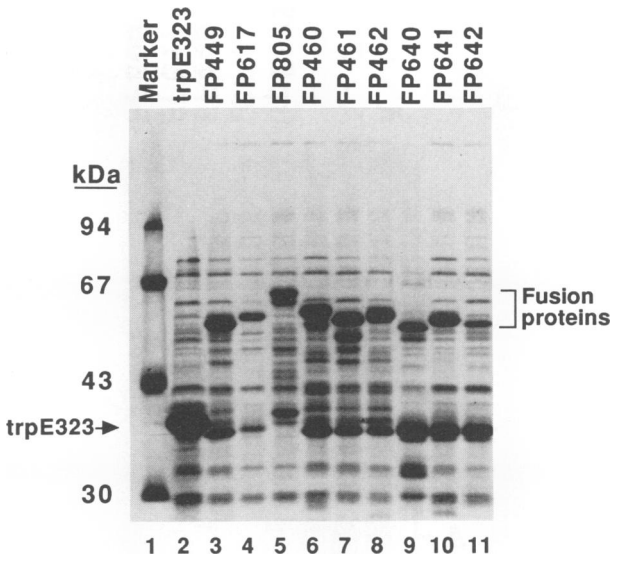

FIG. 2. Analysis of fusion proteins present in *E. coli* extracts. Extracts from cells expressing the indicated proteins were electrophoresed through a 12% SDS-polyacrylamide gel and stained with Coomassie blue. Molecular mass markers (Pharmacia) were loaded in lane 1. The 36-kDa *trpE323* protein encoded by vector pATH2 is indicated by an arrow in lane 2. The *trpE*-region 2 fusion proteins are heavy bands which migrate within the area indicated by a bracket.

We have expressed portions of the ICP4 gene in *E. coli* as both fusion and nonfusion polypeptides and have consistently found that yields are highest when the ICP4 portion is fused to the first 323 residues from the *trpE* protein. Furthermore, the *trpE323* portion has no apparent effect on the DNA-binding properties of the hybrid proteins and provides a convenient tag for immunodetection of various *trpE323* fusion proteins (59). For these reasons, we decided to express portions of the VZV and PRV IE genes as *trpE323* fusion proteins.

The corresponding region 2 domains in VZV ORF62 and PRV IE180 extend from residues 466 to 641 and 493 to 670, respectively. On the basis of the availability of convenient restriction enzyme sites, DNA fragments containing ORF62 codons 417 to 647 and IE180 codons 448 to 696 were fused to the *E. coli trpE* gene in an MCR located 323 codons downstream from the *trpE* translation start codon (Fig. 1). The fusion proteins encoded by the VZV and PRV trpE-region 2 recombinant genes were designated FP617 and FP805, respectively.

Densitometric analysis of stained gels indicated that the relative yields in soluble extracts from *E. coli* of FP449, FP617, and FP805 were approximately 1.0, 0.3, and 0.6 (Fig. 2). These estimates do not include the numerous proteolytic fragments that were observed in some extracts. Using bovine serum albumin (BSA) as a standard on the stained gels, we estimated that the FP449 extract contained 1.3 μg of FP449 per μl.

DNA-binding specificities of region 2 fusion proteins. To test our hypothesis that the VZV ORF62 protein and PRV IE180 bind specifically at or near the transcription initiation site in the promoter for the corresponding gene, we performed DNase I footprinting assays with the region 2 fusion

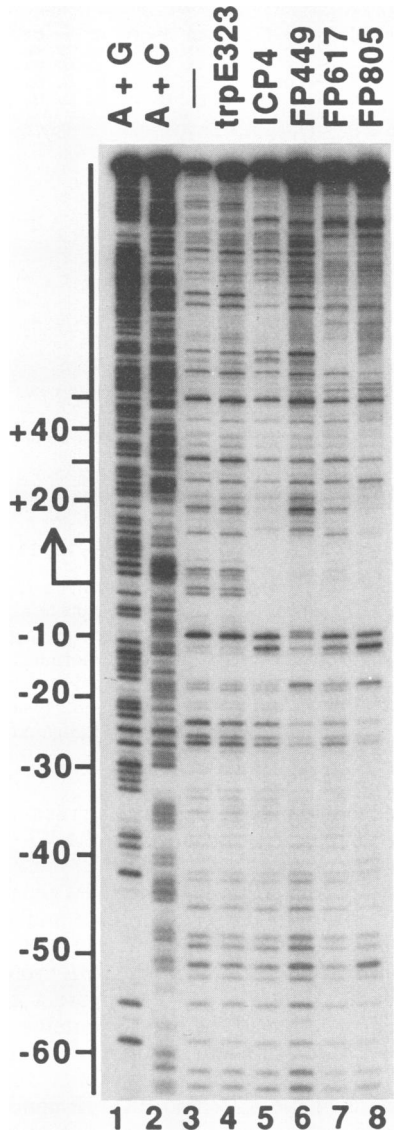

FIG. 3. DNase I footprints on the HSV-1 ICP4 gene promoter. The target DNA was a 301-bp *Hin*dIII-*Eco*RI fragment from pXK26 that contains the ICP4 gene promoter from −107 to +183. The 5′ phosphate at the *Eco*RI site (−107) was radiolabeled. DNA-binding and DNase I footprinting assays were conducted as described in Materials and Methods. The locations of nucleotides relative to the mRNA cap site (arrow) are indicated. Binding reactions contained either no extract (lane 3), *E. coli* extracts as indicated (lanes 4, 6, 7, and 8), or purified native ICP4 (lane 5).

proteins produced in *E. coli* and target DNAs derived from the promoters for these IE genes. Native ICP4, FP449, and the HSV-1 ICP4 gene promoter were included for comparative purposes.

