JOURNAL OF VIROLOGY, Apr. 1988, p. 1145–1157
0022-538X/88/041145-13$02.00/0

Vol. 62, No. 4

Differentiation and DNA Contact Points of Host Proteins Binding at the *cis* Site for Virion-Mediated Induction of α Genes of Herpes Simplex Virus 1

THOMAS M. KRISTIE AND BERNARD ROIZMAN*

The Marjorie B. Kovler Viral Oncology Laboratories, the University of Chicago, Chicago, Illinois 60637

Received 6 July 1987/Accepted 20 December 1987

Transcriptional *trans*-activation of the five herpes simplex virus 1 α genes by the α *trans*-inducing factor requires a *cis*-acting site (αTIC; with the consensus 5′-GyATGnTAATGArATTCyTTGnGGG-3′) located in the promoter-regulatory domains of the α genes. In DNA band shift assays with nuclear extracts from either mock-infected or infected cells, the DNA fragments containing an αTIC sequence from the α0, α4, and α27 genes formed several cellular protein-DNA complexes designated αH1, αH2, and αH3. The host proteins that formed the αH2 and αH3 complexes were differentiated from those that formed the αH1 complex but not from each other by chromatography and specificity of the DNA-binding sites. The αH1 proteins protected the αTIC sequence of all three genes from DNase I digestion. Methylation of the purines in the sequence 5′-GyATGnTAAT-3′ located at the 5′ terminus of the αTIC sites precluded the binding of αH1. The binding site of the αH2-αH3 proteins in the α27 gene αTIC overlapped, in part, with the αH1-binding site. The binding of these proteins was precluded by methylation of the purine residues in the sequence 5′-GCCACGTG-3′ located at the 3′ terminus of the DNase I footprint. The maximum apparent molecular weight of αH1 was 110,000, whereas that of αH2-αH3 was 64,000. A protein designated αH2′, resembling αH2-αH3 with respect to molecular weight and chromatographic properties but differing in sequence specificity, bound to a site adjacent to the αH1 site in the fragment carrying an αTIC sequence of the α4 gene. αH1 and αH2-αH3 or αH2′ bound concurrently, notwithstanding the apparent overlap in the DNase I footprints.

The 70+ genes of herpes simplex virus 1 (HSV-1) expressed during productive infection are tightly regulated in a cascade fashion (17, 18). Interest in the regulation of HSV-1 genes stems from the complexity of the HSV-1 genome, the diversity of the *trans*-activating factors, and the increasing realization that the activation of the productive infection is subordinate to a regulated process that determines whether the virus will replicate or remain latent. Because this process occurs very early in infection, the focus of many of the studies is on the α genes, the first set of genes transcribed after infection (17, 18). Unique among diverse families of viruses, HSV-1 encodes a virion protein which activates in *trans* the transcription of the five α genes (1, 42). The *trans*-activating protein, identified as the infected cell protein 25 (17), or Vmw65 by Campbell et al. (5), and designated as the α *trans*-inducing factor (αTIF [40]), is contained in the tegument (1), a virion component located between the capsid and the envelope of the virus. *trans*-Induction by αTIF requires specific *cis*-acting elements, including a short AT-rich sequence designated as the αTIC (α *trans*-induction *cis* site), which is present in one to several copies in the 5′ promoter-regulatory regions of all α genes (4, 5, 8, 23, 29–31, 34, 40, 45, 49). In these studies in which deletion mutations and chimeric gene constructs were used, it has been demonstrated that the αTIC sequence can confer α-gene regulation on some recipient promoter-indicator genes.

Studies designed to detect the direct binding of the αTIF protein to the αTIC site have not been successful. These experiments have revealed, however, that DNA fragments containing representative αTIC sequences bind host cell rather than viral-encoded proteins (26). Earlier, we reported (26) that in DNA band shift experiments, DNA fragments containing αTIC sites formed several complexes with proteins in nuclear extracts of mock-infected or infected cells. Of the three complexes designated αH1, αH2, and αH3, the most prominent, αH1, was shown to protect the domain of the αTIC sequence of the α27 gene from DNase I digestion (26). Here we report on the properties of the αH1 and αH2-αH3 proteins and on the DNA sequences with which they interact.

The promoter-regulatory domains of α genes possess several significant features. Foremost, the sequences upstream from the transcription initiation site of α genes are separable into promoter domains, which confer upon recipient genes the capacity to be expressed, and regulatory domains, which confer to both α and non-α promoters the capacity to be regulated as α genes (29–31). In addition to the αTIC sequences, the regulatory domains of α genes contain GC-rich regions identified in at least one gene (α4) to be binding sites for the transcriptional factor SP1 (19, 30). Homologs of the immunoglobulin enhancer element (10, 51) were also identified within α-gene promoter domains. Lastly, several of the genes have been shown to contain the binding sites for α4, the major regulatory protein specified by HSV-1 in infected cells (22, 24, 25, 37). The binding sites and the corresponding factors appear to account for the constitutive expression of α genes in cells or, in the case of the α4 protein, for the negative regulation of α-gene expression late in infection (4, 23, 38, 39, 44, 56). They do not account for the virion-mediated induction of these genes. There is, to date, very little information on the mechanism by which αTIF *trans*-activates the transcription of α genes. Inasmuch as αTIF and its *cis*-acting site must play a key role in the initiation of viral gene expression after infection, elucidation of the mechanism by which αTIF initiates the expression of viral genes after infection requires a more thorough under-

* Corresponding author.

standing of the interaction of the cellular proteins with the αTIC sequences.

MATERIALS AND METHODS

Cells and protein extracts. HeLa cells were grown to confluency in 850-cm^2 roller bottles in Dulbecco modified Eagle medium supplemented with 5% fetal bovine serum. The nuclei were washed once in 0.05% Nonidet P-40 prior to protein extraction with 0.42 M KCl (9). The extract was dialyzed for 5 h and cleared by centrifugation at 25,000 × *g* for 30 min. Typical extracts contained 8 to 12 mg of protein per ml.

Clones, DNA probes, and competitor DNAs. The DNA fragments used in this study were cloned by standard techniques. All enzymes were used according to the recommendations of the manufacturer. The derivations of cloned DNAs are referenced or described in the appropriate figure legend. All DNA fragments were extracted from polyacrylamide gels (32) after electrophoretic separation of restriction endonuclease digests of cesium chloride gradient-purified plasmid DNAs. The concentrations of DNA fragments were determined by comparison with known standards in agarose gels. Probe DNA fragments were dephosphorylated with calf intestinal alkaline phosphatase (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) and 5′ end-labeled with [γ-^{32}P]ATP (>5,000 Ci/mmol; New England Nuclear Corp., Boston, Mass.) with T_4 polynucleotide kinase (United States Biochemical Corp., Cleveland, Ohio) to an average activity of 60,000 cpm/ng of DNA fragment. Probe DNA fragments used for the DNase I and methylation studies were prepared from plasmid DNAs digested with a restriction enzyme, 5′ end-labeled as described above, redigested with a second enzyme, and extracted from polyacrylamide gels.

DNA-protein binding assays. DNA-protein binding assays (12–14, 54, 55) were done as follows. A total of 0.5 ng of probe DNA was incubated with 5 μg of protein extract, 3 μg of poly(dI)-poly(dC) or 3 μg of poly(dA-dT)-poly(dA-dT) (P-L Biochemicals, Inc., Milwaukee, Wis., and Pharmacia Fine Chemicals, Piscataway, N.J.), and 5 μg of bovine serum albumin (BSA; Sigma Chemical Co., St. Louis, Mo.) in 20 mM Tris hydrochloride (pH 7.6)–50 mM KCl–1 mM EDTA–0.05% Nonidet P-40–5% (vol/vol) glycerol–5 mM β-mercaptoethanol in a reaction volume of 15 μl for 30 min at 25°C. The reaction mixtures were electrophoretically separated in a 4% polyacrylamide gel (acrylamide-bisacrylamide [29:1]) in 40 mM Tris borate–1 mM EDTA (pH 8.0) at 10 V/cm. The gels were dried and exposed to XS film (Eastman Kodak Co., Rochester, N.Y.) for 12 to 16 h.

DNase I protection assays. Probe DNAs (1 ng), labeled at the 5′ terminus of the coding or noncoding strand, were reacted with either 10 μg of nuclear extract or 10 μg of bovine serum albumin and 7.5 μg of poly(dI)-poly(dC) or 5.0 μg of poly(dA-dT)-poly(dA-dT) in a 30-μl total volume for 30 min at 25°C. The reactions were adjusted to 2.5 mM $MgCl_2$, and DNase I (Cooper Biomedical) was added to 2.5 μg/ml for 2.0 min. The digestion was terminated by the addition of EDTA to 5.0 mM, and the reaction mixtures were electrophoretically separated in a 4% nondenaturing polyacrylamide gel (54). After autoradiography to localize the DNA-protein complexes, the bound and free DNAs were eluted in 0.5 M ammonium acetate–1 mM EDTA–0.1% sodium dodecyl sulfate. The eluted DNA was extracted sequentially with phenol-chloroform-isoamyl alcohol (24:24:1) and chloroform-isoamyl alcohol (24:1) and was precipitated with ethanol in the presence of carrier tRNA (50 μg/ml). The precipitated DNA was suspended in 90% formamide–10 mM EDTA, and equal counts were separated in 8 to 12% sequencing gels. For selected experiments, the appropriate chromatographic fraction was substituted for the crude nuclear extract in the binding reaction, as described in the figure legends.

