JOURNAL OF VIROLOGY, Sept. 1989, p. 3714–3728
0022-538X/89/093714-15$02.00/0

Vol. 63, No. 9

Separation of Primary Structural Components Conferring Autoregulation, Transactivation, and DNA-Binding Properties to the Herpes Simplex Virus Transcriptional Regulatory Protein ICP4

ALYSSA A. SHEPARD,[1,2] ANTHONY N. IMBALZANO,[1,3] AND NEAL A. DELUCA[1,2,3]*

Laboratory of Tumor Virus Genetics, Dana-Farber Cancer Institute,[1] the Committee on Virology,[2] and the Department of Microbiology and Molecular Genetics,[3] Harvard Medical School, Boston, Massachusetts 02115

Received 15 February 1989/Accepted 16 May 1989

A truncated ICP4 peptide which contains the amino-terminal 774 amino acids of the 1,298-amino-acid polypeptide is proficient for DNA binding, autoregulation, and transactivation of some viral genes (N. A. DeLuca and P. A. Schaffer, J. Virol. 62:732–743, 1988) and hence exhibits many of the properties characteristic of intact ICP4. To define the primary sequence important for the activities inherent in the amino-terminal half of the ICP4 molecule, insertional and deletion mutagenesis of the sequences encoding these residues were conducted. The DNA-binding activity of the molecule as assayed by the association with a consensus binding site was sensitive to insertional mutagenesis in two closely linked regions of the molecule. One region between amino acids 445 and 487 is critical for DNA binding and may contain a helix-turn-helix motif. The second region between amino acids 263 and 338 reduces the binding activity to a consensus binding site. When analyzed in the viral background, the DNA-binding activity of a peptide containing an insertion at amino acid 338 to a consensus binding site was reduced while the association with an alternative sequence was eliminated, suggesting a possible mechanism by which ICP4 may recognize a broader range of sequence elements. Mutations which eliminated DNA binding also eliminated or reduced both transactivation and autoregulation, supporting the requirement for DNA binding for these activities. Peptides that retained the deduced DNA-binding domain but lacked amino acids 143 through 210 retained the ability to associate with the consensus site and autoregulatory activity but were deficient for transactivation, demonstrating that the structural requirements for transactivation are greater than those required for autoregulation.

Infected-cell polypeptide (ICP) 4 of herpes simplex virus type 1 (HSV-1) is a large, multifunctional molecule which transcriptionally regulates viral gene expression during productive infection. The 175-kilodalton (kDa) ICP4 protein (6) is expressed in the absence of prior viral protein synthesis and thus is defined as a member of the immediate-early (IE or α) class of HSV-1 gene products (5, 19). ICP4 is required for enhanced expression of both early (β) and late (γ) viral polypeptides, as well as for attenuation of its own transcription and that of other IE genes (13, 41, 42). Viruses carrying temperature-sensitive (*ts*) (13, 41–43), deletion (9, 12), and nonsense (12) mutations that impair the expression or activity of ICP4 fail to synthesize both early and late gene products and exhibit an overproduction of ICP4 and other IE proteins. Similarly, the results of transient expression assays utilizing the cloned gene encoding ICP4 and reporter genes containing the promoters of IE, early, and late viral genes have demonstrated that ICP4 transactivates early and late viral gene promoters while repressing reporter genes under the regulatory control of the ICP4 promoter (10, 14, 16, 30, 35, 36).

The ICP4 polypeptide localizes to the nucleus of infected cells (6, 23, 39) and is phosphorylated such that at least three modified species are observed on sodium dodecyl sulfate-polyacrylamide gels (39, 50). A region of the ICP4 molecule near amino acid residue 700, has been proposed as a nuclear localization signal (12). This region of the ICP4 protein exhibits amino acid similarities with the 140-kDa IE protein of the related herpesvirus, varicella-zoster virus (VZV) (31). The conserved residues specify a sequence which is similar to that required for simian virus 40 large T-antigen nuclear localization (21). While peptides expressed by ICP4 nonsense and deletion mutants encoding this sequence are directed to the nucleus of infected cells, unidentified residues carboxy terminal to the signal are required for correct intranuclear localization (12). Likewise, residues between amino acids 170 and 250, which include the contiguous tract of serine residues conserved in the VZV IE protein (31), have been genetically implicated as a site of ICP4 phosphorylation (12). An ICP4 deletion mutant encoded by the virus *d2*, which lacks the conserved serines within the deletion of residues 185 to 309, is phosphorylated (12), however, suggesting that unidentified residues carboxy terminal to amino acid 309 are also phosphorylated. The effect of this modification of ICP4 structure upon the function of the molecule is unknown.

It also has been demonstrated that the ICP4 molecule can participate in protein-DNA complexes (15, 24, 25, 34). Specific protein-DNA complexes containing ICP4 are formed upon incubation of extracts prepared from HSV-1-infected cells and DNA fragments excised from IE, early, and late HSV-1 gene promoters (15, 24, 25, 33, 34). While some of the identified binding sites, including sequences at the ICP4 transcription initiation site, contain the proposed consensus binding sequence ATCGTCnnnnYCGRC (Y = pyrimidine; R = purine) (15), many HSV-1 gene promoters which are regulated by ICP4 do not. Recent studies have indicated that ICP4 will form protein-DNA complexes with DNA sequences which differ substantially from the consensus (33). Relevant to this study, such nonconsensus binding sites within the domain of the thymidine kinase (*tk*) promoter

* Corresponding author.

have been implied by a previous observation demonstrating an ICP4-DNA complex in the presence of antibody (24).

The mechanism by which ICP4 can both positively and negatively regulate viral gene expression is unclear, but we have previously demonstrated a correlation between the ability of ICP4 to form specific protein-DNA complexes and both transactivation and autoregulation (12). Virus strains containing various ICP4 nonsense and deletion mutations which retained transinduction and autoregulatory activities also retained the ability to form protein-DNA complexes with a DNA fragment spanning the ICP4 start site of transcription, and strains which lacked these regulatory activities failed to form such protein-DNA complexes (12). Moreover, the mutant strains which retained these activities retained the ability to form protein-DNA complexes with DNA fragments in the region of the *tk* promoter, an early gene transactivated by ICP4 which does not possess the ICP4-binding consensus sequence (A. Imbalzano and N. DeLuca, submitted for publication). Given these observations, we conducted a finer genetic dissection of the ICP4 molecule to further test the correlation between DNA binding, transactivation, and autoregulation and to identify and define regions of primary sequence important for each of these functions.

The nonsense mutant *n*208, which specifies a truncated ICP4 polypeptide containing the first 774 amino acids of the intact 1,298-amino-acid molecule (12), was chosen as the parent molecule for the genetic dissection of ICP4 examined in this study. The ICP4 molecule expressed by the *n*208 virus retains the ability to form the specific protein-DNA complexes characteristic of the intact protein, can autoregulate its own expression, and can transactivate a *tk*-chloramphenicol acetyltransferase (tkCAT) chimeric gene to approximately half the level observed with that of wild-type ICP4 (11, 12). The *n*208 ICP4 peptide is not sufficient to support virus growth, however, and when assessed for the ability to enhance the infectivity of the defective ICP4 deletion mutant *d*120, the truncated protein supplied only 3% of the complementing activity observed for intact ICP4 (11). The activities missing in *n*208 specified by the carboxy terminus of the ICP4 protein can, however, be introduced in *trans* by coinfecting cells with a mutant virus specifying the region missing in *n*208 (47). Given that ICP4 exists as a multimer in vitro (32) and that novel heterocomplexes can be observed when analyzing the DNA-binding properties of the ICP4 molecules extracted from coinfected cells, we have suggested that the intragenic complementation results from multimerization of ICP4 molecules in vivo (47). Hence, the *n*208 molecule also possesses the structure required for multimerization. Consistent with its growth defect, *n*208 is impaired in viral DNA synthesis and true late ($\gamma 2$) gene expression (12).

Despite these deficiencies, however, the *n*208 peptide retains many of the properties of the wild-type counterpart and provides a more limited target for further genetic dissection of these properties which are inherent in the intact ICP4 molecule. Initially, this report describes a series of 2-amino-acid insertion mutants generated in the background of the truncated *n*208 molecule, which define two clusters of primary sequence required for DNA-binding and both positive and negative regulatory activity. The subsequent experiments characterize a set of in-frame deletion mutants which retain the identified DNA-binding domains. While all of the deletion mutants retain autoregulatory function, some fail to transactivate viral gene expression, identifying a region of primary sequence important for positive regulation and suggesting a mechanism by which the ICP4 polypeptide can utilize a common DNA-binding domain to regulate differential gene expression.

MATERIALS AND METHODS

Virus and cells. Procedures for the growth and maintenance of Vero, CV-1, and E5 cells were conducted as described previously (8). E5 cells express complementing levels of ICP4 and were derived from Vero cells as described previously (9, 11). The wild-type strain of HSV-1 (KOS), the ICP4-deficient viruses (nonsense mutants *n*12 and *n*208 and the deletion mutant *d*120) (9, 12), and the strains gal4, vi13, and vi16 described in this report were all propagated on E5 cells.

Recombinant plasmids. The plasmid pK1-2 encodes the entire wild-type ICP4 gene and its regulatory sequences (11). The previously described plasmids pn2 and pn7 are derivatives of pK1-2 that contain the insertion of an oligonucleotide specifying a translational termination codon at amino acids 37 and 775, respectively (11). The insertion mutants pi1 through pi19 and the deletion mutants *di*3-*i*4 through *di*3-*i*11, and *di*8-*i*10 described herein were derived from plasmid pn7. The *Eco*RI-to-*Bam*HI fragments from plasmids p4 and p4ΔAT (12) were used as probes for mobility shift assays, as described previously. Plasmid ptkXAP contains HSV-1 sequences from nucleotide positions −385 to −197 relative to the transcription initiation site of the *tk* mRNA. The *Rsa*I-to-*Xba*I fragment from ptkXAP, which spans nucleotide positions −254 to −197 relative to the *tk* transcription initiation site, was used as a probe for gel shift assays. The plasmid pW3-ΔHS8, which encodes the IE gene ICP0, was kindly provided by Wendy Sacks and was used to transactivate the ICP4 promoter in transient assays as described previously (11). Plasmids ptkCAT and pIE3CAT, which contain the bacterial gene for chloramphenicol acetyltransferase (CAT) regulated by the HSV-1 *tk* and ICP4 promoters, respectively, were utilized as reporter genes in transient expression assays as described previously (10). Single-stranded M13 bacteriophage DNAs containing sequences homologous and complementary to the mRNA of HSV-1 genes ICP4, ICP8, and ICP5 for use as probes in nuclear run-on transcription assays were obtained from David Knipe, Harvard Medical School, and have been described previously (17). Single-stranded M13 DNA probes homologous and complementary to the HSV-1 gene encoding *tk*, derived from nucleotide positions +10 to +860 relative to the *tk* transcription initiation site, were also utilized.