The target DNA representing the ICP4 promoter extends from −107 to +183 relative to the mRNA cap site. In agreement with previous results (18), we observed that both native ICP4 and FP449 provided strong protection for the region from −8 to +12 in the ICP4 promoter (Fig. 3, lanes 5 and 6). Interestingly, the region 2 fusion proteins from VZV ORF62 (FP617) and PRV IE180 (FP805) protected essentially the same nucleotides as did native ICP4 and FP449 (Fig. 3, lanes 7 and 8). These results indicate that FP617 and FP805 are sequence-specific DNA-binding proteins which recognize sequences similar to those recognized by FP449 and native ICP4. We also observed additional sites in this target that were weakly protected by either native ICP4, FP617, or FP805. The differences in the patterns of secondary binding sites may reflect the fact that the region 2 domains of these IE proteins are only 50% identical.

To test binding to the VZV ORF62 promoter, three overlapping DNA fragments extending approximately 500 bp upstream from the translation start codon in the VZV ORF62 gene were used as targets (Fig. 4E). These three fragments span the region from −409 to +95 relative to the ORF62 mRNA cap site described by McKee et al. (35) and include a potential ORF62 transcription initiation site proposed by Felser et al. (20; see Discussion) that is located at −287 on the map shown in Fig. 4E.

When these three fragments were used as targets in DNase footprinting assays, native ICP4 and FP449 provided only weak protection at a few sites (Fig. 4A to D). In contrast, both FP617 and FP805 provided strong protection for a region from −311 to −282 (Fig. 4A), which spans the transcription initiation site proposed by Felser et al. (20), and provided moderate to weak protection for additional sites between −190 to +95, including the transcription initiation site described by McKee et al. (35) (Fig. 4B to D). As observed with the ICP4 gene promoter, similar regions were protected by both FP617 and FP805, but the degree of protection was in general greater with FP805. A summary of the regions in the ORF62 gene promoter that were protected by ICP4, FP449, FP617, and FP805 is shown schematically in Fig. 4E. Because a protein in the control extract (*trpE323*) protected nucleotides from −285 to −260 (Fig. 4A), a probe that terminated at position −259 and that was not protected by the control extract was used to demonstrate that protection by FP617 and FP805 extended to position −282 (data not shown).

The target DNA representing the PRV IE180 promoter extends from −154 to +51 relative to the mRNA cap site (1). When this fragment was used for DNase footprinting assays, native ICP4 and FP449 provided partial protection from −20 to +1 and from +20 to +35 (Fig. 5, lanes 5 and 6). In contrast, both FP617 and FP805 provided strong protection for two sites located from −20 to +10 and from +17 to +45; it is noteworthy that one of the protected sites extends across the mRNA cap site (Fig. 5, lanes 7 and 8).

DNA-binding affinities of region 2 fusion proteins. Although the DNase footprinting assays provide strong evidence that the region 2 fusion proteins are sequence-specific DNA-binding proteins, these assays do not give a clear indication of the relative affinities of these proteins for specific sites in DNA. Therefore, we compared the relative affinities of these fusion proteins for representative promoter DNA fragments by DNA mobility shift assays. FP449 yielded well-resolved protein-DNA complexes with each of the targets; the complexes with the ICP4 promoter were the most intense, whereas the complexes with the VZV and PRV IE promoters were weak to barely detectable (Fig. 6). FP617 yielded rather broad smears and some large complexes that failed to enter the gel with three of the targets (Fig. 6A, B, and D). FP617 yielded two fairly sharp bands with the target spanning nucleotides −409 to −242 in the VZV IE promoter (Fig. 6C), which is consistent with the strong footprint that FP617 yielded with this same target DNA (Fig. 4A). FP805 yielded well-resolved complexes with each of the targets; the most