Column chromatography. Double-stranded calf thymus DNA-cellulose (Sigma) was suspended in 2.0 M KCl in buffer A (40 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.6; 5°C], 20% [vol/vol] glycerol, 1 mM dithiothreitol, 1 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride). Nuclear extract (20 mg in 4 ml of 0.1 M KCl in buffer A) was applied to a 1-ml column equilibrated with 0.1 M KCl in buffer A. The column was washed with 10 column volumes of 0.1 M KCl in buffer A, and the proteins were eluted with an 18-ml linear gradient of 0.1 to 0.5 M KCl and then with five column volumes of 1.0 M KCl in buffer A. Fractions of 500 μl were collected and assayed for protein concentration by use of assays from Bio-Rad Laboratories (Richmond, Calif.) and for KCl concentration by conductivity. Fractions were assayed for specific DNA-binding activity, as detailed in the legend to Fig. 4. A 2-ml DEAE-Sepharose CL6B (P-L Biochemicals and Pharmacia) column was equilibrated with 0.1 M KCl in buffer B (40 mM Tris hydrochloride [pH 7.9; 5°C], 20% glycerol, 1 mM dithiothreitol, 1 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride). Nuclear extract (15 mg) was dialyzed for 5 h against 250 volumes of 0.1 M KCl in buffer B and applied to the equilibrated column. The column was washed with 10 column volumes of the same buffer and step eluted with 5 column volumes each of 0.25, 0.5, and 1.0 M KCl in buffer B. Fractions of 500 μl were collected and assayed for protein concentration. Fractions containing the αH1- (flowthrough) and αH2-αH3 (0.25 M KCl elution step)-binding activities were identified by assaying an equal volume (2.5 μl) of alternate fractions with 48α27R or 70α4R DNA probes, in DNA-binding buffer adjusted to a final KCl concentration of 50 mM. Fractions containing the peak activity of αH1 or αH2-αH3 were used for the methylation interference and protection assays.

For selected experiments, fractions containing αH1 and αH2-αH3 were concentrated by precipitation with ammonium sulfate (50% saturation) and suspended in buffer B at 20% of the original fraction volume.

Methylation interference assays. The procedures for the methylation interference assays were similar to those described elsewhere (15, 53, 54). DNA probes were partially methylated in 200 μl of 50 mM sodium cacodylate (pH 8.0)–1 mM EDTA by the addition of 1 μl of dimethyl sulfate (Aldrich Chemical Co., Inc., Milwaukee, Wis.) at 20°C to yield approximately one modification per DNA molecule (32). The reaction was stopped by the addition of 50 μl of 1.5 M sodium acetate (pH 7.0)–1.0 M β-mercaptoethanol–100 μg of poly(dI)-poly(dC) per ml. The DNA was precipitated with ethanol and suspended in 10 mM Tris hydrochloride (pH 7.6)–1 mM EDTA.

The modified probe DNAs (1 ng) were reacted with 8 μl of the DEAE-Sepharose flowthrough fraction (αH1), 6 μg of poly(dI)-poly(dC), and 10 μg of BSA or 10 μl of the DEAE-Sepharose 0.25 M KCl elution fraction (αH2-αH3); 9 μg of poly(dA-dT)-poly(dA-dT); and 10 μg of BSA in DNA-binding buffer adjusted to a final KCl concentration of 70 to 80 mM in a reaction volume of 35 μl for 30 min.

All samples were subjected to electrophoresis in a 4% nondenaturing polyacrylamide gel to separate the bound from the unbound DNA. After autoradiography, the com-

plexed and free DNA was cut from the gel and recast in a 0.8% agarose–Tris acetate gel and electroeluted onto paper (NA45; Schleicher & Schuell, Inc., Keene, N.H.). The DNA was recovered from the paper as specified by the manufacturer, extracted sequentially with phenol-chloroform (1:1) and then with chloroform, and precipitated with ethanol in the presence of carrier tRNA (50 μg/ml).

To reveal the methylated purines, the DNA was suspended in 43 μl of 20 mM sodium acetate–1 mM EDTA and cleaved (14). The DNA was precipitated with ethanol, suspended in 90% formamide–10 mM EDTA, and separated in 8 to 12% sequencing gels.

Cross-linking of proteins to their DNA-binding sites. The *Pvu*II fragments (0.6 pmol) of pRB606, pRB607, and pRB3430 containing the 48α27R, 29α0R, and 70α4R DNA fragments, respectively, were boiled for 5 min in the presence of 1.0 pmol of the M13 universal primer and 1.0 pmol of the M13 reverse primer and allowed to cool slowly to room temperature. The hybridized primers were extended by *Escherichia coli* DNA polymerase I Klenow fragment in the presence of 50 μM each of dGTP, dATP, and 5-bromo-2′-deoxyuridine-5′-triphosphate (Sigma) and 2.5 μM [α-^{32}P]dCTP (800 Ci/mM; New England Nuclear) (6). The reaction mixtures were digested with *Eco*RI and *Hin*dIII, and the αTIC-containing fragment probes and control DNA probes were isolated from polyacrylamide gels. Labeled probe DNA (1.0 ng) was reacted for 30 min with either 30 μg of the DEAE-Sepharose CL6B fraction containing αH1 and 5 μg of poly(dI)-poly(dC) or 40 μg of the fraction containing αH2-αH3 and 7.5 μg of poly(dA-dT)-poly(dA-dT) in a total reaction volume of 30 μl. The reaction mixtures, in polypropylene tubes (12 by 75 mm) covered with Saran wrap, were irradiated with UV light (254 nM) at an intensity of 340 μW/cm^2 (6). Irradiated samples were brought to 10 mM $CaCl_2$, digested with 3.5 μg of DNase I (Cooper Biomedical)–1 U of micrococcal nuclease (Worthington Diagnostics, Freehold, N.J.) for 30 min at 37°C and subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The gels were silver stained, dried, and autoradiographed; or the resolved proteins were electrically transferred to nitrocellulose sheets and autoradiographed.

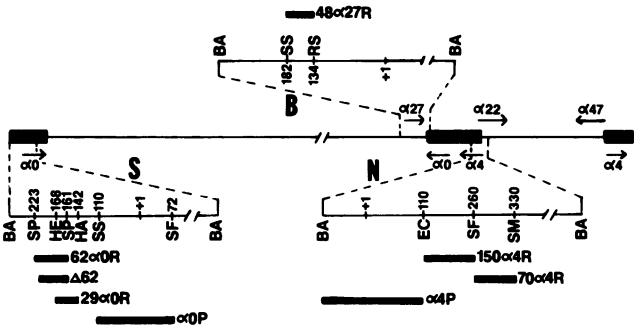

FIG. 1. Schematic diagram of the HSV-1 genome and the location of the α-gene DNA fragments used in this study. The thin lines represent unique sequences of the long and short components, whereas the solid rectangles represent the terminal sequences ab and ca internally repeated as the inverted sequences b′a′c′ (47). The arrows indicate the location and direction of transcription of the five α genes 4, 0, 27, 22, and 47 (28, 57). Expanded scales of *Bam*HI fragments S, B, and N illustrate the locations of the promoter-regulatory domains of the α0, α27, and α4 genes, respectively (29–31). The α-gene DNA fragments (thick lines) cloned and tested in these studies were as follows: 62α0R, the 62-bp *Sph*I fragment of *Bam*HI-S containing the αTIC site at position −223 of the α0 gene cloned in the *Sph*I site of pUC19 as pRB3558; Δ62, the 62-bp *Sph*I fragment of pRB3558 cloned in the *Sma*I site of pUC9 as pRB3429. This fragment lacks five nucleotides from the 5′ and 3′ termini of the 62α0R fragment. 29α0R (pRB607), 48α27R (pRB606), 150α4R (pRB3431), and 70α4R (pRB3430) have been described previously (23, 25) and contain the αTIC homologs from the α0, α27, and α4 (promoter-proximal and promoter-distal) genes, respectively. α0P (pRB3563) and α4P (pRB3353) contain the promoter domains of the α0 and α4 genes, respectively (24). Restriction enzymes are abbreviated as follows: SP, *Sph*I; HE, *Hae*III; HA, *Hae*II; SS, *Sst*II; SF, *Sfa*NI; BA, *Bam*HI; RS, *Rsa*I; EC, *Eco*RI; SM, *Sma*I.

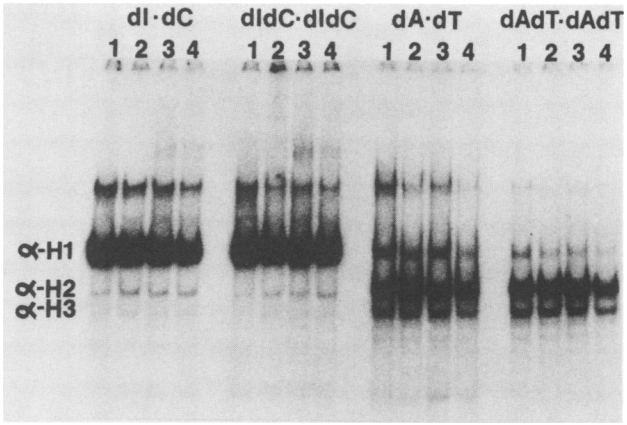

FIG. 2. Autoradiographic image of labeled 48α27R DNA fragments reacted with the HeLa cell nuclear protein extract in the presence of various synthetic DNA polymers. The DNA protein gel electrophoresis assay was done as described in the text, except that increasing amounts of the different competitor nucleic acids indicated at the top of the gel were included in the reactions as follows: lanes 1, 1.3 μg; lanes 2, 2.6 μg; lanes 3, 3.9 μg; lanes 4, 5.2 μg. The positions of the DNA-protein complexes are indicated as αH1, αH2, and αH3. Analyses of the unmarked complex that migrated more slowly than αH1 (see text and Fig. 6 and 9) indicated that it contains both αH1 and αH2-αH3.