Electrophoresis. (i) Viral polypeptides. Viral polypeptides extracted from HSV-1-infected cells were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis as described previously (26, 29). Where indicated, infected-cell monolayers of 5×10^5 Vero cells were labeled with [^{35}S]methionine (20 μCi/ml) (New England Nuclear Corp., Boston, Mass.) as previously described (12).

(ii) Protein-DNA complexes. Protein-DNA complexes formed by incubation of infected-cell extracts and end-labeled DNA probe were resolved in native 4% polyacrylamide gels as previously described (12).

Transfection. (i) Transient expression. For the preparation of extracts for DNA-binding assays, approximately 10^6 CV-1 cells were transfected with 2 μg of a plasmid encoding an ICP4 allele, using the calcium-phosphate coprecipitation procedure previously described (10). Salmon sperm DNA was used to bring the total DNA concentration to 11 μg in 0.5 ml of the transfection buffer. The transfected monolayers

were then infected with the ICP4 deletion mutant *d*120 at a multiplicity of infection (MOI) of 1 PFU per cell. At 12 h postinfection, whole cell extracts to be used in DNA-binding assays were prepared as described for HSV-1-infected Vero cells (12). For the preparation of extracts for in vitro CAT assays, approximately 3×10^6 CV-1 cells were transfected by the same procedure described above with 1 μg of a plasmid encoding an ICP4 allele, and 2 μg of ptkCAT or 2 μg pIE3CAT–1 μg pW3-ΔHS8 in a total volume of 1 ml transfection buffer, with salmon sperm DNA added to bring the final DNA concentration to 22 μg/ml. Monolayers were harvested at approximately 44 h posttransfection. The procedures for the preparation of whole cell extracts from transfected CV-1 cells and the in vitro assay for CAT activity were described previously (10).

(ii) Marker transfer. Approximately 10^6 E5 cells were transfected with 1.5 μg of gal4 viral DNA and 1.5 μg of *Eco*RI-digested plasmids pi13 or pi16 in a total DNA concentration of 11 μg in 0.5 ml, using the procedure described above. The HSV-1 viral strain gal4 contains the insertion of the β-galactosidase (β-gal) gene in both copies of the gene encoding ICP4 and hence gives rise to blue plaques in the presence of 5-bromo-4-chloro-3-indolyl, β-D-galactoside (X-gal). Transfected monolayers were harvested at 2 to 3 days posttransfection. Dilutions of the sonicated, clarified viral lysates were plated on E5 cells and stained with 0.2 mg of X-gal per ml. Clear plaques, indicative of the recombinational exchange of both copies of the β-gal gene for the ICP4 allele encoded by the pi plasmids were isolated and subsequently plaque purified. The identity of the ICP4 allele was confirmed by Southern blot analysis.

DNA-binding assays. Whole cell extracts were prepared from transfected CV-1 cells, as described above, or from approximately 10^6 Vero cells infected at an MOI of 20 PFU per cell, as described previously (12). Then, 3 μl of cell extract was incubated with approximately 0.5 to 1.0 ng of end-labeled DNA (2×10^4 cpm/ng) in the presence of 2 μg of poly(dI)-poly(dC) (Pharmacia) for 30 min and electrophoresed as described above. The procedure for end-labeled probe preparation was as previously described (28).

Western immunoblot analysis. Viral polypeptides separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis as described above were electrophoretically transferred to nitrocellulose filters as previously described (9). The ICP4 polypeptides were visualized by probing filters with either a rabbit polyclonal serum obtained from Richard Courtney, Louisiana State University Medical Center, Shreveport, or from Kent Wilcox, Medical College of Wisconsin, Milwaukee. The conditions for probing with ^{125}I-protein A (New England Nuclear Corp.) and washing the nitrocellulose filters were as described previously (9).

Nuclear run-on transcription analysis. The procedures for the isolation of infected Vero cell nuclei, the in vitro labeling of RNA with [^{32}P]GTP (2 mCi/ml) (New England Nuclear Corp.), and the subsequent isolation of radioactive run-on RNA were as previously described (12, 47, 49). The M13 single-stranded DNA described above was immobilized onto nitrocellulose filters in the manner described previously (12, 49). Conditions for hybridization of run-on RNA to the nitrocellulose filters and subsequent washing to remove nonspecific binding were essentially as described previously (49).

RESULTS

Activities and DNA-binding characteristics of insertion mutants. The plasmid pn7, which encodes the nonsense mutation used to construct the *n*208 virus (11, 12), was mutagenized by the insertion of an oligonucleotide specifying a unique 6-base-pair restriction enzyme recognition site at many loci within the coding sequence for the truncated ICP4 peptide. pn7 was partially digested by the restriction enzyme *Hae*III, which can potentially digest the coding sequence of *n*208 61 times. The synthetic oligonucleotide CTGCAGTCT AGACTGCAG, which can self anneal to form a double-stranded fragment containing, in order, the restriction sites *Pst*I, *Xba*I, and *Pst*I, was ligated to the digested plasmid. The ligation products were then cleaved with excess *Pst*I and electrophoretically fractionated, and the fragment which exhibited the mobility of linearized pn7 was eluted from an agarose gel, recircularized, and transformed into bacteria. The isolated clones were screened for the presence of a unique *Pst*I site and hence an in-frame 2-amino-acid insertion. These clones were reserved for further analysis, whereas clones containing the entire 18-mer or multiple linkers, identified by susceptibility to digestion with *Xba*I, were discarded.

The position of the insertion mutation in the reserved clones was determined by the following methodology. The isolated clones were screened for insertions within the coding region of the *n*208 peptide by double restriction enzyme digestion by using *Pst*I, which marks the insertion, and another enzyme which cleaves the plasmid pn7 once. Clones containing insertions within the vector of the plasmid, in the promoter of the ICP4 gene, or carboxy terminal to the nonsense mutation were discarded. A restriction fragment of pn7 containing the coding sequence of the *n*208 peptide was then radioactively labeled at a unique restriction site (*Sal*I) and then partially digested with the restriction enzyme *Hae*III, to generate a ladder of radioactive markers having a fixed endpoint (*Sal*I) and a variable endpoint designating potential insertion sites. DNA prepared from the chosen clones was labeled at the same restriction site as above (*Sal*I) and then digested with *Pst*I. The sizes of the radioactive *Pst*I fragments were compared with those of the pn7 *Hae*III ladder upon electrophoresis through native polyacrylamide gels, and hence the position of the insertion was precisely determined. Although the location of the hexanucleotide relative to the reading frame of the peptide allows for the insertion of three different pairs of amino acids (Fig. 1), the majority of the clones contained the insertion of two alanine residues and the restoration of an original alanine residue destroyed by the *Hae*III digestion. The positions and the identities of the insertions in the mutants isolated by the above methodology are shown relative to the coding sequence of the *n*208 molecule encoded by pn7 (Fig. 1). The diagram at the top of the figure represents the regions of amino acid similarity with the VZV IE protein. Table 1 indicates the exact position of the insertions by listing the last wild-type residue prior to the insertion.

(i) Transient assays. The effects of these mutations on ICP4 function were studied by transient techniques, and the activities of the mutants were compared with those of the parental molecule encoded by pn7, which specifies the *n*208 partial peptide of ICP4. Mutant ICP4 alleles transiently expressed in CV-1 cells were assessed for the ability to (i) form protein-DNA complexes with a DNA fragment containing the ICP4 transcription initiation site, (ii) repress the expression of a CAT gene regulated by the ICP4 promoter (pIE3CAT), and (iii) transactivate the expression of a tkCAT chimeric gene (ptkCAT).

Prior to examining the effect of the insertion mutations on the DNA-binding activity of the *n*7 peptide, the identity and

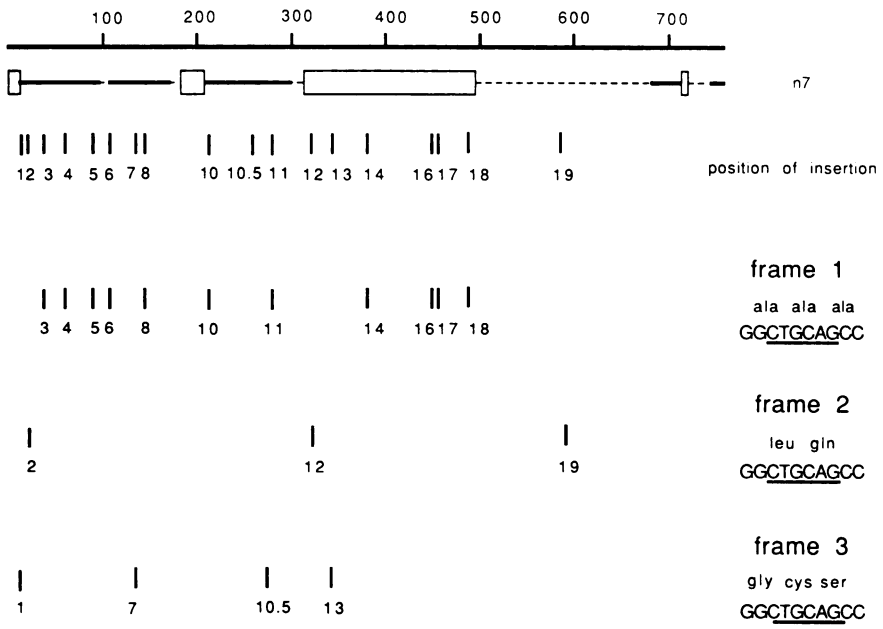


FIG. 1. Positions and amino acid identities of the ICP4 insertion mutants. The primary sequence of the ICP4 nonsense mutant encoded by the plasmid pn7, which was used as the parent molecule for insertional mutagenesis, is indicated above the positions of insertions *i*1 through *i*19. The regions of the amino acid sequence most similar to those of the 140-kDa VZV analog are depicted by open rectangles in the primary sequence encoded by pn7, while dashed lines represent sequences missing in the VZV IE protein (31). The three possible frames in which the hexanucleotide CTGCAG was inserted into pn7 are shown as are the inserted residues specified by each mutant. The majority of the insertion mutants contain two alanine residues inserted in the position where one existed in the parent molecule.

specificity of the *n*7 protein-DNA complex in mobility shift experiments was verified. The *Bam*HI-to-*Eco*RI fragment containing the DNA sequences from −110 to +30 relative to the transcription initiation site of the ICP4 mRNA contains a consensus ICP4-binding site (15). Gel shift experiments performed on cells transfected with the plasmid encoding the *n*7 molecule revealed the formation of a transfected cell-specific complex with the wild-type *Bam*HI-to-*Eco*RI fragment (Fig. 2). When polyclonal antiserum was added to the reaction mixture after an initial incubation period to allow for binding, the transfected cell-specific complex was not observed at the mobility seen in the absence of the antibody. Due to the polyclonal nature of the antibody used in the experiment, discrete shifts in mobility of ICP4-containing complexes were only seen with very dilute antibody preparations. When enough antibody was added to shift all the material from its initial position, most of the protein-DNA complexes remained in the well as is seen in Fig. 2. It should be noted that none of the other lower-abundance complexes seen in Fig. 2 were affected by the presence of antibody, indicating that the antibody has not merely shifted complexes in a nonspecific manner. The consensus binding site for ICP4 is the sequence ATCGTCnnnnYCGRC (15). When the dinucleotide AT was deleted from this sequence, the antibody-sensitive, transfected cell-specific complex no longer formed, reflecting the specificity of the complex.