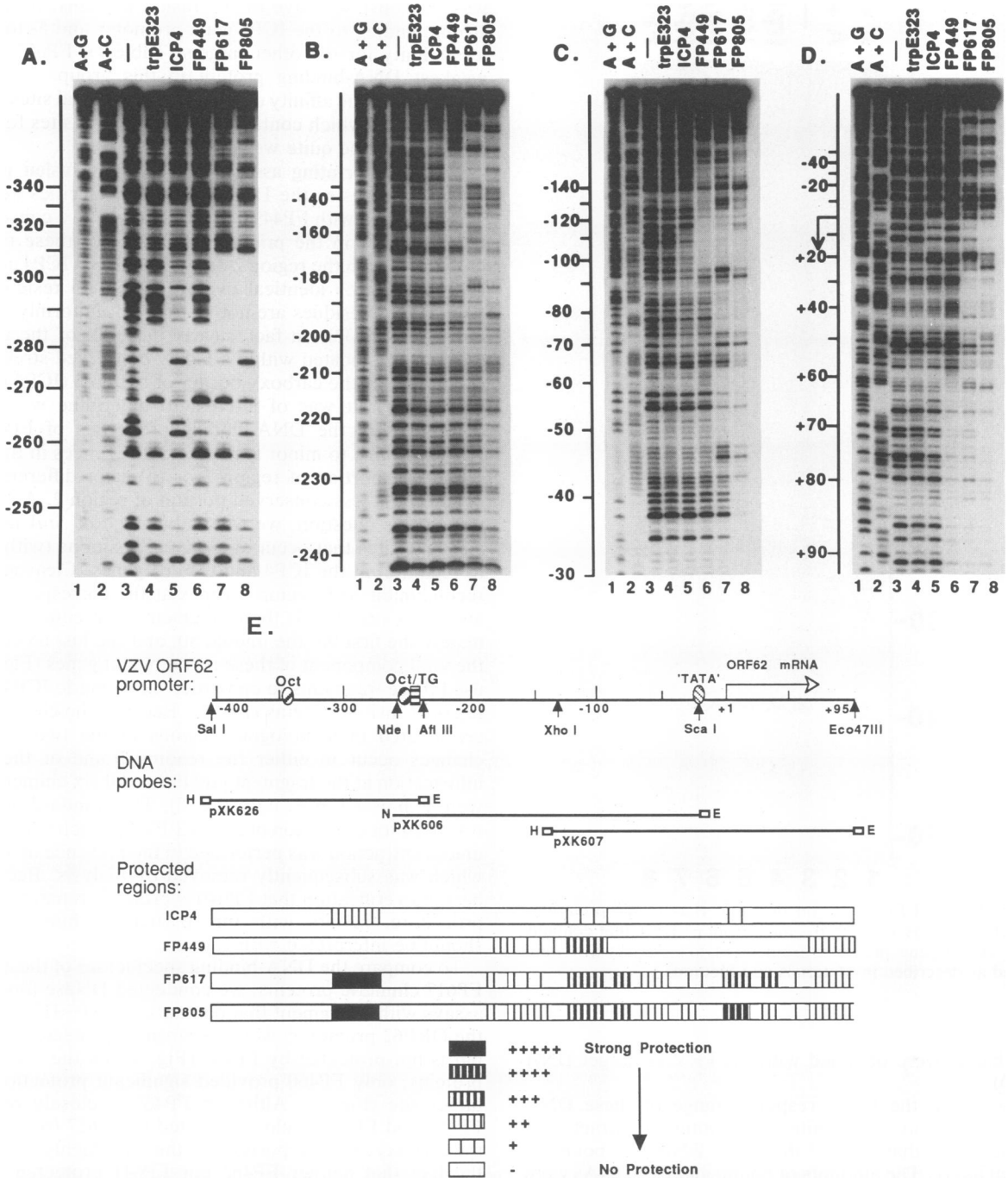

FIG. 4. DNase I footprints on the VZV ORF62 gene promoter. Assays were performed with 5'-end-labeled DNA fragments and bacterial extracts or native ICP4 as described in Materials and Methods. Target DNAs included fragments from pXK626 (A), pXK606 (B and C), and pXK607 (D). In panel E, the locations of restriction enzyme cleavage sites used to generate cloned targets for DNA-binding studies and putative binding sites for transcription factors (Oct, TG, and TATA) relative to the proposed mRNA cap site (35) in the ORF62 promoter are shown schematically on the top line. DNA probes are shown schematically underneath the promoter map; open boxes represent nucleotides derived from vector, and letters indicate cleavage sites at the termini of each target. The lower portion of panel E summarizes in a diagrammatic fashion the footprinting results shown in panels A through D. The open boxes represent the ORF62 promoter drawn to the same scale as in the upper portion of this panel. The density of vertical lines within each box is roughly proportional to the degree of protection by the protein indicated to the left of each box. H, *Hin*dIII; E, *Eco*RI; N, *Nde*I; Oct, octamer; TG, taatgarat.

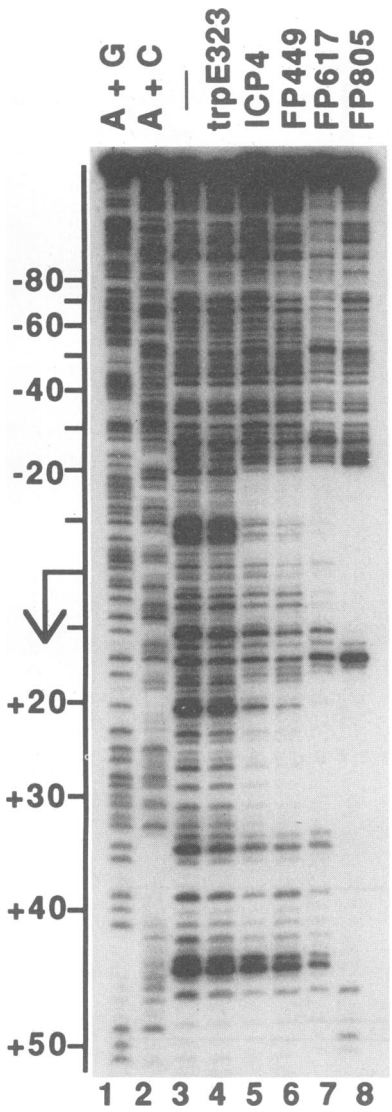

FIG. 5. DNase I footprints on the PRV IE180 gene promoter. The target DNA was a 226-bp fragment from pXK826 that contains the PRV IE180 gene promoter from −154 to +51. Assays were performed as described in Materials and Methods.

intense bands were obtained with the PRV IE target DNA (Fig. 6D).

To be within the linear response range of these DNA mobility shift assays, we limited the amount of extract used so that no more than 80% of the target DNA was bound in any given assay. The amounts of bound and free DNAs were quantitated by scintillation counting. The relative amount of fusion protein in each assay was determined from densitometric scans of stained gels (Fig. 2). The specific activity of each fusion protein was then calculated as femtomoles of DNA bound per staining unit of fusion protein (Table 1). Since the other components of the binding assays remained constant, the specific activities listed in Table 1 should reflect the relative affinities of the fusion proteins for each DNA target. FP805 displayed the highest binding activity of the four proteins (Table 1). It bound to all four targets with higher affinities than did any of the other fusion proteins. The PRV IE180 gene promoter was the best target for FP805 by a factor of 2; this may reflect the presence of two strong binding sites for FP805 in this target (Fig. 5, lane 8). FP449 was the most selective of the fusion proteins; it displayed strong binding to the ICP4 gene promoter and 5- to 15-fold less affinity for the other targets (Table 1). FP617 was the weakest DNA-binding protein in this group. This result suggests that the affinity of FP617 for individual sites in those DNA targets which contain multiple binding sites for FP617 (Fig. 4) must be quite weak.