RESULTS

Characteristics of DNA fragments used for analyses of DNA-binding proteins specific for the αTIC homolog. The genomic locations of the α genes of HSV-1 and of the DNA fragments from the α4, α0, and α27 genes tested in this study are shown in Fig. 1. Of these fragments, one set contained either the intact (29α0R, 48α27R, and 70α4R) or the partial (62α0R, Δ62, 150α4R) αTIC sites which are required for the *trans*-activation of α genes by the HSV-1 αTIF (23, 29–31, 34, 40, 49). Two of these fragments, the 29-base-pair (bp) DNA fragment from the α0 gene (29α0R) and the 48-bp fragment from the α27 gene (48α27R), were previously shown (23, 34) to confer in vivo α-gene regulation on chimeric indicator genes containing the promoter (−110 to +33) of the α4 gene. The other set of fragments derived from the promoter domains of some α genes (α0P, α4P) were used in competition studies but did not contain homologs of the αTIC sites (25, 29–31).

αH1 and αH2-αH3 protein-DNA complexes are detected in the presence of different synthetic DNA competitors. The electrophoretic mobilities of the DNA-protein complexes formed by the labeled 48α27R DNA fragment with proteins contained in mock-infected HeLa cell nuclear extracts in the presence of different deoxyoligonucleotide competitors are shown in Fig. 2. The results indicate that the homopolymer

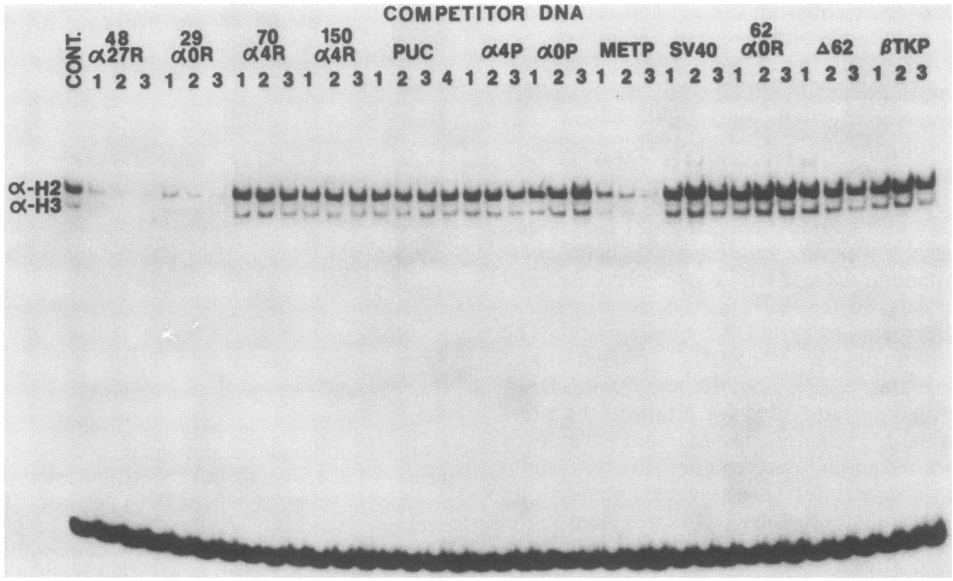

FIG. 3. Autoradiographic image of labeled 48α27R DNA reacted with the HeLa cell nuclear protein extract in the presence of various unlabeled competitor DNA fragments. Labeled 48α27R fragments (0.5 ng) were reacted with nuclear extracts from HeLa cells in the presence of 3.2 μg of poly(dA-dT)-poly(dA-dT) and unlabeled competitor DNA fragments indicated at the top of the gel in molar concentrations of 0.3 pM (lanes 1), 0.6 pM (lanes 2), 0.9 pM (lanes 3), 1.4 pM (lane 4), for a molar excess of competitor over probe DNA of 20, 40, 60, and 90 times, respectively. 48α27R, 29α0R, 70α4R, 150α4R, α4P, α0P, 62α0R, and Δ62 were described in the legend to Fig. 1. The remaining competitor DNAs were as follows: PUC, the 175-bp *Pvu*II-*Hin*dIII fragment from pUC9; METP, the 216-bp *Sac*I-*Bgl*II fragment derived from pMK (3) and containing the mouse metallothionein gene promoter domain from positions +64 to −152; SV40, the 200-bp *Sph*I-*Hin*dIII fragment derived from pSV2Neo (27) and containing the SV40 early promoter and approximately 30% of the promoter-proximal 72-bp enhancer element; βTKP, the 160-bp *Bgl*II-*Bam*HI fragment of LS119/109 (35) and containing the βTK gene promoter domain from positions +50 to −110; CONT, reaction in the absence of additional competitor DNA. αH2 and αH3 indicate the locations of the respective 48α27R-protein complexes.

poly(dA)-poly(dT) and the alternating copolymer poly(dA-dT)-poly(dA-dT) effectively competed with the labeled DNA probe for the DNA-binding activity designated αH1, but not for the DNA-binding activity designated αH2-αH3. In contrast, the homopolymer poly(dI)-poly(dC) or the alternating copolymer poly(dI-dC)-poly(dI-dC) competed for the αH2-αH3 but not for the αH1 DNA-binding activities. In these competition studies, we noted a DNA-protein complex with an electrophoretic mobility slower than that formed by αH1. The DNA-binding activity in this complex has properties similar to those of the αH1 activity and is identified and discussed in more detail later in the text. The results of these studies are consistent with the expectations that the αH1 and αH2-αH3 DNA-binding activities recognize different sequences within the α27-gene DNA fragment.

Specificity of proteins contained in αH2-αH3 protein-DNA complexes. In the competition studies illustrated in Fig. 3, labeled 48α27R DNA was reacted with nuclear extracts of mock-infected cells in the presence of poly(dA-dT)-poly(dA-dT) and increasing molar quantities of unlabeled competitor DNA fragments derived from the promoter domains of various genes. Both the 29α0R and the homologous 48α27R DNA fragments efficiently competed for the αH2-αH3 proteins. In contrast, equimolar concentrations of the plasmid vector sequences (PUC) and the HSV-1 DNA fragments containing the α4 regulatory (70α4R and 150α4R) or promoter (−110 to +33; α4P) domains, the α0 gene promoter (−110 to +72α0P), or the βTK gene promoter domain (−110 to +50; βTKP) did not effectively compete for the αH2-αH3 DNA-binding activities. The mouse metallothionein promoter, which failed to compete for the αH1 DNA-binding activity (26), did compete for the αH2-αH3 proteins (Fig. 3). In contrast, the simian virus 40 (SV40) early region, which competed efficiently for the αH1 proteins, failed to compete with the 48α27R fragment for the αH2-αH3 proteins (26). These results indicate that the αH1 and αH2-αH3 proteins recognize specific sequences located within the αTIC sites of the 48α27R and 29α0R HSV-1 DNA fragments and within the promoter domain of the mouse metallothionein gene.

Results of the αTIC site in the 29α0R fragment also support this conclusion. Although the 29α0R fragment effectively competed with the 48α27R DNA fragment for the αH2-αH3 proteins (Fig. 3), it failed to form demonstrable DNA-protein complexes with the αH2-αH3 DNA-binding activities in crude nuclear extracts (data not shown). Another fragment, 62α0R, contained the second distal αTIC homolog of the α0 gene. This homolog is nearly identical to the αTIC homolog of the 29α0R fragment, but it lacks the five 5′ nucleotides (CCGTG) present in the αTIC sequence of the 29α0R fragment. The 62α0R fragment does not compete for the αH2-αH3 proteins (Fig. 3) but does compete for and bind the αH1 proteins (data not shown). Deletion of the five nucleotides (GCATG; fragment Δ62) adjacent to those missing in the 62α0R fragment abolished the ability of this fragment to compete for the αH1 protein (data not shown). These results suggest that the DNA sequences critical for the binding of αH1 and αH2-αH3 either overlap or are directly juxtaposed within this αTIC site of the α0 gene.

At the molar concentrations tested, the 70α4R DNA fragment did not significantly compete with the 48α27R DNA fragment for the αH2-αH3 DNA-binding activity in the crude extract. As shown below, this fragment does bind proteins in chromatographic fractions, with characteristics similar to those of the αH2-αH3 proteins.

Chromatographic separation of host proteins forming the αH1 and αH2-αH3 DNA-protein complexes. The separation of the αH1- from the αH2-αH3-binding activities was obtained by several chromatographic procedures. In the one illustrated in Fig. 4, the nuclear extract from mock-infected cells was chromatographed on a column of calf thymus double-stranded DNA-cellulose. αH1 eluted with most of the DNA-binding proteins at 150 to 300 mM KCl, whereas the αH2-αH3 proteins eluted at 300 to 500 mM KCl (Fig. 4). Separation of the αH1 from the αH2-αH3 proteins was also obtained by DEAE-Sepharose and heparin-agarose chromatography. The separation of αH1 from αH2-αH3 on heparin-agarose was qualitatively similar to that seen on DNA-cellulose, but αH1 and αH2-αH3 eluted at 250 and 500 mM KCl, respectively. On DEAE-Sepharose, the αH1 proteins flowed through the column at 100 mM KCl, whereas the αH2-αH3 proteins eluted from the column at 250 mM KCl. In none of these studies were we able to separate the αH2 from the αH3 activities.