TABLE 1. Positions of 2-amino-acid insertions

ICP4 mutant	Last wild-type amino acid prior to insertion[a]
pi1	Pro-12
pi2	Gly-17
pi3	Gly-30
pi4	Arg-63
pi5	Glu-93
pi6	Ala-104
pi7	Gly-138
pi8	Glu-142
pi10	Ala-210
pi10.5	Gly-263
pi11	Val-274
pi12	Gly-320
pi13	Gly-338
pi14	Glu-386
pi16	Ala-449
pi17	Val-452
pi18	Gly-487
pi19	Arg-603

[a] The numbering for the amino acids and their identities are taken from McGeoch et al. (31).

Shown in Fig. 3C are the protein-DNA complexes formed by incubating the *Bam*HI-to-*Eco*RI probe containing the wild-type ICP4-binding site with extracts prepared from cells transfected with pi1 through pi19. The parent plasmid pn7, encoding the *n*208 nonsense mutation, served as the positive control (last lane), while pUC19 DNA served as the negative control (first lane). The insertion mutants which retain DNA-binding activity exhibit a specific protein-DNA complex of similar mobility to that observed with extracts prepared from pn7-transfected cells.

With the exception of pi1, insertions that clustered in two regions of the ICP4 molecule between amino acid residues 263 and 487 inhibited the formation of protein-DNA complexes with the ICP4 transcription initiation site probe. The ICP4 peptides encoded by mutant plasmids pi1, pi10.5, pi11, pi12, and pi13 exhibited greatly diminished DNA-binding activity, and the ICP4 alleles encoded by pi16 through pi18

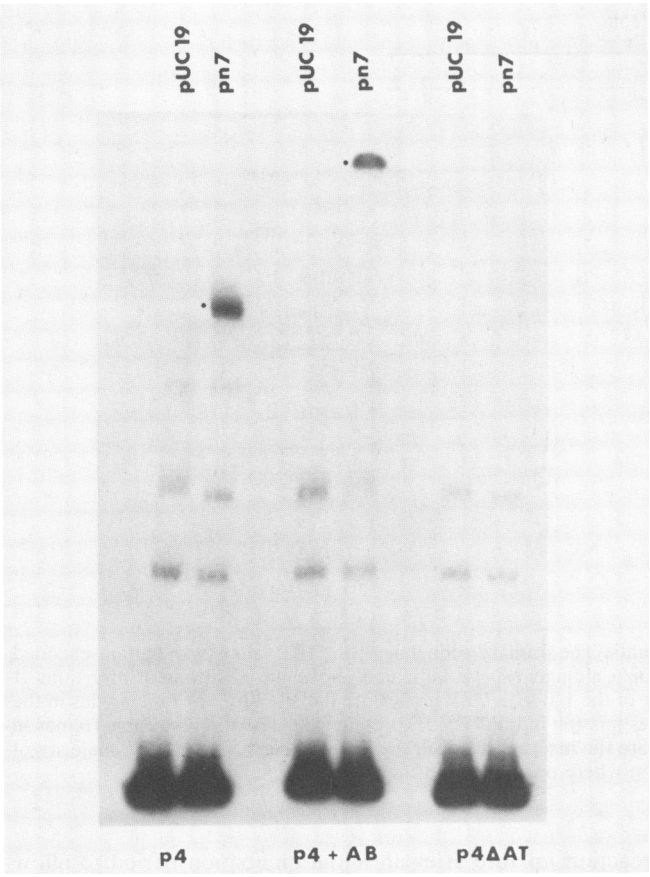


FIG. 2. Identity and specificity of the transfected *n*7 peptide. Extracts were prepared from cells transfected with the parent plasmid pn7 and pUC19 DNA as a negative control, as described in Materials and Methods. The protein-DNA complexes formed from these extracts when incubated with the DNA probe p4 are shown in the first two lanes. p4 contains the *Bam*HI-to-*Eco*RI fragment from the ICP4 promoter, spanning nucleotides −110 to +30 relative to the ICP4 transcription initiation site. A polyclonal antibody (AB) directed against the ICP4 polypeptide was added to the binding assays represented by the indicated lanes (p4 + AB). The last two lanes exhibit the protein-DNA complexes formed upon incubation of the extracts with the DNA probe p4ΔAT (12). p4ΔAT contains the same *Bam*HI-to-*Eco*RI fragment from the ICP4 transcription initiation site yet lacks the dinucleotide AT from the consensus binding site ATCGTC. The novel complexes formed with the *n*7 extracts are indicated by dots.

failed to form detectable protein-DNA complexes. The ICP4 peptide encoded by pi14, which specifies an insertion after amino acid residue 386 which is located between the mutations encoded by the two sets of previously mentioned plasmids, retains substantial DNA-binding activity. The region of the ICP4 gene that contains the insertions *i*16, *i*17, and *i*18 exhibits extensive amino acid homology with the ICP4 counterpart of VZV (31) and contains the locus of the inactive ICP4 temperature-sensitive mutant *ts*K (7). Western blot analysis of the extracts used for the DNA-binding analysis indicated that similar levels of ICP4 from each transfected insertion mutant were found in the extracts and that each ICP4 polypeptide exhibited electrophoretic mobilities in sodium dodecyl sulfate-polyacrylamide gels similar to those observed for the *n*208 ICP4 molecule encoded by pn7 (data not shown). The altered DNA-binding activity observed for some of the ICP4 insertion mutants relative to that of *n*208 most probably reflects a structural perturbation resulting from the addition of 2 amino acids. Therefore, these results suggest two closely linked regions of primary sequence which are important to confer the ability to form protein-DNA complexes with a DNA probe containing a consensus binding site.

The expression of ICP4 and other IE genes, initially stimulated possibly by a virion component referred to as Vmw65 (2, 3, 40) (also identified as VP16 and α-transinducing factor) or ICP0 (36), is subsequently repressed later in infection (19). This negative regulation is believed to be a direct consequence of the function of the ICP4 polypeptide (13, 42). It has previously been demonstrated that wild-type ICP4 when transiently expressed in CV-1 cells will marginally repress the expression of a plasmid that contains the bacterial CAT gene under the regulatory control of the ICP4 promoter (pIE3CAT) (10, 36). The HSV-1 IE polypeptide ICP0 has been shown to be a potent transactivator of ICP4 and many other genes in transient assays (14, 16, 30, 35, 36, 44). The negative regulatory effect of ICP4 is dominant to the positive effect of ICP0 (36). Therefore, when a plasmid encoding ICP4 is cotransfected with pIE3CAT and the plasmid pW3-ΔHS8, which encodes ICP0, the expression of CAT is attenuated 30- to 50-fold (11). The ICP4 nonsense mutant encoded by pn7 represses the ICP4 promoter in transient assays slightly less effectively than wild-type ICP4 (11).

Figure 3B examines the ability of the ICP4 insertion mutants to repress the expression of pIE3CAT in the presence of ICP0. The numbers above each sample in the figure represent the ratios of CAT activity in the extract of the transfected ICP4 mutant to the CAT activity resulting from transfected pUC19 DNA (first lane), averaged over four separate trials. Mutants impaired in the ability to negatively regulate pIE3CAT were also defective in DNA-binding activity. The products of plasmids pi16 through pi18, which failed to form protein-DNA complexes (Fig. 3C), exhibited levels of CAT activity similar to that seen in the transfection with just ICP0 and pIE3CAT (compare the first lane of Fig. 3B with the lanes for pi16, pi17, and pi18). Three mutant proteins (*i*1, *i*12, and *i*13) that were severely impaired in the ability to form the protein-DNA complex (Fig. 3C) were also impaired with respect to repressive activity. Proteins proficient for DNA binding repress pIE3CAT expression 5- to 10-fold (i.e., pi2 through pi10, pi14, pi19, and pn7). The products of pi10.5 and pi11 were also apparently reduced in their abilities to bind to DNA yet they retain the ability to repress pIE3CAT expression. The DNA-binding and autoregulatory activities of these proteins expressed from recombinant viruses are being evaluated to further address the behavior of these molecules.

These results support the observations that (i) the *cis*-acting signals responsible for ICP4 autoregulation in transient assays are located in the region of the ICP4 transcription initiation site (36, 45) and (ii) partial peptides resulting from nonsense mutations in the viral ICP4 genes, which localize to the nucleus of infected cells and retain the ability to associate with the start site of ICP4 transcription, specify autoregulatory activity (11). The finding that the primary sequence of ICP4 important for autoregulatory activity is also necessary for associating with the ICP4 transcription initiation site strengthens the correlation between DNA binding and autoregulation and supports the theory that at least one mechanism conferring ICP4 autoregulation involves blocking the start site of transcription.