DNase footprinting assays with chimeric fusion proteins. The differences in the DNA-binding specificities and affinities observed with FP449 and FP617 must be a consequence of differences in the primary structures of these two proteins. Although the region 2 domains of HSV ICP4 and VZV ORF62 are 48% identical overall (84 of 175 residues), the homologous residues are not distributed uniformly throughout region 2 (33). In fact, nearly half (40) of the identical residues are located within a highly conserved stretch of 54 residues near the carboxy border of region 2 (ICP4 residues 422 to 475). It was of interest to determine whether the differences in the DNA-binding properties of FP449 and FP617 are due to minor amino acid differences in the highly conserved portion of region 2 or to major differences that occur in the less conserved portion of region 2. As a partial test of this question, we took advantage of *Stu*I and *Nco*I cleavage sites that occur at identical positions (with respect to region 2) in the ICP4 and ORF62 genes. Cleavage of the recombinant *trpE*-region 2 genes at the corresponding *Stu*I and *Nco*I sites yields three fragments that contain approximately the first 90, the middle 80, and the last 65 codons of the viral component of these recombinant genes (Fig. 1). We used these fragments to construct six chimeric ICP4-ORF62 region 2 fusion proteins (Fig. 1). Because the cleavage sites are located in homologous regions of the two genes, no changes occur in either the reading frame or the coding information at the fragment junctions. All six chimeric genes were expressed in *E. coli* (Fig. 2). The standard procedure was used for extraction of all but FP640, which was insoluble unless extraction was performed in the presence of 4 M urea, which was subsequently removed by dialysis. Because we have no verification that FP640 is properly renatured by this procedure, results with this particular chimeric protein should be interpreted with caution.

To compare the DNA-binding specificities of these FP449-FP617 chimeric proteins, we conducted DNase footprinting assays with a fragment that contains a site (−311 to −282 in the ORF62 promoter) which is strongly protected by FP617 but is not protected by FP449 (Fig. 4). Of the six chimeric proteins, only FP460 provided significant protection of the target site (Fig. 7). Although FP462 is closely related to FP460 and FP641 is closely related to FP617 (residues near the carboxy terminal portion are the most highly conserved), the fact that neither FP462 nor FP641 protected this site suggests that specific residues located on both sides of the *Nco*I site govern the DNA-binding specificities of ICP4 and ORF62 proteins. The fact that FP460 binds DNA in a manner that is nearly indistinguishable from that of FP617 suggests that much of the binding specificity encoded in region 2 is governed by residues downstream from the *Stu*I site. If this latter hypothesis were true, then one would predict that FP640, which is the reciprocal of FP460, would have the same DNA-binding specificity as FP449. Tests of this prediction using the ICP4 gene promoter as the target DNA were negative (data not shown), but we suspect that the

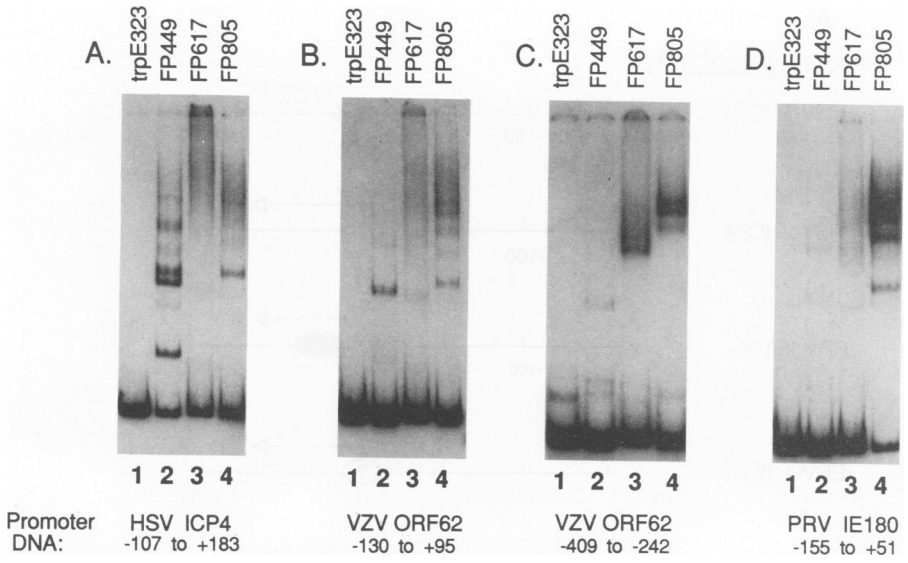

FIG. 6. DNA mobility shift assays. Binding reaction mixtures contained 4 fmol of target DNA, 1 μg of sonicated calf thymus DNA, and bacterial extract as follows: lane 1, *trpE323* (1 μl); lane 2, FP449 (0.1 μl); lane 3, FP617 (1 μl); lane 4, FP805 (0.05 μl). Assays were conducted as described before (59) except that gel electrophoresis was performed in 45 mM Tris-borate buffer plus 1 mM EDTA for 4 h at 150 V. Autoradiograms of the gels are shown. The nucleotide sequences included in each radiolabeled target DNA are indicated at the bottom of each panel and correspond to fragments from pXK26 (A; 301 bp), pXK607 (B; 267 bp), pXK626 (C; 204 bp), and pXK826 (D; 226 bp).

solubility problems with FP640 may have adversely affected this test.

DISCUSSION

Comparing structurally related biological systems has proven to be very useful in elucidating the essential role of highly conserved features of such systems. We have taken this approach to characterize similarities in the recognition sites and DNA-binding domains of three related IE proteins from three different herpesviruses. Although the data presented here are based on the use of truncated portions of these viral proteins expressed in *E. coli*, we believe that similar results, at least to a first approximation, would be obtained with the corresponding native protein. Given this assumption, one can then address the question that was raised in the introduction: do the region 2 domains in the ICP4 analogs encoded by PRV, VZV, and EHV-1 recognize a specific sequence of conserved nucleotides at or near the transcription start site in the corresponding gene? A partial answer to this question is provided by comparison of the sequences and locations of nucleotides strongly protected by

TABLE 1. Specific activities of region 2 fusion proteins

Fusion protein	Sp act with given promoter DNA[a]			
	HSV ICP4, −107 to +183	VZV ORF62, −130 to +95	VZV ORF62, −409 to −242	PRV IE180, −155 to +51
FP449	20.3	1.6	4.1	1.6
FP617	4.7	3.0	2.5	1.9
FP805	39.5	39.0	41.4	81.9

[a] Expressed as femtomoles of target DNA bound per staining unit of fusion protein. One staining unit yields the same intensity as 1 μg of BSA on an SDS-polyacrylamide gel stained with Coomassie blue. Bound DNA was determined by counting gel slices from mobility shift assays; background values contributed by *E. coli* proteins were determined in control lanes and subtracted.