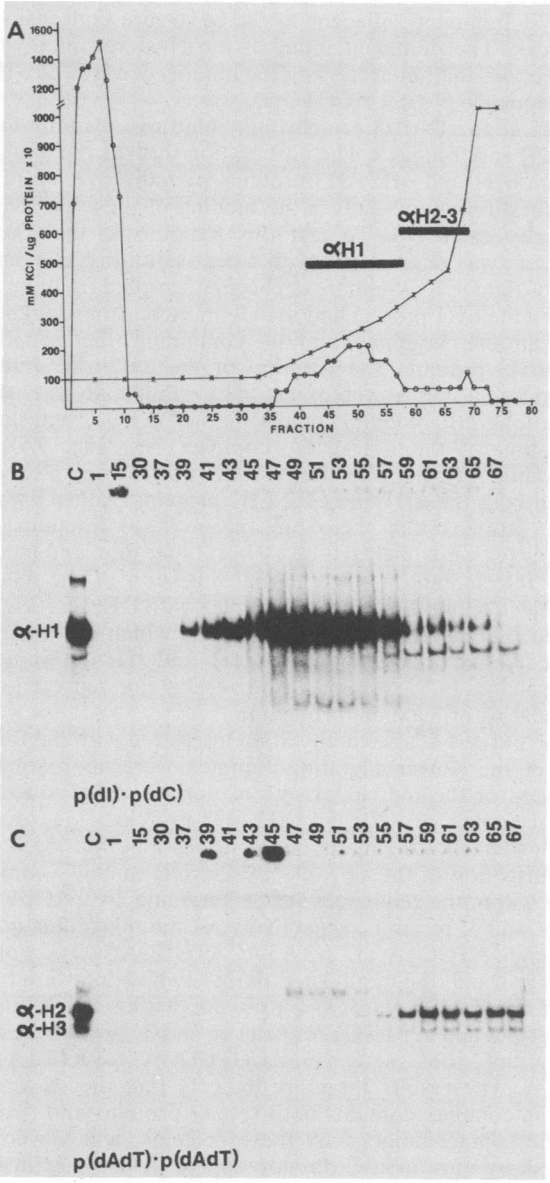

FIG. 4. Separation of αH1 and αH2-αH3 proteins by chromatography on DNA-cellulose. (A) Elution profile of the proteins. Symbols: X, KCl as determined by conductivity readings of the fractions: O, micrograms of protein per 500-μl fraction. The fractions containing the peak activities of the αH1 and αH2-αH3 proteins are indicated by the filled bars. (B and C) Autoradiographic image of the labeled 48α27R probe reacted with selected fractions from DNA-cellulose chromatography in the presence of 3.0 μg of poly(dI)-poly(dC) for detection of the αH1 protein or 4.5 μg of poly(dA-dT)-poly(dA-dT) for detection of the αH2-αH3 proteins, respectively. Binding assays were done as described in the text with 2.5 μl of the fractions indicated at the top of the gel. The lanes designated C contained the complex formed by the labeled probe with proteins in 5.0 μg of crude HeLa cell nuclear extract in the presence of the respective synthetic polymers.

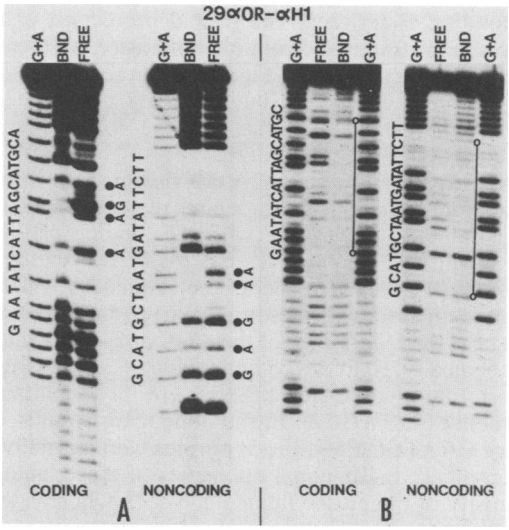

FIG. 5. Analyses of the αH1-binding site in the 29α0R DNA fragment by methylation interference and DNase I protection tests. (A) The 29α0R DNA fragment was labeled at the 5′ terminus of the coding or noncoding strands, partially methylated, and reacted with the DEAE chromatographic fraction containing αH1. The positions of the methylated purines which partially or completely precluded the binding of the αH1 protein are indicated by the filled circles to the right of the panel. Abbreviations: BND, DNA extracted from the αH1-bound complex; FREE, DNA excluded from the αH1 complexes. (B) The 29α0R DNA fragment, labeled at the 5′ terminus of the coding or noncoding strands, was reacted with 10 μg of bovine serum albumin (FREE) or 10 μg of HeLa cell nuclear extract (BND) in the presence of 7.5 μg of poly(dI)-poly(dC) and partially digested with DNase I. The maximum protected domain is indicated by the vertical line. The relevant nucleotide sequence is indicated to the left of each reaction set. G+A, Purine-specific sequencing reaction. The procedures used in this and in the other studies for which the results are shown in Fig. 6 and 7 are described in the text.

Binding sites of αH1 proteins within the regulatory domains of α genes. We have shown previously (26) by DNase I protection studies that the αH1 protein binds across the αTIC consensus homolog that is contained within the 48α27R DNA fragment. A partial DNase I digestion of the 29α0R DNA fragment labeled at the 5′ terminus of the coding or the noncoding strands and incubated with nuclear extracts of mock-infected cells is shown in Fig. 5B. The vertical line in Fig. 5B designates the maximum boundaries of the DNase protection. For both noncoding and coding strands, the 5′ termini of the indicated sequences (GCAT GCTAAT and CATTAGCATG, respectively) were more strongly protected from DNase I digestion than were the

corresponding 3′ termini. The partial protection of the 3′ domain can be inferred from the intensity of the bands outside of the vertical lines (Fig. 5B). The regions totally or partially protected by αH1 in the 29α0R fragment correspond to the equivalent domain of the 48α27R fragment protected by the αH1 protein (26).

In Fig. 5A, 6A and D, and 7A are shown the positions of the methylated guanine and adenine residues which interfered completely or partially with the binding of the αH1 protein to the 29α0R, 48α27R, and 70α4R DNA fragments, respectively. In these fragments the sequences which contain the methylated purines are nearly identical and are located near the 5′ terminus of the αTIC consensus sequence (30). Specifically, (i) in the coding strand of 29α0R (Fig. 5A), the interfering methylated purines were those contained in the sequence ATTAGCA. In the noncoding strands, in the sequence GCATGCTAA, the 5′ purines were partially interfering (compare the unequal intensities of these bands with the intensity of the purine immediately 5′), whereas the two methylated adenines 3′ completely interfered with the binding of αH1. (ii) In the coding strand of 48α27R, the methylated purines which interfered with the binding of αH1 (Fig. 6A) were in the sequence ATTAGCA. In the noncoding strand (Fig. 6D), the three adenines in the sequence ATGC TAA nearly completely blocked binding, whereas the methylated guanine was only mildly interfering. (iii) In the coding strand of the 70α4R fragment (Fig. 7A), the methylated purines which interfered with the binding of αH1 were contained in the sequence GCATGCTAACG. In the noncoding strand (Fig. 7A), the methylated purines which interfered with the binding of αH1 were in the sequence GTTAG CATG.

The dominant motif of the αH1-binding site is the 5′ domain GyATGnTAAT of the consensus sequence for the αTIC site (30). The near identity of the purine contact points of the three αTIC sites with the αH1 proteins and the results of the previously described competition experiments (26) indicate that the αH1 proteins which bind to the αTIC sites of the α0, α27, and α4 genes are identical.

Binding site of αH2-αH3 proteins within the 48α27R DNA fragment. The 48α27R DNA fragment sequences protected from DNase I digestion by the αH2-αH3 proteins contained in the DEAE-Sepharose chromatographic fraction are shown in Fig. 6G and H. In this instance, the electrophoretically separated αH2 and αH3 protein-DNA complexes were each individually excised from the nondenaturing gels and processed as described above. The vertical lines shown in Fig. 6G and H represent the region that is more highly protected by the αH2 protein. The sequence protected from DNase I digestion by the αH2 proteins overlapped and extended 3′ to the αTIC homolog sequence protected by the αH1 protein and reported previously (26) (Fig. 6G and H). Identical results were obtained with the αH3 protein-48α27R DNA complex (data not shown), suggesting that the binding sites of the αH2 and αH3 proteins are identical and different from that of the αH1 proteins.

In Fig. 6B, C, E, and F are identified the methylated purines which completely or partially interfered with the binding of the αH2 and αH3 proteins to the 48α27R DNA fragment. Specifically, in the coding strand, the methylated purines which interfered with the binding of αH2 (Fig. 6B) or αH3 (Fig. 6C) resided within the sequence ACGTGG. The 3′-terminal guanine only partially interfered with the binding of these proteins. In the noncoding strand, the methylated purines which interfered with the binding of αH2 (Fig. 6E) or αH3 (Fig. 6F) resided within the sequence GCCACGTG. For both proteins, the 5′-terminal guanine was only partially interfering.

Comparison of the methylation interference patterns for the 48α27R-αH2 and αH3 protein-DNA complexes reinforces the conclusion concerning the identity of the binding sites. The purines critical for the binding of these proteins reside near the center of the pallindromic sequence 5′-AATACATGcCACGTACT-3′ at the 3′ terminus of the 48α27R fragment, adjacent to the αH1 core consensus sequences. The dominant motif of the clustered methylated nucleotides that interfere with the binding of αH2 and αH3 is the sequence 5′-GCCACGTG-3′.