Shown in the CAT assays in Fig. 3A are the relative

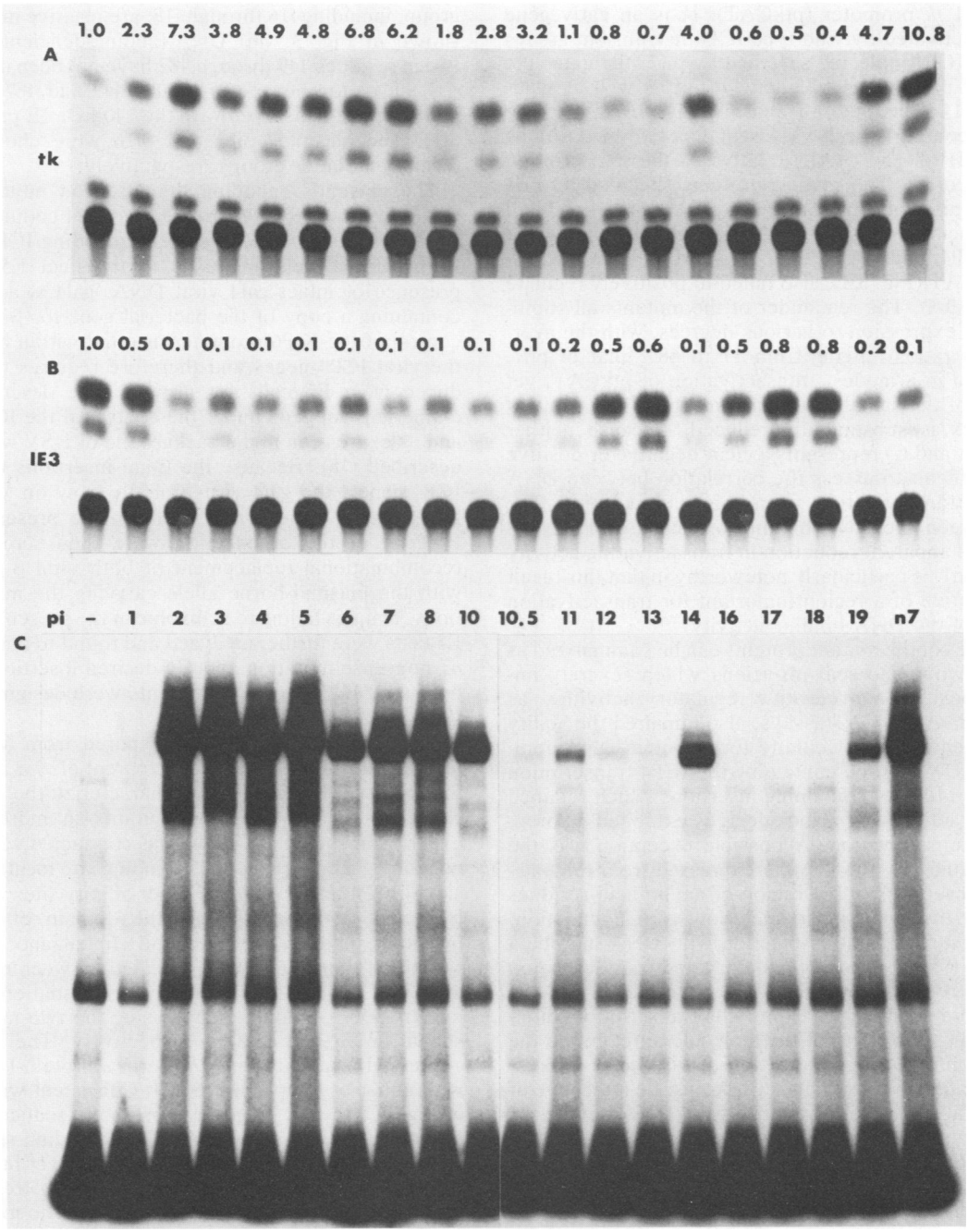


FIG. 3. Transient analysis of the insertion mutants. (A) Mutant ICP4 alleles transiently expressed in CV-1 cells as described in Materials and Methods were tested for the ability to induce the expression of the CAT gene regulated by the HSV-1 *tk* promoter. The first lane indicates the basal uninduced level of ptkCAT expression, while the last lane signifies the level of transactivation by the ICP4 nonsense mutant encoded by pn7. The relative abilities of the ICP4 insertion mutants encoded by pi1 through pi19 to transactivate ptkCAT are shown in the remaining lanes, in the order listed above the figure in panel C. The ratio of CAT activity resulting from the induction of the transfected ICP4 mutant to the CAT activity resulting from transfected pUC19 DNA is given above each lane. (B) The ability of each ICP4 allele to repress the expression of CAT under the regulatory control of the ICP4 (IE3) promoter is shown in the manner described for panel A. Each transfection experiment included the gene encoding ICP0 to induce the expression of the ICP4 promoter so that the autoregulatory response of ICP4 could be observed (11). The first lane contains just the ICP0 plasmid and pIE3CAT and thus gives the level of unrepressed pIE3CAT. For both panels A and B, data from representative experiments are shown. The numbers given are averages from four such experiments rounded to two significant figures. (C) Protein-DNA complexes formed by using extracts prepared from CV-1 cells transfected with the ICP4 insertion mutants, as described in Materials and Methods. The DNA probe (p4) contained the *Bam*HI-to-*Eco*RI fragment of the ICP4 promoter, which contains sequences spanning nucleotide positions −110 to +30 relative to the ICP4 transcription initiation site.

abilities of the mutant ICP4 alleles to induce the expression of a plasmid encoding the bacterial gene for CAT regulated by the HSV-1 *tk* promoter (ptkCAT). *tk* is an early gene readily transactivated by wild-type ICP4 in transient assays (16, 35). The ICP4 molecule specified by pn7 stimulates the expression of ptkCAT approximately half as well as wild-type ICP4 (11). The ptkCAT plasmid (10) used in these experiments contains the HSV-1 sequences from positions −725 (*Bam*HI) to +54 (*Bgl*II) relative to the transcription initiation site of *tk*. The mutant proteins specified by pi1, pi11 through pi13, and pi16 through pi18, which exhibited impaired association with a DNA fragment containing the consensus ICP4-binding site (Fig. 3C) and impaired repression of pIE3CAT (Fig. 3B), also failed to positively regulate ptkCAT (Fig. 3A). The remainder of the mutants all stimulated ptkCAT expression to various degrees, with the average induction ratios ranging from 17 to 68% that of pn7. pi10.5 resulted in a modest transactivation of ptkCAT, yet the intensity of the signal resulting from the pi10.5 protein-DNA complex is substantially reduced. Because neither result (Fig. 3A and C) represents a clear negative or positive value, it is difficult to assess the correlation between DNA binding and transactivation from this datum point. The contrast between the DNA-binding proficiency of the pi8 protein and its apparent reduction in transactivating activity (17% that of pn7) is particularly noteworthy in that this result may be indicative of a region important for transactivation independent of the DNA-binding activity.

The activities of the mutant proteins can be summarized as follows. (i) Two-amino-acid insertions which severely impaired both positive and negative regulatory activities, as analyzed by transient CAT assays, also impaired the ability of the truncated ICP4 molecule to form protein-DNA complexes with a DNA probe containing the ICP4 transcription initiation site. Mutants containing insertions which retained DNA-binding activity to various degrees exhibited autoregulation and transactivation. A possible exception was the activities specified by the peptide expressed from the plasmid pi8. (ii) Insertions in the region of amino acid residues 449 through 487 were the most deleterious to ICP4 function. These results suggest that association with DNA may be a requirement for both the autoregulatory and transactivation functions inherent in the *n*208 ICP4 molecule. While the ability to form a complex with the ICP4 transcription initiation site may be a property of the molecule unrelated to the mechanism conferring transinduction of viral gene expression, this specific DNA-binding activity may be indicative of a more general ability to associate with DNA sequences exhibiting a nucleotide character divergent from the consensus binding site, such as the ICP4-binding sites identified in the *tk* gene promoter (Imbalzano and DeLuca, submitted).

(ii) Recombinant viruses. The regions between the insertions *i*10.5 through *i*13 (residues 263 through 338) and *i*16 through *i*18 (residues 449 through 487) are sensitive to insertional mutagenesis, rendering the molecule deficient for DNA binding as determined by the experiment shown in Fig. 3C. The ambiguities associated with transfection experiments and the desire to more closely examine the DNA-binding properties of the mutant peptides led us to recombine two selected alleles into the viral genome. Given the two apparent regions indicated by the insertions *i*10.5 through *i*13 and *i*16 through *i*18, we decided to construct mutant viruses containing a member from each region with the following observations in mind. (i) The mutants *i*12 and *i*13 are inactive in the transient assays with both ptkCAT and pIE3CAT, but retain a low affinity for the binding site at the start site of ICP4 transcription relative to the molecule expressed from pn7 (Fig. 3). (ii) Mutants of the second group, including *i*16 through *i*18, are inactive in all the assays shown in Fig. 3. (iii) DNA-binding-deficient mutants between residues 449 through 487 have not been described, and (iv) the region between residues 449 and 487 is within the domain most conserved with the 140-kDa IE protein of VZV (31). The plasmids pi13 and pi16 were chosen from the former and latter groups, respectively.

The plasmids encoding the insertion mutants were digested with *Eco*RI, separating the ICP4 coding region from the vector. The *Eco*RI fragment encoding ICP4 was eluted from agarose gels and used to transfect E5 cells in the presence of intact gal4 viral DNA. gal4 is a mutant virus containing a copy of the bacterial gene for β-gal under the control of the ICP6 promoter inserted within both copies of the viral ICP4 genes and therefore requires the ICP4 produced from E5 cells for growth. The description of the chimeric β-gal gene under the control of the ICP6 promoter and its expression in the background of HSV was previously described (18). Because the β-gal insertions inactivate the ICP4 genes, the gal4 virus fails to grow on Vero cells yet forms blue plaques on E5 cells in the presence of X-gal. Progeny of the transfection were then screened for the recombinational replacement of both viral β-gal insertions with the plasmid-borne allele carrying the insertion mutations. Plaques failing to stain blue in the presence of X-gal on E5 cells were further analyzed and found to contain both the *n*7 nonsense mutation and the desired insertion mutation at both viral ICP4 loci. The mutants were designated vi13 and vi16.

Whole cell extracts were prepared from Vero cells infected with vi13, vi16, *n*208 (the parent of vi13 and vi16), and the wild-type strain KOS to characterize the DNA-binding properties of the insertion mutants in more detail. The experiment shown in Fig. 3C was conducted with a fragment containing a consensus ICP4-binding site located at the start site of ICP4 transcription. Lower-affinity sites which deviate from this consensus within the domain of the early *tk* promoter have also been found (Imbalzano and DeLuca, submitted). One such site is located between residues −197 and −254 relative to the transcription initiation site of the *tk* mRNA. DNA-binding assays using the two target sites are shown in Fig. 4A and B, respectively. The ICP4 peptide expressed from vi16 failed to form a protein-DNA complex with either of the probes tested. Consistent with the results of Fig. 3C, the vi13 peptide possesses a reduced but detectable affinity for the consensus ICP4-binding site within the ICP4 promoter region. This peptide however failed to form a complex with the binding site located upstream of the *tk* promoter. The quantity and electrophoretic mobility of the mutant peptides in the extract were determined by Western blot analysis. Figure 4C shows that greater quantities of ICP4 peptide were present in the insertion mutant extracts than in the extract prepared from cells infected with the parental virus *n*208. Moreover, similar electrophoretic species were observed for *n*208, vi13, and vi16. Therefore, despite the greater quantities of the vi13 and vi16 peptides relative to the *n*208 peptide, the vi16 peptide is absolutely deficient in DNA binding and the vi13 peptide possesses a reduced ability to form a complex with the recognized consensus binding site (15) and does not form a complex with the site within the *tk* gene under the conditions of our assay. A vi13-tk probe complex was not evident despite overexposures, various quantities and types of nonspecific