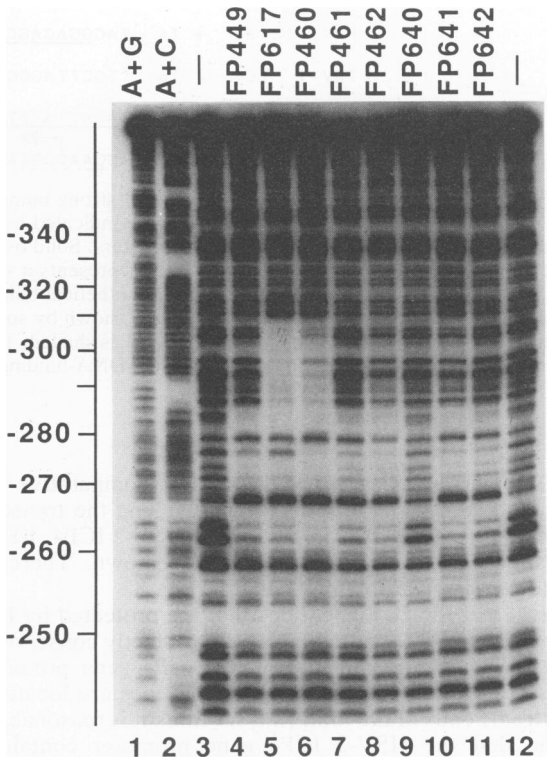

FIG. 7. DNase I footprints with chimeric fusion proteins. The target DNA was a radiolabeled fragment from pXK626 which includes the VZV ORF62 gene promoter from −409 to −242. Reactions were conducted as described in Materials and Methods with extracts containing the indicated proteins.

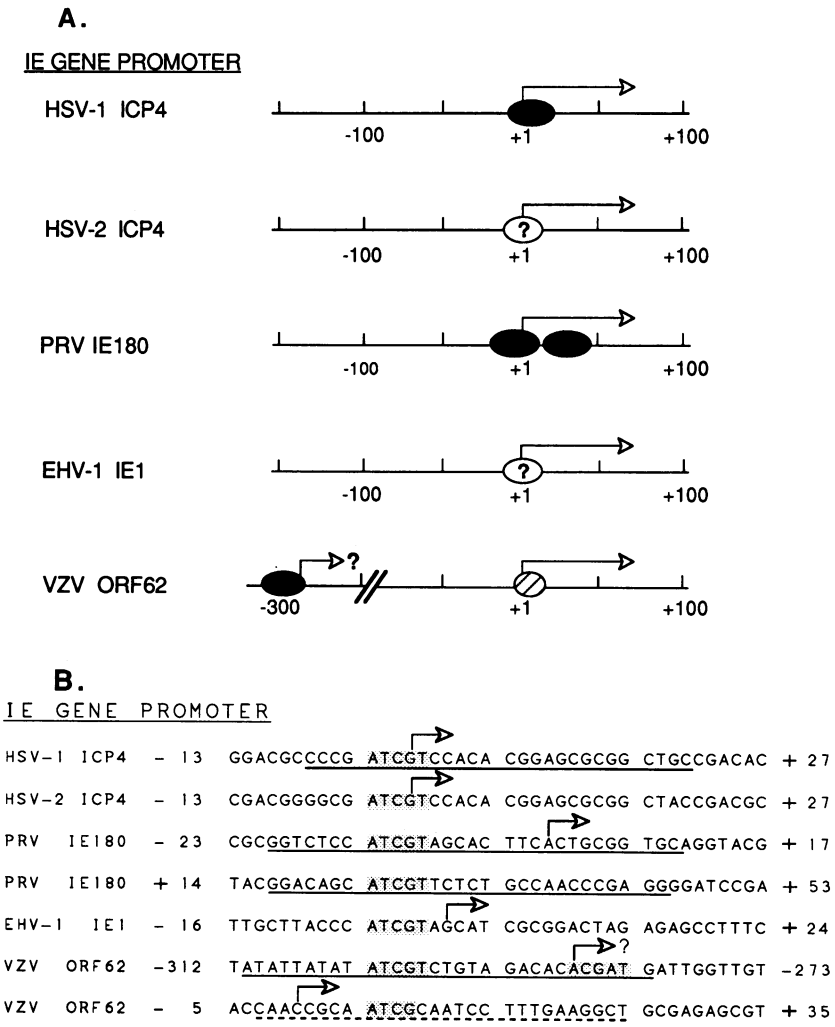

FIG. 8. Locations and nucleotide sequences of strong binding sites. The diagram summarizes the DNase I footprinting results shown in Fig. 3 through 5. Transcription initiation sites are indicated by arrows. An alternative start site proposed for the VZV ORF62 gene is also shown. (A) Schematic diagrams of IE gene promoters. Solid ovals represent sites that are strongly protected by the region 2 protein encoded by the corresponding gene; the circle with lines represents a site that is relatively weakly protected by FP617; circles with question marks in the HSV-2 and EHV-1 gene promoters are hypothetical binding sites for the corresponding protein. (B) Nucleotide sequences within the protected sites. Strongly protected sequences are shown by solid underlines. Sequences protected relatively weakly are shown by a dashed underline. The conserved pentanucleotide ATCGT is shaded. This sequence is also present on the complementary strand of the VZV ORF62 promoter from −283 to −287. To our knowledge, DNA-binding assays with the HSV-2 ICP4 and EHV-1 IE1 gene promoters have not been reported.

FP449, FP617, and FP805 (Fig. 8). For comparative purposes, the sequences of nucleotides flanking the transcription start sites in the promoters for the HSV-2 ICP4, EHV-1 IE1, and VZV ORF62 genes are also shown. There are several striking features presented in Fig. 8.