αH1 and αH2-αH3 proteins may bind concurrently to the 48α27R and fragment. The observation that the binding sites for the αH1 and αH2-αH3 proteins partially overlap raised the question as to whether the binding of these proteins is mutually exclusive. To test this hypothesis, the 48α27R fragment was reacted first with a near saturating concentration of the DEAE-Sepharose chromatographic fraction containing the αH1 protein and then with increasing amounts of the chromatographic fraction containing the separated αH2-αH3 protein. The addition of the αH2-αH3 fraction generated a DNA-protein complex similar to the slow-electrophoretic-mobility complex shown in Fig. 2 and slower than that of the αH1-DNA complex (see Fig. 9A). To determine whether the larger DNA-protein complex contained both the αH1 and αH2-αH3 proteins, rather than the other proteins that were present in these fractions, the complexes were analyzed by DNase I digestion and methylation interference assays. Partial digestion of the 48α27R low-electrophoretic-mobility complex with DNase I (Figures 6I and J) generated a protected domain which encompassed the individual boundaries of the αH1- and αH2-αH3-protected domains. In Fig. 6I and J, the vertical line represents the boundaries of maximal protection.

The purines whose methylation interfered with the formation of the slower-migrating complex were approximately the sum of the individual sets of purines that formed the contact points of αH1 and αH2-αH3 proteins (Fig. 6K and L). Specifically, the methylated purines that interfered with the formation of the slow-electrophoretic-mobility complex were within the sequences ATGCTAA and GCCACGTG at the 5′ and 3′ termini, respectively, of the noncoding-strand sequence ATGCTAATTAAATACATGCCACGTG. For the coding strand, the methylated purines which interfered with the formation of this complex were within the sequence ACGTGG and ATTAGCA at the 5′ and 3′ termini, respectively, of the sequence ACGTGGCATGTATTTAATT AGCA. The results therefore indicate that the slower-migrating complex contains both sets of proteins and that the purine contact points with these proteins, which were resolved by this assay, do not appear to be significantly altered.

Binding of the 70α4R DNA fragment to proteins in fractions containing αH2-αH3 proteins. Experiments similar to those described above were also done with the 70α4R fragment. As in the case of the binding of αH2-αH3 to the 48α27R DNA fragments, the cytosine-rich deoxyoligonucleotides effectively competed with the 70α4R fragment for the binding to these proteins (data not shown). In this instance, however, the methylated purines which interfered with the binding of proteins in the chromatographic fraction containing the αH2-αH3 proteins were different from those of the 48α27R fragment (Fig. 6B, C, E, and F and Fig. 7B). Thus, the methylated purines which interfered with the binding of αH2 to the coding strand of the 70α4R fragment reside within

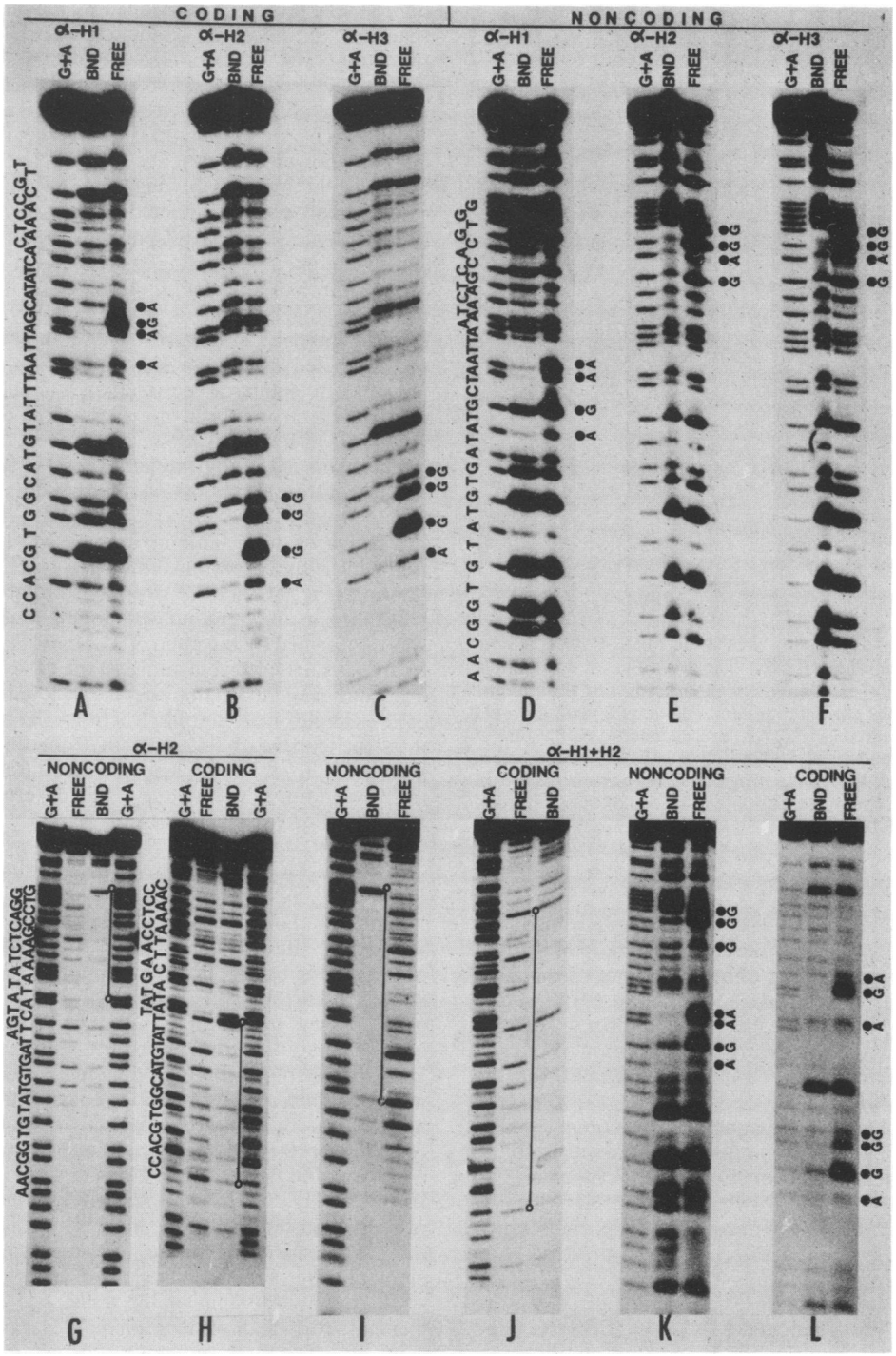

FIG. 6. Analyses of the αH1- and αH2-αH3-binding sites in the 48α27R DNA fragment by methylation interference and DNase I protection assays. The positions of the methylated purines which partially or completely precluded the binding of αH1 (A and D), αH2 (B and E), and αH3 (C and F) on the coding and noncoding strands are shown. (G and H) DNA sequences protected from DNase I digestion by αH2 on the coding and noncoding strands. (I and J) Positions of DNA sequences protected from DNase I digestion by a reconstituted mixture of αH1 and αH2-αH3 (identified as αH2 in the figure). (K and L) Positions of the methylated purines which interfered with the binding of a reconstituted mixture of αH1 and αH2-αH3. In all of these experiments, the DEAE-Sepharose chromatographic fractions were the source of separated αH1 and αH2-αH3. The reactions shown in panels B, C, E, and F were done on αH2 or αH3 protein-DNA complexes that were individually excised from the gels. The position of interfering methylated purines is shown by solid circles. The minimal domains protected from DNase I digestion are indicated by the vertical lines. BND and FREE are as defined in the legend to Fig. 5.

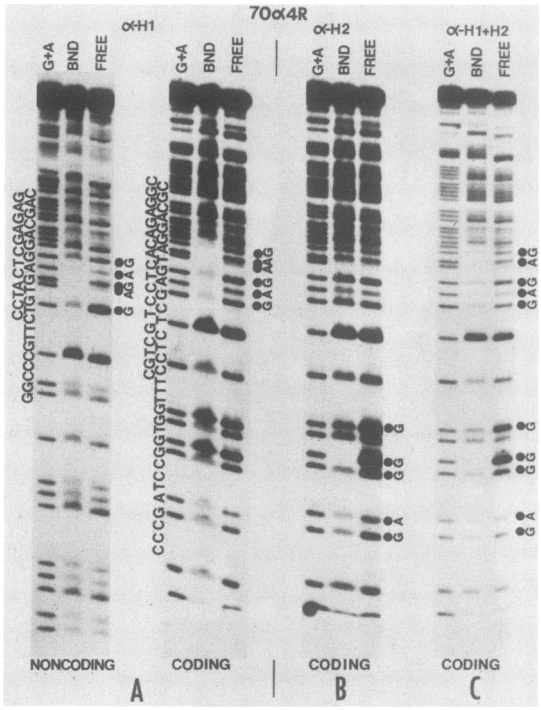

FIG. 7. Analyses of the binding sites in the 70α4R DNA fragment by methylation interference assays. The locations of methylated purines which partially or completely interfered with the binding activities in the DEAE chromatographic fractions containing αH1 (A) or αH2-αH3 (indicated as αH2 [B]) or a reconstituted mixture of both activities (indicated as αH1+H2 [C]) are shown. The relevant nucleotide sequence of the 70α4R fragment is shown to the left of each panel. The solid circles indicate the positions of the interfering methylated purines. BND and FREE are as defined in the legend to Fig. 5.

the sequence GATCCGGTGG. Only the fourth purine in this sequence completely blocked binding. The possibility that these proteins represent either entirely distinct or modified αH2-αH3 proteins is supported also by the observation that unlabeled 70α4R competed with the homologous labeled fragment for this binding activity (data not shown) but did not compete with the labeled 48α27R fragment for the αH2-αH3 proteins (Fig. 3).