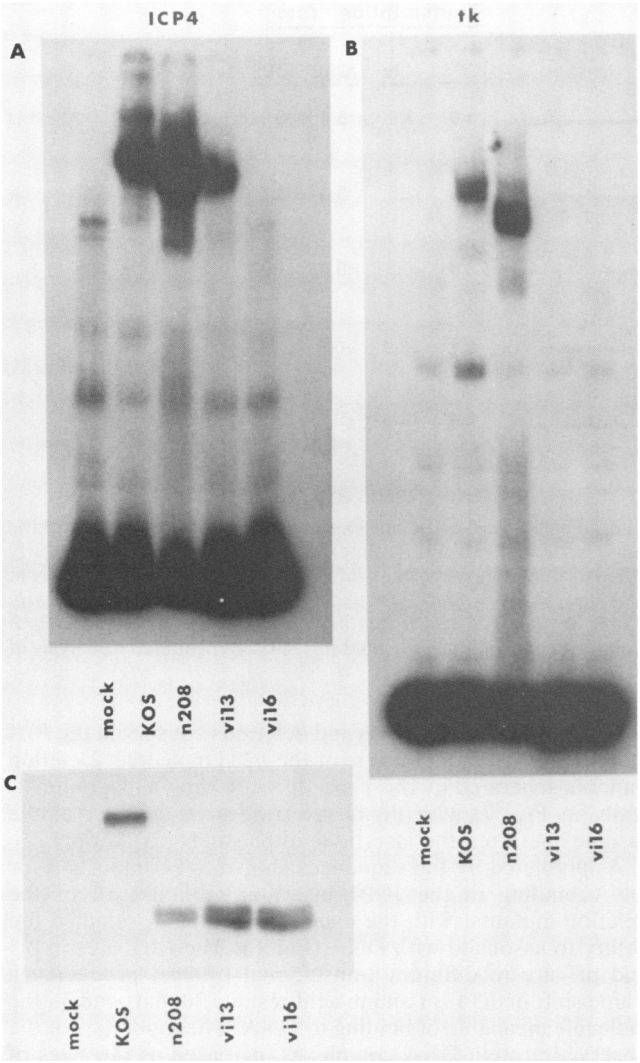


FIG. 4. DNA-binding properties of mutant ICP4 polypeptides produced in infected Vero cells. Protein-DNA complexes formed with extracts prepared from Vero cells infected with the indicated viruses at an MOI of 20 PFU per cell. KOS is the wild-type HSV-1 strain, and *n*208 specifies the truncated ICP4 molecule used as the parent for the insertion mutants specified by vi13 and vi16. The DNA probe used in panel A contained the *Bam*HI-to-*Eco*RI fragment of the ICP4 promoter described in the legend of Fig. 2, whereas the DNA probe used in panel B contained sequences spanning nucleotide positions −197 to −254 relative to the *tk* transcription initiation site. (C) ICP4 polypeptides produced by mutant and wild-type virus detected by an anti-ICP4 polyclonal antibody, as described in Materials and Methods.

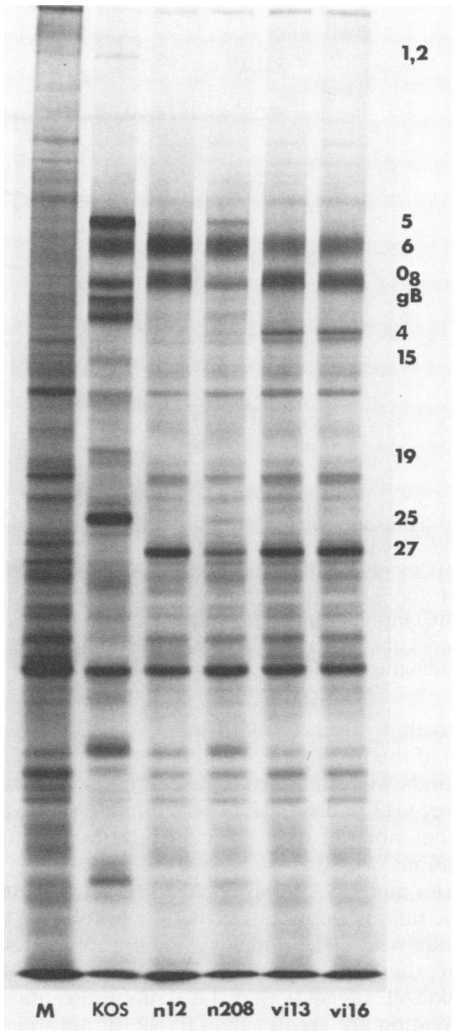


FIG. 5. Polypeptide profiles characteristic of ICP4 mutant viruses. Vero cells infected with the indicated viruses at an MOI of 20 PFU per cell were labeled with 20 μCi of [^{35}S]methionine per ml from 6 to 9 h postinfection. Sodium dodecyl sulfate polypeptides solubilized from cells were resolved in a 9% polyacrylamide gel. Peptides labeled in uninfected cells (mock infected) are shown in the lane labeled M. Indicated on the right are the ICP designations for selected viral polypeptides. The peptides designated 4 represent the *n*208, vi13, and vi16 proteins. The identities of these species as the result of the translation of the 774-amino-acid *n*208 protein have previously been established (12).

DNA used in the binding reactions, and various salt concentrations (data not shown).

Two experiments were conducted to examine the effects of the insertions in vi13 and vi16 on gene expression in the background of the virus. The first experiment examined the polypeptides expressed in KOS-, *n*208-, vi13-, and vi16-infected cells. Figure 5 shows that at the sensitivity of this analysis both vi13 and vi16 were unable to stimulate the expression of early (β) and late (γ_1) genes, as demonstrated by the underrepresentation of ICP1, ICP2, ICP5, ICP8, and ICP25. Figures 5 and 4C also indicate that vi13 and vi16 are deficient in autoregulation. Both conclusions are consistent with the transient results shown in Fig. 3A and B.

To more closely examine the activities of the mutant peptides, we determined the rates of transcription for the ICP4, ICP8, *tk*, and ICP5 genes by nuclear run-on transcription. ICP8, *tk*, and ICP5 are all genes induced directly by ICP4. The run-on experiment shown in Fig. 6 was performed from nuclei of cells infected with viruses KOS, *n*208, vi13, vi16, and *n*12 in the presence of the DNA synthesis inhibitor phosphonoacetic acid to restrict the measurement to input templates and also to minimize the often-observed nonspecific transcription occurring after the onset of DNA synthesis (17, 49). The *n*12 virus is completely deficient with respect to ICP4 activity (12). The results of Fig. 6 are consistent with the results of Fig. 3A and B and 5 in that the transcription of the ICP4 gene from vi13 and vi16 was elevated compared with that of KOS and *n*208 and was

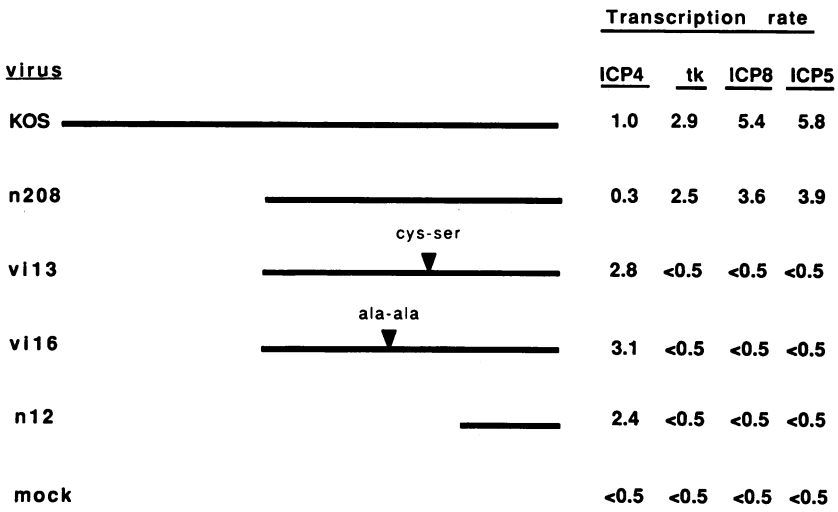
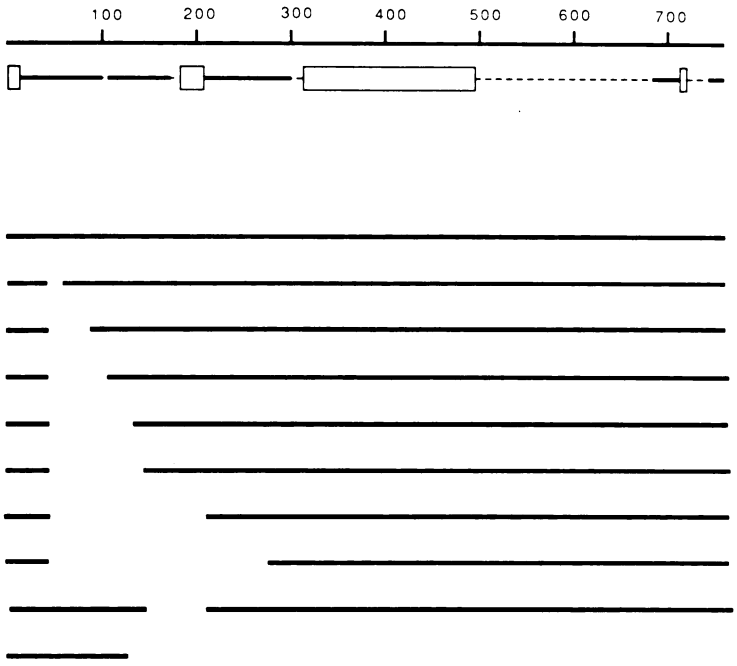


FIG. 6. Transcription rates of selected viral genes in wild-type- and mutant-infected Vero cells. The expressed amino acid sequence of the ICP4 alleles in wild-type (KOS), *n*208, vi13, vi16, and *n*12 viral strains are represented next to the transcription rates of ICP4 and the early genes *tk*, ICP8, and ICP5. Nuclear run-on analysis was performed on nuclei isolated from Vero cells infected with the indicated viruses at an MOI of 20 PFU per cell, in the presence of 300 μg of the DNA synthetic inhibitor phosphonoacetic acid per ml. The amount of radioactivity incorporated into the run-on RNA did not vary by more than 20% from sample to sample. Thus, the entire run-on RNA sample was used to probe filters containing serial threefold dilutions of single-stranded DNA complementary to the indicated mRNAs. Densitometry analysis of the autoradiographic images of in vitro ^{32}P-labeled run-on RNA is given in arbitrary units.

similar to that observed with inactive *n*12 virus. The transcription of the ICP8, ICP5, and *tk* genes in the vi13 and vi16 backgrounds was similar to that seen in the *n*12 background. Therefore, vi13 and vi16 did not stimulate the expression of these genes nor did they attenuate the expression of ICP4 transcription.

Activities and DNA-binding characteristics of deletion mutants. The mutant peptide specified by pi8 retained the ability to associate with DNA and specified autoregulatory activity. However, the reduced ability of the pi8 peptide to transactivate ptkCAT suggests that the structural requirements for transactivation are greater than those for autoregulation. To test this hypothesis, a nested set of in-frame deletion mutants was constructed from the insertion mutants described above. Mutants were designed to retain the predicted DNA-binding domain while harboring progressively larger in-frame deletions in the residues amino terminal to amino acid residue 274. Deletion mutants were constructed as follows. Insertion mutants pi3 through pi11, with the exception of pi7 and pi10.5, all contain a *Pst*I recognition site encoding two alanine residues inserted between the first and second nucleotides of codons which also specify alanine (Fig. 1). The *Pst*I site of pi3 at amino acid residue 30 served as the amino-terminal endpoint for the deletions; the carboxy-terminal endpoints were determined by the loci of the insertions in pi4 through pi11. pi7 was used in this analysis as a negative control since we had predetermined that a gene constructed with pi3 and pi7 would contain a frame shift amino terminal to the DNA-binding domain. The individual pi plasmids were digested with the restriction enzyme *Pst*I, which cleaves the inserted hexanucleotide, and the enzyme *Hin*dIII, which digests the vector sequences outside the gene for ICP4. Pairs of digested plasmids were combined in ligation reaction mixtures in a one-to-one ratio and introduced into bacteria. Four products of the ligation were isolated in roughly equal proportions. In addition to the two original insertion mutants, ICP4 molecules containing duplications and deletions of the residues between the two insertions were generated from each ligation reaction. Each of the resulting duplications and deletions conserved the *Pst*I linker and the reading frame of the *n*208 molecule. Deletion mutants generated by this procedure are shown diagrammatically in Fig. 7, with di*x*-i*y* referring to a deletion of the residues between the insertions in plasmids pi*x* and pi*y*.