First, 25 of the 28 nucleotides that are protected by ICP4 in the HSV-1 ICP4 gene promoter are perfectly conserved at the identical location in the HSV-2 ICP4 gene promoter. Because the three nonconserved nucleotides are located at the extreme ends of the footprint, we think it is reasonable to assume that the HSV-2 ICP4 gene promoter contains a strong binding site for HSV-1 ICP4. Although the sequence encoding region 2 of HSV-2 ICP4 has not been reported, there is substantial evidence indicating that the ICP4 proteins encoded by HSV-1 and -2 are functionally interchangeable (53, 54). Therefore it is reasonable to expect that HSV-2 ICP4 is capable of binding to the transcription start site in its own promoter.

A second feature shown in Fig. 8 is that sites with relatively high affinity for FP449, FP617, and/or FP805 contain the sequence 5′-ATCGT-3′. This pentanucleotide occurs on both strands in a 20-bp imperfect palindrome from −302 to −283 in the VZV ORF62 promoter; it also occurs twice on the same strand in a 41-bp region of the PRV IE180 promoter (−13 to +28). A site in the VZV ORF62 promoter which binds FP617 and FP805 weakly contains a degenerate version (ATCGc) of this pentanucleotide from position +6 to +10. This pentanucleotide is also located immediately upstream from the transcription initiation site in the EHV-1 IE1 gene promoter, which suggests that it may function as a binding site for the EHV-1 IE1 protein.

A third feature indicated in Fig. 8 is that the region 2

domains of the HSV-1 ICP4 and PRV IE180 proteins bind with relatively high affinity to sequences that flank the +1 position in the corresponding genes, whereas the region 2 domain of the VZV ORF62 protein binds less well to the ORF62 gene +1 site. The results with FP449 and FP805 imply that autoregulation of transcription from the HSV-1 ICP4 and PRV IE180 genes may occur by a similar mechanism. The fact that FP617 (region 2 portion of the VZV ORF62 protein) provided only weak protection of the reported (35) transcription initiation site in the VZV ORF62 gene is less satisfying in terms of our proposed unifying hypothesis. There are several ways that the results for DNA binding with FP617 might be reconciled with our hypothesis that ORF62 gene expression is down-regulated by formation of a complex between the ORF62 protein and the ORF62 gene transcription initiation site. A trivial explanation is that the binding specificity of the ORF62-region 2 fusion protein produced in *E. coli* differs somewhat from that of the native protein because of lack of additional residues or lack of specific modifications. Given the numerous weak FP617-binding sites that were observed in the ORF62 gene promoter (Fig. 4E), another explanation is that down-regulation of ORF62 gene expression occurs only when there is sufficient ORF62 protein to saturate several weak sites in the promoter. Finally, if there is an alternative transcription initiation site around position −287 in the ORF62 gene, as suggested by data from Jeff Ostrove's laboratory (20; personal communication), then one could speculate that the strong FP617-binding site between −311 and −282 might provide a mechanism to preferentially repress transcription from this upstream site. Thus, our results raise the possibility that the ORF62 protein influences the selection of transcription initiation sites by binding with differential affinity to two alternative transcription start sites within the ORF62 gene promoter.

The results presented here also yield some information regarding the DNA-binding domains in these related IE regulatory proteins. It is tempting to speculate that the highly conserved stretch of 55 amino acids at the carboxy portion of region 2 represents the residues which are required for recognition of the pentanucleotide sequence ATCGT, which is highly conserved at a similar location in each of the strong binding sites (Fig. 8). Shepard et al. (51) proposed that this portion of region 2 may contain a helix-turn-helix motif that is a feature of some DNA-binding proteins (41). One can further speculate that the less highly conserved amino acid residues upstream from position 435 (in ICP4) are responsible for recognition of specific nucleotides located one full turn of the DNA helix downstream from the ATCGT sequence. The fact that only one of the ICP4-ORF62 chimeric proteins maintained the ability to bind specific DNA sequences demonstrates that equivalent portions of the region 2 domains in these proteins are not uniformly interchangeable without consequence. The fact that the chimeric protein FP460 maintained the same binding activity and specificity as FP617 suggests to us that the amino acid residues which read the DNA sequence are located downstream from residue 354 in ICP4 (or residue 507 in the ORF62 protein). However, analysis of point and insertion mutations in the ICP4 gene have clearly demonstrated that additional residues in the region from positions 262 to 354 in ICP4 also play a role in the DNA-binding activity of ICP4 (42, 43, 51). One overly simplistic interpretation of these results is that residues from 262 to 354 provide a framework that stabilizes the proper conformation required for recognition of specific nucleotides by residues 355 to 490.

In summary, we have shown that the region 2 domains encoded by the HSV-1 ICP4, PRV IE180, and VZV ORF62 genes can be synthesized in *E. coli* as fusion proteins which recognize and bind to related but not identical nucleotide sequences at or near the transcription initiation site in the corresponding viral genes. By extension, we predict that the native forms of the proteins encoded by the HSV-2 ICP4, VZV ORF62, PRV IE180, and EHV-1 IE1 genes are DNA-binding proteins which bind specifically to a sequence that overlaps the transcription start site in the corresponding gene.

ACKNOWLEDGMENTS

We thank Jeff Ostrove, Andrew Cheung, and Joan Betz for generous contributions of plasmids. Roger Everett and Chris Preston kindly communicated experimental results prior to publication and provided helpful discussions. Ming Cao and Lisa Hoffman participated in the initial constructions for the recombinant gene encoding FP805. Mary Jo Rademacher assisted in many of the recombinant gene constructions and provided expert technical assistance throughout this project.

This research was supported by Public Health Service research grant AI 17246 from the National Institutes of Health.