Both αH1 and the activity present in the fractions containing αH2-αH3 were capable of binding simultaneously to the 70α4R DNA fragments in reconstitution experiments (Fig. 9B). The methylated purines which interfered with the formation of this complex were the sum of the purines which interfered with the binding of the individual proteins (Fig. 7C). Thus, the methylated purines which interfered with the formation of the slow-migrating complex (Fig. 9B) were contained within the sequences GATCCGGTGG and GCATGCTAACG representing the 5′ and 3′ termini, respectively, of the sequence GATCCGGTGGTTTCCGCTTCCG TTCCGCATGCTAACG. The patterns of partial and complete interference observed with αH1 and αH2 were similar to those obtained with the individual proteins.

Apparent molecular weights of proteins binding to the 29α0R, 48α27R, and 70α4R DNA fragments cross-linked to labeled deoxynucleotides. The 29α0R, 48α27R, and 70α4R DNA fragments labeled with [α-^{32}P]dCTP and 5-bromo-2′-deoxyuridine-5′-triphosphate were reacted with the chromatographic fractions containing either the αH1 or the αH2-αH3 proteins. The incorporation of 5-bromo-2′-deoxyuridine-5′-triphosphate did not significantly affect the binding of these proteins, as assayed by gel electrophoresis. The reaction mixtures were irradiated with UV light for various time intervals to cross-link the protein-DNA complexes, digested with nucleases, and electrophoretically separated in denaturing polyacrylamide gels. The lanes containing the 29α0R-αH1 protein complex, which was irradiated for 15 to 60 min, showed the presence of a labeled protein band with an apparent molecular weight of 110,000 (Fig. 8A). This protein band was not detected in lanes containing the control 91-bp fragment derived from pUC9 and reacted with the same fraction. The labeling of this protein was not affected in the presence of a 60-fold molar excess of the 91-bp control fragment (Fig. 8B, lane 4 versus lane 2) but was reduced in labeling intensity in the presence of an equimolar amount of the unlabeled 29α0R DNA fragment (Fig. 8B, lane 3 versus lane 2). In similar experiments, a labeled band with an identical electrophoretic mobility was also obtained with the 70α4R DNA fragment probe. Moreover, competition experiments with a 150-fold molar excess of the 70α4R and the 91-bp control fragments yielded results similar to those obtained with the 29α0R fragment (data not shown).

Results of cross-linking studies on the 48α27R DNA fragment with the proteins in the αH2-αH3 chromatographic fraction (Fig. 8C and D) revealed the labeling of a specific band with an apparent molecular weight of 64,000. This band was not formed by the control 91-bp pUC9 fragment that was treated in an identical fashion. In the case of the 70α4R DNA fragment (Fig. 8D), the specific labeled protein from the αH2-αH3 chromatographic fraction migrated slightly slower than the corresponding protein from this fraction labeled with the 48α27R DNA fragment. The differential electrophoretic mobilities could be due to the difference in the number of residual cross-linked nucleotides.

DISCUSSION

Differentiation and binding sites of host proteins interacting with αTIC sites. The five HSV-1 α genes contain in their promoter-regulatory domains homologs of a sequence designated αTIC which is required for the induction of these genes by the structural component of the virus designated αTIF. Our studies centered on the proteins that bind to small DNA fragments containing the single αTIC homolog of the α27 gene, one of the three homologs of the α0 gene, and the most distal of the homologs of the α4 gene. We have reported previously (26) that the host proteins designated αH1 bind and protect the αTIC site from DNase I digestion. In this study, we have extended the analysis of the αH1 proteins and the additional host cell proteins which bind to the αTIC site and to the sequences juxtaposed to this element. The salient conclusions of this study are as follows.

(i) The environment of the αTIC site contains sequence elements for which we identified at least two sets of proteins (Fig. 10). The first sequence, corresponding to the αTIC consensus homolog, binds to a protein designated αH1. The protein cross-linked to labeled deoxynucleotides has an apparent molecular weight of 110,000, but at this time it is not known whether it is a monomeric or a polymeric protein. The dominant motif of the αH1-binding site is in the 5′ domain (GyATGnTAAT) of the αTIC consensus sequence (30) and is similar to the binding site for nuclear factor III (46).

The 70α4R fragment the αTIC homolog is inverted relative to the transcription initiation site of the α4 gene (30). The

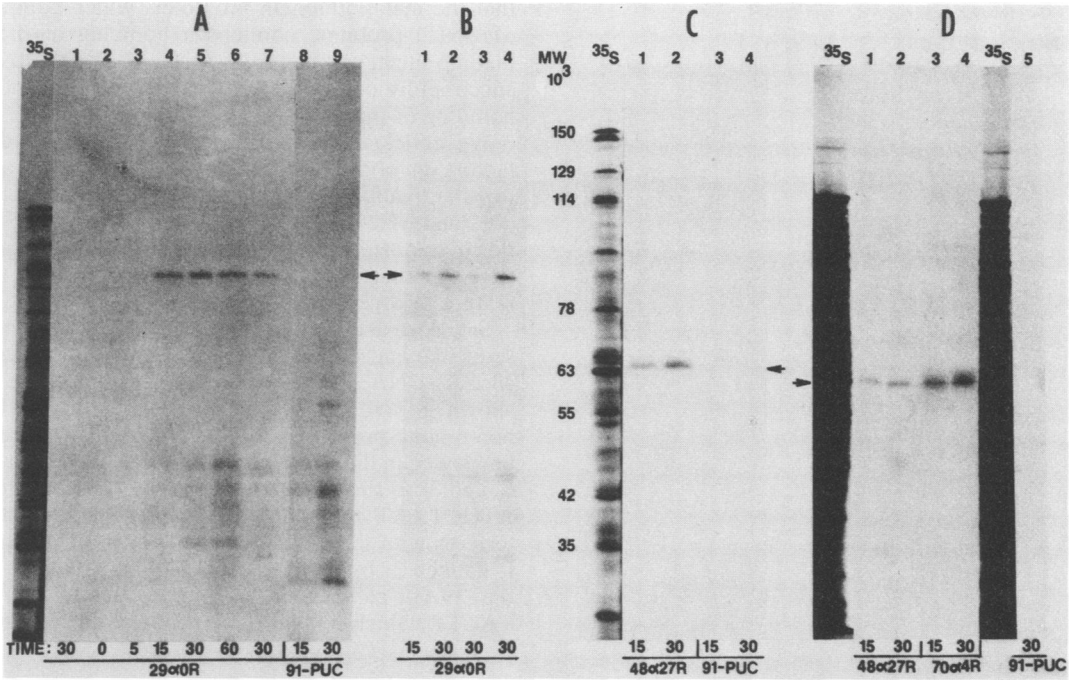

FIG. 8. Autoradiographic images of the proteins labeled by cross-linking to internally labeled DNA probes containing αTIC sites and electrophoretically separated in denaturing gels. In the experiments for which the results are shown here, the labeled probe DNAs were reacted with the DEAE-Sepharose chromatographic fractions containing αH1 (A and B) or αH2-αH3 (C and D) in buffer containing 10 μg of BSA (complete reaction mixture), irradiated with UV light for the time intervals (in minutes) indicated at the bottom of the gel, digested with nucleases, and subjected to electrophoresis in denaturing polyacrylamide gels. (A) The electrophoretic mobilities of the proteins cross-linked to the 29α0R fragment in complete mixture (lanes 2 to 6), in the absence of the chromatographic fraction (lane 1), or in the absence of BSA (lane 7). Lanes 8 and 9 show the electrophoretic mobilities of the proteins from the same fraction cross-linked to the control 91-bp fragment from pUC9 in the complete mixture. (B) Electrophoretic mobilities of the proteins cross-linked to the 29α0R fragment in the absence (lanes 1 and 2) or in the presence of a 60-fold molar excess of unlabeled 29α0R (lane 3) or the 91-bp pUC9 (lane 4) competitor DNA fragments. (C) Proteins cross-linked to the 48α27R (lanes 1 and 2) or the 91-bp pUC9 fragments (lanes 3 and 4). (D) Proteins cross-linked to the 48α27R (lanes 1 and 2), the 70α4R (lanes 3 and 4), and the 91-bp fragment from pUC9 (lane 5). The lanes marked with ^{35}S contain an electrophoretically separated lysate of cells labeled at late times postinfection with [^{35}S]methionine for use as a molecular weight (MW) marker based on the apparent molecular weights of HSV-1 proteins (36).

significance of this inversion is not clear. In the in vivo studies on chimeric gene constructs, αTIC sequences were capable of conferring αTIF-dependent induction in an orientation-independent manner (23). The αTIC sequence of the 70α4R fragment was the most distal homolog relative to the α4 gene transcription initiation site and was located in a domain which could serve as the regulatory sequence for both the α4 and the α22 or α47 genes (Fig. 1).

The second sequence, located in the α27 gene (48α27R DNA fragment) and deduced from results of competition studies to be present in the α0 gene, binds to a protein designated αH2-αH3. Except for the electrophoretic mobilities of the DNA-protein complexes, we were not able to demonstrate a difference between the two DNA-binding activities. The αH2-αH3 protein cross-linked to labeled deoxynucleotides had an apparent molecular weight of 64,000. The dominant motif of the αH2-αH3 binding site was ACGTGGCATGT (Fig. 10).