As predicted by the sequences shown to be important for DNA binding in the ICP4 insertion mutants, all of the deletion mutants, with the exception of di3-i7, retained the ability to associate with DNA (Fig. 8). The *Pst*I sites in pi3 and pi7 are in different translational frames, producing a frame shift in di3-i7 at amino acid residue 30 that renders the molecule incapable of binding to DNA. The mobilities of the observed protein-DNA complexes increased as the sizes of the constructed deletions increased, as would be expected if the deletions were in frame, conserving the integrity of the DNA-binding domain. The DNA-binding assays shown in Fig. 8 were conducted by the same procedure described for the experiment in Fig. 3C, and the interaction between ICP4 deletion mutants and the DNA probe spanning the ICP4 start site of transcription was compared with that observed with *n*7. The ICP4 nonsense mutant encoded by pn2 has been previously shown to lack sequences necessary for DNA binding (11). The retention of DNA-binding activity by all of the ICP4 deletion mutants supports the suggestion based on the insertion mutant data that sequences carboxy terminal to amino acid residue 274 specify a structure essential for associating with DNA. The ICP4 molecules specified by di3-i10 and di3-i11, however, exhibited reduced binding compared with that of the other deletion mutants and *n*7 (first and second panels of Fig. 8). The interaction between ICP4 deletion mutants and the ICP4 start site of transcription was specific, as shown by failure of the ICP4 molecule encoded by di3-i10 to associate with an analogous DNA probe containing a deletion of the AT dinucleotide from the consensus binding site ATCGTC (compare protein-DNA complexes observed with wild-type and mutant probes, p4 and p4ΔAT, respectively, shown in the right panel of Fig. 8). The plasmids containing duplications between amino acids 31 and 274 also associated with the probe, further demon-



	A	B	C
	IE3 CAT	tk CAT	DNA:protein complex
n7	0.1	6.1	+
di3-i4	0.1	6.1	+
di3-i5	0.1	5.0	+
di3-i6	0.1	3.9	+
di3-i7	nd	nd	-
di3-i8	0.1	6.0	+
di3-i10	0.1	1.7	+
di3-i11	0.2	0.6	+
di8-i10	0.1	1.1	+
n2	1.2	1.1	-

FIG. 7. Structures and activities of ICP4 deletion mutants. The expressed amino acid sequences of the ICP4 deletion and nonsense mutants is shown, where dix-iy refers to a deletion of residues between insertions ix and iy. The ability of each ICP4 deletion mutant to repress the expression of CAT under the regulatory control of the ICP4 (IE3) promoter is shown relative to that of the nonsense mutants specified by *n*7 and *n*2 in column A. The data in column B indicate the relative abilities of the deletion mutants to induce the expression of CAT regulated by the *tk* promoter. These transient CAT assays were conducted in the manner described in the legend to Fig. 3 and in Materials and Methods. Column C summarizes the DNA-binding data shown in Fig. 8.

strating that the DNA-binding domain of the constructions was retained, in frame, carboxy terminal to amino acid 274 (data not shown). These mutants were not analyzed further. The DNA-binding properties of the deletion mutants are summarized in column C of Fig. 7.

The ability of these ICP4 deletion mutants to function in the transient expression assays described in the first section of this report is also shown in Fig. 7. In column A of Fig. 7, the ability of these ICP4 molecules to attenuate the expression of the ICP0-induced pIE3CAT gene is indicated relative to that of the molecules encoded by pn7 and pn2. All of the deletion mutants tested retained the ability to repress the expression of the ICP4 promoter in the pIE3CAT plasmid with an efficiency similar to that observed for *n*7, relative to the inactive *n*2 molecule. In that these molecules can associate with the binding site located at the start site of ICP4 transcription, this result supports the previously observed correlation between DNA-binding and autoregulatory activity.

The level of induction of tkCAT expression exhibited by the ICP4 deletion mutants relative to *n*7 and *n*2 is also indicated in Fig. 7. The ICP4 molecule encoded by pn7 stimulated the expression of tkCAT approximately sixfold over the basal level, whereas the inactive peptide encoded by pn2 failed to induce tkCAT above this level. ICP4 molecules lacking increasing amounts of sequence between insertions *i*3 and *i*8 all retained the transactivation function characteristic of *n*208. Mutants containing the deletion of residues carboxy terminal to the insertion in pi8 (amino acid residue 142) failed to induce the expression of the tkCAT chimeric gene (Fig. 7, column B). Since the ICP4 molecule encoded by di3-i8 is as proficient as *n*7 in transactivating tkCAT, yet the additional deletion of the residues between *i*8 and *i*10 in di3-i10 resulted in a protein that is impaired in this capacity, a deletion mutant solely lacking the sequence between *i*8 and *i*10 was constructed to more rigorously test the requirement of these residues for transactivation. The deletion in di8-i10 eliminated amino acid residues 143 through 210, which includes a region of serine residues conserved with the VZV IE protein (31). As shown in column B of Fig. 7, di8-i10 also failed to induce the expression of tkCAT, yet this deletion mutant specifies autoregulatory activity and can participate in protein-DNA complexes (Fig. 8, third panel). This observation suggests that (i) DNA binding as measured by the association with the consensus binding site is not sufficient for transactivation and (ii) the region of primary sequence containing the conserved serine residues, proposed to serve as a phosphorylation site (12), is required for conferring positive regulatory function but is not necessary for the autoregulatory function inherent in the *n*208 ICP4 molecule.

DISCUSSION

The ICP4 protein of HSV is a complex multifunctional regulatory protein which stimulates the transcription of many HSV-1 genes and is therefore required for viral growth (13, 41). The induced genes are often of the early and late kinetic classes (19), although the promoter requirements for ICP4 induction have yet to be elucidated. ICP4 is also autoregulatory (13, 42), acting in part by binding to a strong binding site located at the start site of ICP4 transcription (45).

The complexity of the ICP4 protein is indicated by several independent genetic analyses. The study of McGeoch et al. (31) examined the regions of amino acid conservation between the 175-kDa ICP4 protein of HSV-1 and the analogous 140-kDa protein of VZV. Sequence similarities were found

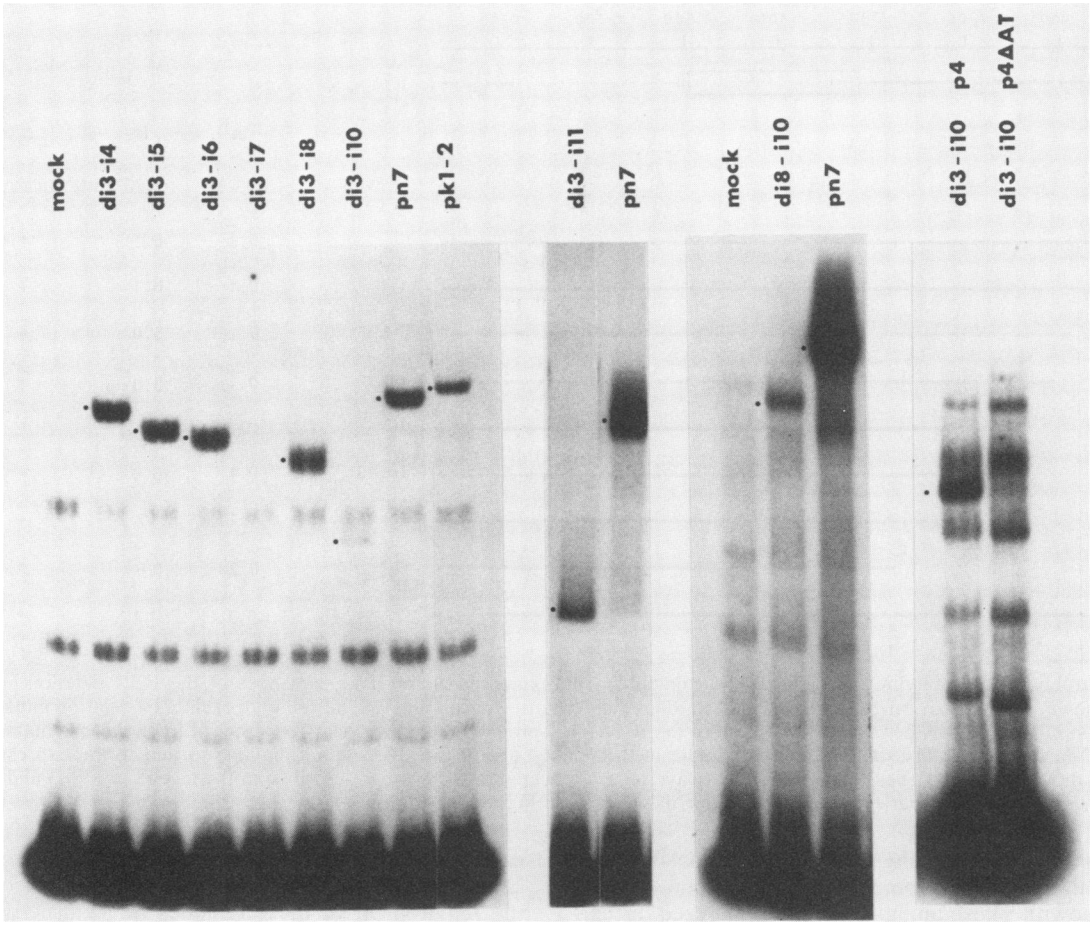


FIG. 8. DNA-binding properties of ICP4 deletion mutants. Extracts prepared from CV-1 cells transfected with ICP4 deletion mutants were incubated with the DNA probe p4, spanning the ICP4 transcription initiation site as described in the legend to Fig. 3C. The left and middle two panels reveal the protein-DNA complexes formed with the ICP4 deletion mutants described in the text. The right panel shows a longer exposure of the complex formed with extract prepared from cells transfected with the mutant di3-i10 incubated with the ICP4 start site probe (p4) and the analogous probe which lacks the dinucleotide AT from the consensus binding site ATCGTC (p4ΔAT). The protein-DNA complexes specific to the transiently expressed ICP4 species are indicated by dots.

to be confined to discrete regions of the two proteins rather than distributed in a more dispersive fashion, indicating the evolutionary conservation of functionally important regions or domains of the protein. Mutational analysis has begun to ascribe functional and biochemical significance to the conserved domains of the ICP4 molecule (11, 12, 37, 38). Of note and germane to the present study is the observation that the amino-terminal 774 amino acids expressed by the mutant virus *n*208 specify a protein which binds to DNA, is autoregulatory, and transactivates a number of HSV-1 genes (11, 12). The present study evaluates the contribution of various regions of the *n*208 molecule in the processes of DNA binding, autoregulation, and transactivation.