REFERENCES

1. **Campbell, M. E., and C. M. Preston.** 1987. DNA sequences which regulate the expression of the pseudorabies virus major immediate early gene. Virology **157**:307–316.
2. **Caughman, G. B., J. Staczek, and D. J. O'Callaghan.** 1985. Equine herpesvirus type 1 infected cell polypeptides: evidence for immediate early/early/late regulation of viral gene expression. Virology **145**:49–61.
3. **Cheung, A. K.** 1989. DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus. Nucleic Acids Res. **17**:4637–4646.
4. **Chlan, C. A., C. Coulter, and L. T. Feldman.** 1987. Binding of the pseudorabies virus immediate-early protein to single-stranded DNA. J. Virol. **61**:1855–1860.
5. **Clements, J. B., R. J. Watson, and N. M. Wilkie.** 1977. Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on the viral genome. Cell **12**:275–285.
6. **Cromlish, W. A., S. M. Abmayr, J. L. Workman, M. Horikoshi, and R. G. Roeder.** 1989. Transcriptionally active immediate-early protein of pseudorabies virus binds to specific sites on class II gene promoters. J. Virol. **63**:1869–1876.
7. **Davison, A. J., and J. E. Scott.** 1986. The complete DNA sequence of varicella-zoster virus. J. Gen. Virol. **67**:1759–1816.
8. **Davison, A. J., and N. M. Wilkie.** 1983. Location and orientation of homologous sequences in the genomes of five herpesviruses. J. Gen. Virol. **64**:1927–1942.
9. **DeLuca, N. A., and P. A. Schaffer.** 1985. Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4. Mol. Cell. Biol. **5**:1997–1208.
10. **DeLuca, N. A., and P. A. Schaffer.** 1987. Activities of herpes simplex virus type 1 (HSV-1) ICP4 genes specifying nonsense peptides. Nucleic Acids Res. **15**:4491–4511.
11. **DeLuca, N. A., and P. A. Schaffer.** 1988. Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. J. Virol. **62**:732–743.
12. **DiDonato, J. A., and M. T. Muller.** 1989. DNA binding and gene regulation by the herpes simplex virus type 1 protein ICP4 and involvement of the TATA element. J. Virol. **63**:3737–3747.
13. **Dieckmann, C. L., and A. Tzagoloff.** 1985. Assembly of the mitochondrial membrane system. J. Biol. Chem. **260**:1513–1520.
14. **Dixon, R. A., and P. A. Schaffer.** 1980. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. J. Virol. **36**:189–203.