As discussed below, a protein with many attributes of αH2-αH3 but differing in its affinity for DNA sequences was detected in studies of protein binding to the 70α4R fragment, which contains an αTIC site from the α4 gene. For convenience, and pending further analyses of the relationship of this protein to αH2-αH3, we designate this protein αH2′.

(ii) αH1 is common to all the αTIC sequences analyzed to date. However, both αH1 and αH2-αH3 proteins may be required for the induction of α genes. This conclusion rests on the observation that the 48α27R fragment was able to impart α-gene regulation to both α and β promoters, whereas 29α0R was able to impart α-gene regulation only in the case of an α promoter (α4P) (Fig. 1) which contains a portion of the αTIC sequence at its 5′ terminus. While the 48α27R fragment contains an intact αH2-αH3-binding site at its 3′ terminus, the corresponding sequence is only partially present in the 29α0R fragment.

(iii) The position of the αH1 and αH2-αH3 or αH2′ binding sites suggests that the proteins binding to these sequences may interact. In the 48α27R fragment, the binding site for the αH2-αH3 protein overlaps, in part, with the αH1-binding site. Although the 29α0R fragment tested does not have an intact αH2-αH3-binding site, the presence of a binding site adjacent to the αH1-binding site was deduced from the results of competition studies. The analyses described above indicate that the αH2-αH3-binding site in this instance is 5′ to the αH1-binding site and may also overlap in part with an αH1-binding site. In the case of the 70α4R fragment, αH2′, while nearly identical to αH2-αH3 with respect to apparent molecular weight and chromatographic properties, differed from the defined αH2-αH3 proteins with respect to its affinity for particular DNA sequences, as shown by results of the competition and methylation interference studies described above. The binding site for this protein is at the 3′

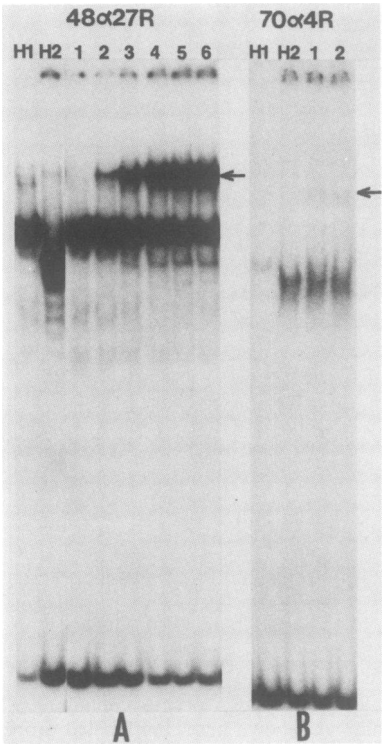

FIG. 9. Autoradiographic images of the 48α27R and 70α4R DNA fragments reacted with the reconstituted extracts containing αH1 and αH2-αH3. (A) The labeled 48α27R DNA fragment was reacted with 9 μg of the chromatographic fraction containing the αH1 protein and increasing amounts of the proteins in the fraction containing the αH2-αH3-binding activities as follows: lanes 1 to 6, 0, 3, 9, 15, 21, and 27 μg, respectively; lanes H1 and H2 contained 9 and 21 μg of the αH1 or the αH2-αH3 fraction, respectively. (B) The 70α4R DNA fragment was reacted with 3 μg of the chromatographic fraction containing αH1 (lane H1), 3 μg of the fraction containing αH2-αH3 (lane H2), or 6 μg of the αH1 and 3 μg of the αH2-αH3 fractions (shown in duplicate in lanes 1 and 2). The arrows show the location of the slow-migrating complex described in the text.

terminus of the fragment, immediately adjacent to the second αTIC sequence located at position −260 from the transcription initiation site of the α4 gene. In all the instances in which we demonstrated the binding of αH1 and αH2-αH3 or αH2′, both sets of proteins were shown to be able to bind concurrently.

(iv) The 70α4R fragment contains several homologs of a sequence, GGAACGG, that is shared with the 48α27R sequence and with the immunoglobulin enhancer. This sequence is located 5′ to the αH1 site in the 48α27R DNA fragment and between the αH1- and the αH2′-binding sites of the 70α4R DNA fragment. A possible interaction between αH2′ and αH1 with a protein which may bind to this sequence is suggested by the sequence motif arrangement and by the observation that the homolog adjacent to the αH2′-binding site is partially protected by αH2′ from DNase I digestion (data not shown). Although we could not show protection of the GA homolog shared by the immunoglobulin enhancer element that is present in the 48α27R fragment, the possibility exists that a required portion of the binding site is in sequences adjacent to this fragment.

(v) Sequences homologous to the αTIC sites have been detected in a variety of constitutively expressed or inducible cellular and viral genes (49). As might be expected, we found that the metallothionein promoter, which competed for the αH2-αH3 proteins, contains a homolog of the αH2-αH3-binding site at positions −90 to −101. The SV40 72-bp repeat successfully competed for the αH1 protein and contains a homolog of this binding site.

We conclude that αH1 and αH2-αH3 are distinct host factors involved in transcriptional regulation and that the *trans*-activation of α genes may involve the function of both αH1 and αH2-αH3 proteins at the αTIC site. Furthermore, the juxtaposition of the αH1- and αH2-αH3-binding sites suggests that interaction of these factors is an essential feature of α-gene regulation.

Sequence arrangement in the 48α27R DNA fragment. The sequence and the implied *cis*-functional arrangement in the 48α27R DNA fragment are of interest because it has been shown in vivo to confer α-gene regulation to both α- and non-α-gene promoters. To understand its function, it must be viewed in the context of the sequences which extend both 5′ and 3′ to the fragment. The center point is the sequence ATGCTAAT containing the purines whose methylation interferes with the binding of the αH1 protein. This sequence is flanked 3′ by the αH2-αH3-binding site contained within the 19-bp pallindromic sequence AAATACATGcCACG TACTT which extends four nucleotides outside of the

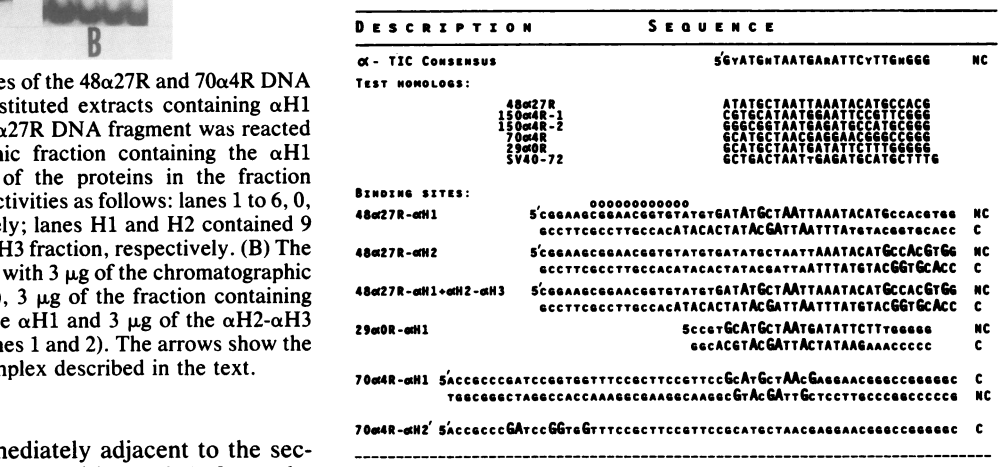

FIG. 10. Summary of the binding sites for the host proteins in the α0, α27, and α4 DNA fragments identified in this study. The top line shows the consensus sequence of the αTIC sites of the HSV-1(F) α genes (30). The list below the consensus sequence includes the nucleotide sequence of the αTIC homologs of α4, α0, and α27 genes tested in this study and a homolog identified in the 72-bp repeat of the SV40 DNA. The entire HSV-1 nucleotide sequence of the 48α27R and the 29α0R fragments is shown, but only the relevant portions of the 70α4R fragment are shown. It should be noted that the last two nucleotides of the 48α27R fragment (GG of the noncoding strand) were donated by the plasmid vector sequence; the authentic nucleotides are AC (29, 30). The large letters identify the purines whose methylation partially or completely interfered with the binding of the respective proteins. The intermediate-sized letters show the domains protected from DNase I digestion. The circles indicate a sequence motif that is homologous to the immunoglobulin gene enhancer element. As noted in the text, the αTIC sequence contained in 70α4R is inverted relative to the transcription initiation site of the α4 gene. Abbreviations: NC, noncoding strand; C, coding strand; Y, pyrimidine; R, purine; N, any nucleotide. The binding site of the αH1 and αH2-αH3 proteins on the 48α27R fragment was derived from data shown in Fig. 6I, J, K, and L. The corresponding patterns for the 70α4R fragment are not shown. The αH1 protein DNase 1 footprint on the 48α27R fragment is from Kristie and Roizman (26).

48α27R fragment. Homologs of portions of this pallindrome (AAATACAT and TACATGCCACG) are present in an inverted orientation at the 5′ domain of the DNA protected from DNase I digestion by αH1. The 5′ end of the αH1-protected domain is juxtaposed to the 12-bp sequence (5′-CGGAACGGTGTA-3′) that is homologous to the immunoglobulin enhancer sequence (7, 10, 51). The proximity of this sequence to the αTIC site, and particularly to the αH2-αH3- and αH2′-binding sites, may be relevant to the overall regulation of α-gene expression. Finally, the 5′-terminal sequence (5′-CGGAAGCGGAA-3′) of this fragment corresponds exactly to the sequence between the purine-protein contact points of the αH1 and αH2′ of the 70α4R DNA fragment and is partially protected from DNase I digestion by αH2′ (data not shown).