The *n*208 molecule was chosen for this analysis because of its inherent properties and activities. In addition, it has long been recognized that mutations carboxy terminal to the nonsense mutation in *n*208 can have a deleterious effect on the activities of the ICP4 molecule. A number of ICP4 *ts*-mutants whose loci map in the carboxy-terminal domain do not autoregulate and do not stimulate gene expression (8). Consistent with the observations from the studies with *ts* mutants, it was also shown that a nonsense mutation at amino acid 1060 has a more deleterious effect on ICP4 activity than a nonsense mutation at amino acid 774 (11, 12). Presumably these mutations alter the molecule such that it may be unstable, improperly localized, or conformationally changed, thus rendering the active domains nonfunctional. Therefore, limiting the analysis to the *n*208 molecule simplifies the analysis in this regard.

In the present study, the insertion mutants pi1 through pi19 reveal information about the regions of the ICP4 molecule which are involved in DNA binding. Analysis of a nested set of in-frame deletion mutants addresses the functional requirements for structures in addition to the DNA-binding domain for the autoregulatory and transactivity of the *n*208 molecule. The interpretation of our results is given diagrammatically in Fig. 9.

DNA binding. The ability of the ICP4 molecule to associate with specific DNA sequences has been examined in mobility shift, immunoprecipitation, and southwestern blot experiments (15, 24, 25, 33, 34). A consensus binding site motif (ATCGTCnnnnYCGRC) has been proposed which specifies a relatively strong interaction (15), however, it has recently been shown that ICP4 will associate specifically with sequences which deviate considerably from the consensus (33). The data presented in Fig. 3 show that the region between insertion mutants pi10.5 (residue 263) and pi18 (residue 487) contributes to the DNA-binding property of the

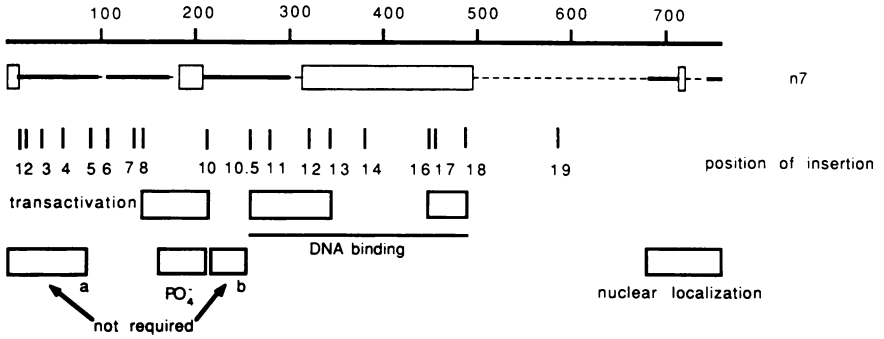


FIG. 9. Functional domains of the truncated ICP4 polypeptide. The positions of the 2-amino-acid insertions described in this report are shown under the expressed amino acid sequence of the truncated ICP4 molecule expressed by pn7. The regions of amino acid similarity between ICP4 and the VZV IE protein (31) are indicated by open rectangles in the *n*7 coding sequence. Regions of primary sequence shown to be important for DNA binding (as determined by the data on insertion mutants shown in Fig. 3) and transactivation (as determined by the data on the deletion mutants shown in Fig. 7 and 8) are indicated below the positions of insertions. Sequences previously shown to be important for phosphorylation and nuclear localization (12), as well as two regions found to be nonessential for viral growth (11, 12, 46), are also indicated. An ICP4 mutant viral strain lacking the region "a" which is proficient for viral growth has been described previously (11, 12), resulting from translational restart at residue 90 due to the insertion of a translation termination codon at amino acid residue 12. A spontaneous deletion mutant proficient for growth, which lacks residues 209 through 236 specifying region "b", has also been described previously (46).

*n*208 molecule. pi1 was also drastically reduced in the ability to form a protein-DNA complex. Because we have previously demonstrated that the region containing the pi1 insertion is not essential for DNA-binding activity (11, 12), it is possible that the pi1 mutation introduces a conformational change that drastically alters the molecule. Additional mutations in this region will be necessary to further test this possibility.

The subtleties of Fig. 3C and the analysis of the viral insertion mutants vi13 and vi16 divide this segment of primary structure into two regions with respect to the ability to affect DNA binding by this mutagenic approach. The region defined by insertions *i*16, *i*17, and *i*18 is the most critical for DNA binding as revealed by its great sensitivity to insertional mutagenesis. In our experiments, these were the only mutants for which no protein-DNA complex was evident. This is especially evident in Fig. 4, which shows that despite relatively large quantities of the mutant protein in the extract, the *i*16 mutation absolutely abolishes DNA binding. The region containing the insertions *i*16, *i*17, and *i*18 is highly conserved with that of the VZV 140-kDa IE protein. In fact, when the rules of Chou and Fasman (4) are applied to the deduced amino acid sequence, the predicted secondary structure in this region of ICP4 is also similar to that of the VZV protein. The insertional loci of *i*16 (residue 449) and *i*17 (residue 452) fall in a region predicted to form a 10-amino-acid hydrophobic α helix in HSV-1 between residues 445 and 454. This structure is conserved in the VZV counterpart. A predicted β turn consisting of the residues Ser-Arg-Arg-Tyr (positions 455 through 458) followed by an aspartic acid residue follows the above helical structure. These five residues are exactly conserved in a homologous position in the VZV protein. A longer predicted amphiphilic α-helical domain of moderate amino acid conservation follows between residues 460 and 487. This region contains the loci for *i*18 and the temperature-sensitive mutant *ts*K (7).

It is evident that the region between residues 445 and 487 constitutes the most important region with respect to DNA binding as measured by our analysis. By analogy to studies on 434 bacteriophage repressor protein (1), the predicted helix-turn-helix motif may indicate that this region is directly involved in the specific recognition of DNA. Studies are in progress to test this hypothesis.

In a similar study (38), it was found that the insertion of four amino acids was most deleterious to DNA binding at residues 329, 337, and 373. This study did not analyze mutants in the region between residues 445 and 487 described above. Our results indicate that while this former region is important for DNA binding, all of the insertion mutants within this region retain some DNA-binding properties. These are *i*10.5 (residue 263), *i*11 (residue 274), *i*12 (residue 320), and *i*13 (residue 338). When the *i*13 mutation is recombined into the viral genome, the ICP4 peptide is expressed in greater quantities than the parent *n*208 (Fig. 4), indicative of impaired autoregulation. The vi13 polypeptide is able to form a complex with reduced efficiency with the binding site located at the start site of ICP4 transcription. Interestingly, it is not able to form a complex with the site upstream of the *tk* promoter. Therefore, while the mutation in vi13 affects the association with the ICP4 consensus site, it apparently has a greater deleterious effect on the association with the alternative site upstream of the *tk* promoter, suggesting that the structural requirement for the latter interaction is greater than that for the former. A greater number of nonspecific protein-DNA interactions may be required for the association with sites which are of lower affinity and deviate from the consensus. Alternatively, the region defined by insertions *i*10.5 to *i*13 may act to broaden the scope of specific ICP4-DNA interactions. The consensus sequence ATCGTCnnnnYCGRC (15) represented in B-form DNA would put the specific nucleotide bases (underlined) on the same side of the helix across 1.5 turns. Given the lack of similarity or symmetry between both halves of the consensus, it is possible that the protein recognizing this sequence contains separate domains which engage in protein-DNA contacts across two major grooves. The summation of interactions across two major grooves may broaden the scope of binding sites which confer a stable and significant association. It should also be considered that either or both of these regions do not directly make contact with DNA but rather specify a structure required for binding to DNA (i.e., multimerization).

Autoregulation. The ability to repress expression of pIE3CAT correlates well with the ability of the ICP4 peptide to associate with the binding site located at the start site of ICP4 transcription. This result is similar to those obtained in

previous studies (12, 38, 45). The results of the present study extend on previous work in two ways. (i) The structural requirements for autoregulation are less than those required for transactivation and may only require DNA binding. The deletion mutants di3-i10, di3-i11, and di8-i10 do not transactivate ptkCAT but retain the ability to repress pIE3CAT. As a test of this hypothesis in viral infection, we have observed that the di3-i11 molecule stably expressed from transformed cells can bind specifically to DNA and attenuate the expression of ICP4 from an infecting virus in the absence of later gene expression (unpublished observations). (ii) The vi13 molecule retains a certain degree of affinity for the above binding site yet it does not elicit autoregulatory activity as measured by the run-on analysis shown in Fig. 6. Given the observations mentioned above, the affinity of the vi13 peptide may not be great enough in vivo to result in a measurable effect. Perhaps this mode of transcriptional repression requires a relatively strong interaction. The dissociation constant for purified ICP4 interacting with a consensus binding site has been reported to be 1.1×10^{-9} M (22).

The mechanism by which the expression of ICP4 is attenuated during viral infection, in part, requires the strong interaction of ICP4 with the sequence at its own transcription initiation site. While it has recently been shown that an ICP4 promoter-CAT chimera containing a deletion of the dinucleotide pair AT of the sequence ATCGTC at the ICP4 transcription initiation site is not repressed by ICP4 in transient assays (45), this mutation is not sufficient to eliminate the repression of ICP4 transcription during viral infection (N. DeLuca, manuscript in preparation). Therefore, the shutoff of ICP4 transcription and possibly the transcription of other IE genes may be mediated by more than one mechanism.

Transactivation. The insertion mutations which eliminate DNA binding also eliminate the ability to transactivate tkCAT. The results of this and previous studies (12, 38) strengthen the previously proposed hypothesis (12) that DNA-binding activity is a necessary condition for transactivation. While most ICP4-inducible genes do not contain the consensus ATCGTCnnnYCGRC, several sites within the domain of the *tk* promoter which deviate from the consensus by several nucleotides have been found (Imbalzano and DeLuca, submitted). Therefore, these sites may represent sequences requiring the same region of ICP4 that is also required for binding to the start site of ICP4 transcription. The functional significance of these sites within ICP4-stimulated genes has yet to be determined.