15. **Everett, R. D.** 1987. A detailed mutational analysis of Vmw110, a trans-acting transcriptional activator encoded by herpes simplex virus type 1. EMBO J. **6**:2069–2076.
16. **Everett, R. D.** 1987. The regulation of transcription of viral and cellular genes by herpesvirus immediate-early gene products. Anticancer Res. **7**:589–604.
17. **Faber, S. W., and K. W. Wilcox.** 1986. Association of the herpes simplex virus regulatory protein ICP4 with specific nucleotide sequences in DNA. Nucleic Acids Res. **14**:6067–6083.
18. **Faber, S. W., and K. W. Wilcox.** 1988. Association of herpes simplex virus regulatory protein ICP4 with sequences spanning the ICP4 gene transcription initiation site. Nucleic Acids Res. **16**:555–570.
19. **Feldman, L., F. J. Rixon, J. H. Jean, T. Ben-Porat, and A. S. Kaplan.** 1979. Transcription of the genome of pseudorabies virus (a herpesvirus) is strictly controlled. Virology **97**:316–327.
20. **Felser, J. M., P. R. Kinchington, G. Inchauspe, S. E. Straus, and J. M. Ostrove.** 1988. Cell lines containing varicella-zoster virus open reading frame 62 and expressing the "IE" 175 protein complement ICP4 mutants of herpes simplex virus type 1. J. Virol. **62**:2076–2082.
21. **Gelman, I. H., and S. Silverstein.** 1987. Dissection of immediate-early gene promoters from herpes simplex virus: sequences that respond to the virus transcriptional activators. J. Virol. **61**: 3167–3172.
22. **Grundy, F. J., R. P. Baumann, and D. J. O'Callaghan.** 1989. DNA sequence and comparative analyses of the equine herpesvirus type 1 immediate early gene. Virology **172**:223–236.
23. **Hardwicke, M. A., P. J. Vaughan, R. E. Sekulovich, R. O'Conner, and R. M. Sandri-Goldin.** 1989. The regions important for the activator and repressor functions of herpes simplex virus type 1 alpha protein ICP27 map to the C-terminal half of the molecule. J. Virol. **63**:4590–4602.
24. **Harty, R. N., C. F. Colle, F. J. Grundy, and D. J. O'Callaghan.** 1989. Mapping the termini and intron of the spliced immediate-early transcript of equine herpesvirus 1. J. Virol. **63**:5101–5110.
25. **Honess, R. W., and B. Roizman.** 1973. Proteins specified by herpes simplex virus. XI. Identification and relative molar rates of synthesis of structural and nonstructural herpes virus polypeptides in the infected cell. J. Virol. **12**:1347–1365.
26. **Ihara, S., L. Feldman, S. Watanabe, and T. Ben-Porat.** 1983. Characterization of the immediate-early functions of pseudorabies virus. Virology **131**:437–454.
27. **Kattar-Cooley, P., and K. W. Wilcox.** 1989. Characterization of the DNA-binding properties of herpes simplex virus regulatory protein ICP4. J. Virol. **63**:696–704.
27a. **Kinchington, P., I. Hay, and W. Rucheyan.** Personal communication.
28. **Knipe, D. M., D. Senechek, S. A. Rice, and J. L. Smith.** 1987. Stages in the nuclear association of the herpes simplex virus transcriptional activator protein ICP4. J. Virol. **61**:276–284.
29. **Kristie, T. M., and B. Roizman.** 1986. Alpha 4, the major regulatory protein of herpes simplex virus type 1, is stably and specifically associated with promoter-regulatory domains of alpha genes and of selected other viral genes. Proc. Natl. Acad. Sci. USA **83**:3218–3222.
30. **Levine, M., and J. L. Manley.** 1989. Transcriptional repression of eukaryotic promoters. Cell **59**:405–408.
31. **McCarthy, A. M., L. McMahan, and P. A. Schaffer.** 1989. Herpes simplex virus type 1 ICP27 deletion mutants exhibit altered patterns of transcription and are DNA deficient. J. Virol. **63**:18–27.
32. **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.
33. **McGeoch, D. J., A. Dolan, S. Donald, and D. H. Brauer.** 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res. **14**:1727–1745.
34. **McGeoch, D. J., A. Dolan, S. Donald, and F. J. Rixon.** 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J. Mol. Biol. **181**:1–13.
35. **McKee, T. A., G. H. Disney, R. D. Everett, and C. M. Preston.** 1990. Control of expression of the varicella-zoster virus major immediate early gene. J. Gen. Virol. **71**:897–906.
36. **Metzler, D. W., and K. W. Wilcox.** 1985. Isolation of herpes simplex virus regulatory protein ICP4 as a homodimeric complex. J. Virol. **55**:329–337.
37. **Morse, L. S., T. G. Buchman, B. Roizman, and P. A. Schaffer.** 1977. Anatomy of herpes simplex virus DNA. IX. Apparent exclusion of some parental DNA arrangements in the generation of intertypic (HSV-1 × HSV-2) recombinants. J. Virol. **24**:231–248.
38. **Muller, M. T.** 1987. Binding of the herpes simplex virus immediate-early gene product ICP4 to its own transcription start site. J. Virol. **61**:858–865.
39. **O'Hare, P., and G. S. Hayward.** 1985. Three *trans*-acting regulatory proteins of herpes simplex virus modulate immediate-early gene expression in a pathway involving positive and negative feedback regulation. J. Virol. **56**:723–733.
40. **O'Hare, P., and G. S. Hayward.** 1987. Comparison of upstream sequence requirements for positive and negative regulation of a herpes simplex virus immediate-early gene by three virus-encoded *trans*-acting factors. J. Virol. **61**:190–199.
41. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
42. **Paterson, T., and R. D. Everett.** 1988. Mutational dissection of the HSV-1 immediate-early protein Vmw175 involved in transcriptional transactivation and repression. Virology **166**:186–196.
43. **Paterson, T., and R. D. Everett.** 1988. The regions of the herpes simplex virus type 1 immediate early protein Vmw175 required for site specific DNA binding closely correspond to those involved in transcriptional regulation. Nucleic Acids Res. **16**: 11005–11025.
44. **Paterson, T., V. G. Preston, and R. D. Everett.** 1990. A mutant of herpes simplex virus type 1 immediate early polypeptide Vmw175 binds to the cap site of its own promoter in vitro but fails to autoregulate in vivo. J. Gen. Virol. **71**:851–861.
45. **Pereira, L., M. H. Wolff, M. Fenwick, and B. Roizman.** 1977. Regulation of herpesvirus macromolecular synthesis. V. Properties of alpha polypeptides made in HSV-1 and HSV-2 infected cells. Virology **77**:733–749.
46. **Preston, C. M.** 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK. J. Virol. **29**:275–284.
47. **Rice, S. A., and D. M. Knipe.** 1990. Genetic evidence for two distinct transactivation functions of the herpes simplex virus alpha protein ICP27. J. Virol. **64**:1704–1715.
48. **Roberts, C. R., A. C. Weir, J. Hay, S. E. Straus, and W. T. Ruyechan.** 1985. DNA-binding proteins present in varicella-zoster virus-infected cells. J. Virol. **55**:45–53.
49. **Roberts, M. S., A. Boundy, P. O'Hare, M. C. Pizzorno, D. M. Ciufo, and G. S. Hayward.** 1988. Direct correlation between a negative autoregulatory response element at the cap site of the herpes simplex virus type 1 IE175 (alpha 4) promoter and a specific binding site for the IE175 (ICP4) protein. J. Virol. **62**:4307–4320.
50. **Roizman, B., L. E. Carmichael, F. Deinhardt, G. de-The, A. J. Nahmias, W. Plowright, F. Rapp, P. Sheldrick, M. Takahashi, and K. Wolf.** 1981. Herpesviridae. Definition, provisional nomenclature, and taxonomy. The Herpesvirus Study Group, the International Committee on Taxonomy of Viruses. Intervirology **16**:201–217.
51. **Shepard, A. A., A. N. Imbalzano, and N. A. DeLuca.** 1989. Separation of primary structural components conferring autoregulation, transactivation, and DNA-binding properties to the herpes simplex virus transcriptional regulatory protein ICP4. J. Virol. **63**:3714–3728.
52. **Shiraki, K., and R. W. Hyman.** 1987. The immediate early proteins of varicella-zoster virus. Virology **156**:423–426.
53. **Smith, C. A., and P. A. Schaffer.** 1987. Intertypic recombinants of herpes simplex virus types 1 and 2 infected-cell polypeptide

4. Virology **160**:176–182.
54. **Smith, C. A., and P. A. Schaffer.** 1987. Mutants defective in herpes simplex virus type 2 ICP4: isolation and preliminary characterization. J. Virol. **61**:1092–1097.
55. **Stow, N. D., and E. C. McMonagle.** 1982. Propagation of foreign DNA sequences linked to a herpes simplex virus origin of replication, p. 199–204. *In* Y. Gluzman (ed.), Eukaryotic viral vectors. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
56. **Vlèek, C., V. Paèes, and M. Schwyzer.** 1989. Nucleotide sequence of the pseudorabies virus immediate early gene, encoding a strong transactivator protein. Virus Genes **2**:335–346.
57. **Watson, R. J., and J. B. Clements.** 1980. A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. Nature (London) **285**:329–330.
58. **Wilcox, K. W., A. Kohn, E. Sklyanskaya, and B. Roizman.** 1980. Herpes simplex virus phosphoproteins. I. Phosphate cycles on and off some viral polypeptides and can alter their affinity for DNA. J. Virol. **33**:167–182.
59. **Wu, C. L., and K. W. Wilcox.** 1990. Codons 262 to 490 from the herpes simplex virus ICP4 gene are sufficient to encode a sequence-specific DNA binding protein. Nucleic Acids Res. **18**:531–538.
60. **Yao, F., and R. J. Courtney.** 1989. A major transcriptional regulatory protein (ICP4) of herpes simplex virus type 1 is associated with purified virions. J. Virol. **63**:3338–3344.