The observation that the deletion mutants di3-i10, di3-i11, and di8-i10 retain DNA-binding properties and autoregulatory activity but fail to transactivate tkCAT indicates that although DNA binding appears necessary for transactivation it may not be sufficient. Therefore the structural requirements for transactivation by the *n*208 molecule include the DNA-binding domain but also include a part of the protein amino terminal to the DNA-binding domain. As shown above, the amino-terminal region is not required for autoregulation. The mutations in di3-i10, di3-i11, and di8-i10 remove amino acids 31 through 210, 31 through 274, and 143 through 210, respectively. The region deleted in all three mutants is that between residues 143 and 210. This deletes a serine-rich region which is conserved in VZV (31) and also is proposed to be a phosphorylation site (12) (Fig. 9). The residues amino terminal to amino acid 143 contain many charged amino acids, which appear to be of minor importance for transactivation. The serine-rich region is predicted by the rules of Chou and Fasman (4) to be in a loop structure. The negative charge afforded to this structure by phosphorylation may serve as a transactivating domain in a manner similar to that of many regulatory proteins (20, 27, 48). Alternatively, ICP4 may serve as a phosphate donor for the modification of cellular transcription factors. The phosphorylation of the *Escherichia coli* RNA polymerase by the bacteriophage T7-encoded protein kinase has been suggested as a mechanism of transcriptional control (51). A third possibility could be that this region is necessary to specify interactions with non-consensus binding sites. It is possible that phosphorylation could alter the DNA-binding properties of ICP4 in a manner similar to that previously implied (35) or may allow for the interaction with other DNA-binding proteins. We have not investigated the interaction of peptides expressed from transfected cells with the *tk* site because the lower relative affinity for this site makes the interpretation of signals from transfected cells difficult. Preliminary experiments with the di3-i11 molecule expressed from transformed cells suggests that it is deficient in forming complexes with the binding sites in *tk* (data not shown). We are currently investigating other deletion mutants in this area in an attempt to distinguish among these possibilities.

The serine-rich tract is required for transactivation in the context of the *n*208 molecule. It is not absolutely required in the context of the entire ICP4 protein (N.A.D., unpublished observations). The *i*8-*i*10 deletion in the context of the entire protein results in 40 to 60% of the activation observed with the entire molecule. A virus containing this mutation is capable of growing in the absence of complementing ICP4 supplied by transformed cells (N.A.D., unpublished observations). Therefore both sequences carboxy terminal to amino acid 774 and the serine-rich tract contribute to transactivation.

The process of transactivation by ICP4 requiring DNA binding may be complicated, involving contact with a number of cellular transcription factors. This hypothesis would be consistent with multiple domains of the protein conferring transactivity in the presence of the DNA-binding activity and the observation that multiple ICP4-binding sites are found in ICP4-inducible genes (Imbalzano and DeLuca, submitted) both upstream and downstream of the *cis*-acting signals for cellular transcription factors.

ACKNOWLEDGMENTS

We thank Priscilla Schaffer for helpful discussions and for reviewing the manuscript and Sandra Weller for providing a plasmid containing the *lacZ* gene under the control of the ICP6 promoter. We also thank Lucy Goldberg for manuscript preparation.

The work was supported by Public Health Service grant A127431 from the National Institutes of Health and by American Cancer Society grant JRFA-195 to N.A.D. A.A.S. and A.N.I. are predoctoral students funded by Public Health Service institutional training grants 5T32AI0724 and 5T32GM07196-14, respectively, from the National Institutes of Health.

LITERATURE CITED

1. **Anderson, J. E., M. Ptashne, and S. C. Harrison.** 1985. The structure of a phage repressor-operator complex at 7A resolution. Nature (London) **316**:596–601.
2. **Batterson, W., and B. Roizman.** 1983. Characterization of the herpes simplex virion-associated factor responsible for the induction of α genes. J. Virol. **46**:371–377.
3. **Campbell, M. E. M., J. W. Palfreyman, and C. M. Preston.** 1984. Identification of herpes simplex virus DNA sequences which encode a *trans*-acting polypeptide responsible for stimulation of immediate early transcription. J. Mol. Biol. **180**:1–19.
4. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of

protein conformation. Annu. Rev. Biochem. **47**:251–276.
5. **Clements, J. B., R. J. Watson, and N. M. Wilkie.** 1977. Temporal regulation of herpes simplex virus type 1 transcription: localization of transcripts on the viral genome. Cell **12**:275–285.
6. **Courtney, R. J., and M. Benyesh-Melnick.** 1974. Isolation and characterization of a large molecular weight polypeptide of herpes simplex virus type 1. Virology **62**:539–551.
7. **Davison, M.-J., V. G. Preston, and D. J. McGeoch.** 1984. Determination of the sequence alteration of herpes simplex virus type 1 temperature-sensitive mutant *ts*K. J. Gen. Virol. **65**:859–863.
8. **DeLuca, N. A., M. A. Courtney, and P. A. Schaffer.** 1984. Temperature-sensitive mutants in herpes simplex virus type 1 ICP4 permissive for early gene expression. J. Virol. **52**:767–776.
9. **DeLuca, N. A., A. McCarthy, and P. A. Schaffer.** 1985. Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. J. Virol. **56**:558–570.
10. **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**:1997–2008.
11. **DeLuca, N. A., and Schaffer, P. A.** 1987. Activities of herpes simplex virus type 1 (HSV-1) ICP4 genes specifying nonsense peptides. Nucleic Acids Res. **15**:4491–4511.
12. **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.
13. **Dixon, R. A. F., 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.
14. **Everett, R. D.** 1984. Transactivation of transcription by herpes virus products: requirements for two HSV-1 immediate-early polypeptides for maximum activity. EMBO J. **3**:3135–3141.
15. **Faber, S. W., and K. W. Wilcox.** 1986. Association of the HSV-1 regulatory protein ICP4 with specific nucleotide sequences in DNA. Nucleic Acids Res. **14**:6067–6082.
16. **Gelman, I. H., and S. Silverstein.** 1985. Identification of immediate early genes from herpes simplex virus that transactivate the virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA **82**:5265–5269.
17. **Godowski, P. J., and D. M. Knipe.** 1986. Transcriptional control of herpes virus gene expression: gene functions required for positive and negative regulation. Proc. Natl. Acad. Sci. USA **83**:256–260.
18. **Goldstein, D. J., and S. K. Weller.** 1988. Herpes simplex virus type 1-induced ribonucleotide reductase activity is dispensable for growth and DNA synthesis: isolation and characterization and of an ICP6 *lacZ* insertion mutant. J. Virol. **62**:196–205.
19. **Honess, R. W., and B. Roizman.** 1974. Regulation of herpes virus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. **14**:8–19.
20. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46**:885–894.
21. **Kalderon, D., B. Roberts, W. D. Richardson, and A. E. Smith.** 1984. Sequence requirements for nuclear localization of SV40 large T antigen. Nature (London) **311**:33–38.
22. **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.
23. **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.
24. **Kristie, T. M., and B. Roizman.** 1986. α4, the major regulatory protein of herpes simplex virus type 1, is stably and specifically associated with the promoter-regulatory domains of α genes and of selected other viral genes. Proc. Natl. Acad. Sci. USA **83**:3218–3222.
25. **Kristie, T. M., and B. Roizman.** 1986. DNA-binding site of major regulatory protein α4 specifically associated with promoter-regulatory domains of α genes of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **83**:4700–4704.
26. **Laemmli, U. K.** 1979. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
27. **Ma, J., and M. Ptashne.** 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48**:847–853.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 122–123. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Manservigi, P., P. Spear, and A. Buchan.** 1977. Cell fusion induced by herpes simplex virus is promoted and suppressed by different viral glycoproteins. Proc. Natl. Acad. Sci. USA **74**: 3913–3917.
30. **Mavromara-Nazos, P., S. Silver, J. Hubenthal-Voss, J. C. McKnight, and B. Roizman.** 1986. Regulation of herpes simplex virus 1 genes: α gene sequence requirements for transient induction of indicator genes regulated by β or late (γ_2) promoters. Virology **149**:152–164.
31. **McGeoch, D. J., A. Dolan, S. Donald, and D. H. K. Brauer.** 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res. **14**:1727–1744.
32. **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.
33. **Michael, N., D. Spector, P. Mavromara-Nazos, T. M. Kristie, and B. Roizman.** 1988. The DNA-binding properties of the major regulatory protein α4 of herpes simplex virus. Science **239**: 1531–1534.
34. **Muller, M. T.** 1987. Binding of the herpes simplex virus type 1 gene product ICP4 to its own transcription start site. J. Virol. **61**:858–865.
35. **O'Hare, P., and G. S. Hayward.** 1985. Evidence for a direct role for both the 175,000- and 110,000-molecular-weight immediate-early proteins of herpes simplex virus in the transactivation of delayed-early promoters. J. Virol. **53**:751–760.
36. **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.
37. **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.
38. **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.
39. **Pereira, L., M. H. Wolff, M. Fenwick, and B. Roizman.** 1977. Regulation of herpes virus macromolecular synthesis. V. Properties of α polypeptides made in HSV-1 and HSV-2 infected cells. Virology **77**:733–749.
40. **Post, L. E., S. Mackem, and B. Roizman.** 1981. Regulation of α genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with α gene promoters. Cell **24**:555–565.
41. **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 *ts*K. J. Virol. **29**:275–284.
42. **Preston, C. M.** 1979. Abnormal properties of an immediate early polypeptide in cells infected with the herpes simplex virus type 1 mutant *ts*K. J. Virol. **32**:357–369.
43. **Preston, V. G.** 1981. Fine-structure mapping of herpes simplex virus type 1 temperature-sensitive mutations within the short repeat region of the genome. J. Virol. **39**:150–161.
44. **Quinlan, M., and D. Knipe.** 1985. Stimulation of expression of a herpes simplex virus DNA-binding protein by two viral functions. Mol. Cell. Biol. **5**:957–963.
45. **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 (α4) promoter and a specific binding site for the IE175 (ICP4) protein. J. Virol. **62**:4307–4320.

46. **Schroder, C. H., J. DeZazzo, K. W. Knopf, H. C. Kaerner, M. Levine, and J. Glorioso.** 1985. A herpes simplex virus type 1 mutant with a deletion in the polypeptide-coding sequences of the ICP4 gene. J. Gen. Virol. **66**:1589–1593.
47. **Shepard, A. A., and N. A. DeLuca.** 1989. Intragenic complementation among partial peptides of herpes simplex regulatory protein ICP4. J. Virol. **63**:1203–1211.
48. **Triezenberg, S. J., R. C. Kingsbury, and S. L. McKnight.** 1988. Functional dissection of VP16, the *trans*-activator of herpes simplex virus immediate early gene expression. Genes Dev. **2**:718–729.
49. **Weinheimer, S. P., and S. L. McKnight.** 1987. Transcriptional and post-transcriptional controls establish the cascade of herpes simplex virus protein synthesis. J. Mol. Biol. **195**:819–833.
50. **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.
51. **Zillig, W., H. Fujiki, W. Blum, D. Janekovic, M. Schweiger, H.-J. Rahmsdorf, H. Ponta, and M. Hirsch-Kauffmann.** 1975. In vivo and in vitro phosphorylation of DNA dependent RNA polymerase of Escherichia coli by bacteriophage-T7-induced protein kinase. Proc. Natl. Acad. Sci. USA **72**:2506–2510.