This sequence analysis does not take into account other possible matches in which the sequence divergence is greater. The conservation of nucleotide stretches, some of which are within known binding sites, suggests that additional proteins might bind to this fragment. In this context, the role of the αTIC sequence should be viewed as necessary but potentially not in itself sufficient to confer α-gene regulation. Further analyses of the domains of this fragment should establish the role of these sequences in the induction of α genes.

Function of αTIF, αH1, and αH2-αH3 proteins in the induction of α genes. Assessment of the functions of αTIF, αH1, and αH2-αH3 proteins in the induction of α genes in the naturally infected host are hampered by the fact that these studies have been done in cell culture, which may not reflect the transcriptional capabilities of cells in tissues infected by HSV. For example, α genes are expressed constitutively on introduction into animal cells even when the selectable marker is another non-α gene (16, 23, 34, 43, 52). There is no evidence to suggest that the cells in tissues naturally infected by HSV have this capability. Notwithstanding these limitations, the significant observation is that the fusion of αTIC sequences to an α-gene promoter does not significantly increase the expression of the chimeric gene in the absence of αTIF (23). Any model of the mechanism of induction of α genes by αTIF must, however, take into account the recent observation that αTIF forms specific complexes with DNA fragments containing the αTIC site in the presence of αH1, but not in the presence of αH2-αH3 alone or in the presence of competitors which preclude the binding of αH1 (33). The identification of the αH1 and αH2-αH3 proteins and their DNA-binding sites should permit genetic and biochemical analyses of the function of these host proteins.

Comparison of functions and requirements of the *trans*-acting proteins involved in regulation of expression of HSV-1 genes. Until the discovery of αTIF, the α4 protein was the only HSV-1 gene product known to specifically *trans*-activate viral genes. This protein regulates either positively or negatively the expression of HSV genes (for a review, see reference 48). In this instance, however, there is evidence for the binding of α4 to the viral DNA at specific sites (2, 11, 24, 25, 37, 41), even in the absence of other cellular proteins (22).

αTIF differs from α4 in three important respects. First, αTIF appears to induce a specific set of coordinately regulated genes (the α genes). Second, αTIF is a component of the virion, whereas α4 is expressed after infection. Lastly, it does not appear to bind, by itself, to the domains of the genes on which it acts, even though a specific *cis*-acting site is required for *trans*-activation. The first two properties of αTIF are without precedent. With respect to the last property mentioned above, the most notable example of such a *trans*-acting factor is the E1A protein of adenoviruses that has been reported to *trans*-activate adenovirus genes expressed later in infection by mobilization or activation of host transcriptional factors (20, 21).

The molecular mechanisms of the *trans*-activation of α genes by αTIF remain to be established. In a different dimension, there remains the question as to why herpesviruses have evolved a structural component packaged in a different virion compartment than that of the viral DNA to *trans*-activate the first set of genes to be expressed. Fundamentally, the choice of regulatory genes must be governed by biological necessity. Mobilization of transcriptional factors for effective expression of the first transcribed set of genes in other viruses (e.g., adenoviruses, papovaviruses, retroviruses) is accomplished by the presence of specific *cis* sites such as enhancer elements. In such systems, however, initiation of expression of the first set of genes is irrevocable in permissive cells in which transcriptional factors are, by definition, present. In its natural host, HSV is capable of both productive and latent infections, the determinants of which may be, in part, viral. As a *trans*-acting element, αTIF may play a role in determining whether infection is productive or latent. When αTIF is present, it could rapidly effect the initiation of transcription of the α genes containing the αTIC sites, leading to productive infection and obligatory cell death. Conceivably, suppression of α-gene expression is a requirement for the establishment of latency, and in the absence of αTIF, latency-specific genes would have a selective advantage. The biological necessity of such a regulatory system is exemplified in HSV infections of sensory neurons in which the efficient expression of α genes may lead to productive infection rather than to the latent state in which the expression of these genes is not detected uniformly (for a review, see reference 50). We note that in cultured cells, the distance between the site of entry of the virus into the cell and the nucleus is measured in micrometers, whereas the distance between the site of entry at the nerve endings of sensory neurons and the neuronal nucleus harboring latent virus is measured in centimeters. The separation of αTIF and viral DNAs in the virion may well imply that they make their way into the intracellular environment independently, with interesting biological consequences. For example, in a model in which αTIF must interact with αH1 proteins for the induction of α genes (33), the failure of αTIF to reach the nucleus of the sensory neuron could explain the dominance of latency-specific functions over those leading to productive infection (50). The ability to establish latency may well be a requirement for the survival of HSV in their natural hosts.

ACKNOWLEDGMENTS

We thank H. Singh, L. Chodosh, R. Carthew, and P. A. Sharp for the methylation interference and protein-DNA cross-linking protocols and S. L. McKnight, R. Palmiter, and the late G. Khoury for the gifts of plasmids.

This study was supported by Public Health Service grants CA08494 and CA19264 from the National Cancer Institute and by grant MV2-W from the American Cancer Society.

LITERATURE CITED

1. **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.
2. **Beard, P., S. Faber, K. W. Wilcox, and L. I. Pizer.** 1986. Herpes simplex virus immediate early infected cell polypeptide 4 binds

to DNA and promotes transcription. Proc. Natl. Acad. Sci. USA **83**:4016–4020.
3. **Brinster, R. L., H. Y. Chen, R. Warren, A. Sarthy, and R. D. Palmiter.** 1982. Regulation of metallothionein-thymidine kinase fusion plasmids injected into mouse eggs. Nature (London) **296**:39–42.
4. **Bzik, D. J., and C. M. Preston.** 1986. Analysis of DNA sequences which regulate the transcription of herpes simplex virus immediate early gene 3: DNA sequences required for enhancer-like activity and response to *trans*-activation by a virion polypeptide. Nucleic Acids Res. **14**:929–943.
5. **Campbell, M. E. M., J. W. Palfreyman, and C. M. Preston.** 1984. Identification of herpes simplex virus sequences which encode a *trans*-acting polypeptide responsible for stimulation of immediate early transcription. J. Mol. Biol. **180**:1–19.
6. **Chodosh, L. A., R. W. Carthew, and P. A. Sharp.** 1986. A single polypeptide possesses the binding and transcription activities of the major late transcription factor of adenovirus. Mol. Cell. Biol. **6**:4723–4733.
7. **Church, G. M., A. Ephrussi, W. Gilbert, and S. Tonegawa.** 1985. Cell type specific contacts to immunoglobulin enhancers in nuclei. Nature (London) **313**:798–801.
8. **Cordingley, M. G., M. E. M. Campbell, and C. M. Preston.** 1983. Functional analysis of a herpes simplex virus type 1 promoter: identification of far upstream regulatory sequences. Nucleic Acids Res. **11**:2347–2365.
9. **Dignam, J. D., R. M. Lebovitz, and R. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
10. **Ephrussi, A., G. M. Church, S. Tonegawa, and W. Gilbert.** 1985. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. Science **227**:134–140.
11. **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.
12. **Fried, M., and D. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
13. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions. Nucleic Acids Res. **9**:3047–3060.
14. **Hendrickson, W., and R. F. Schlief.** 1984. Regulation of the Escherichia coli L-arabinose operon by gel electrophoresis DNA binding assay. J. Mol. Biol. **179**:611–628.
15. **Hendrickson, W., and R. F. Schlief.** 1985. A dimer of AraC protein contacts three adjacent major groove regions of the AraI DNA site. Proc. Natl. Acad. Sci. USA **82**:3129–3133.
16. **Herz, C., and B. Roizman.** 1983. The α promoter regulator-ovalbumin chimeric gene resident in human cells is regulated like the authentic α4 gene after infection with herpes simplex virus 1 mutants in α4 gene. Cell **33**:145–151.
17. **Honess, R. W., and B. Roizman.** 1974. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. **14**:8–19.
18. **Honess, R. W., and B. Roizman.** 1975. Regulation of herpesvirus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides. Proc. Natl. Acad. Sci. USA **72**:1276–1280.
19. **Jones, K. A., and R. Tjian.** 1985. SP1 binds to promoter sequences and activates herpes simplex virus 'immediate-early' gene transcription in vitro. Nature (London) **317**:179–185.
20. **Kovesdi, I., R. Reichel, and J. Nevins.** 1987. Role of an adenovirus E2 promoter binding factor in E1A mediated coordinate gene control. Proc. Natl. Acad. Sci. USA **84**:2180–2184.
21. **Kovesdi, I., R. Reichel, and J. R. Nevins.** 1986. E1A transcription induction: enhanced binding of a factor to upstream promoter sequences. Science **231**:719–722.
22. **Kristie, T. M., N. Michael, D. Spector, and B. Roizman.** 1987. Analysis of the proteins binding to the cis-acting sites of α-TIF and α4, two major herpes simplex virus 1 trans-acting factors. In Mechanisms of control of gene expression, UCLA Symp. Mol. Cell. Biol. **67**:201–212.
23. **Kristie, T. M., and B. Roizman.** 1984. Separation of sequences defining basal expression from those conferring α gene recognition within the regulatory domains of herpes simplex virus 1 α genes. Proc. Natl. Acad. Sci. USA **81**:4065–4069.
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 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. The binding site of the major regulatory protein α4 specifically associated with the promoter-regulatory domains of α genes of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **83**:4700–4704.
26. **Kristie, T. M., and B. Roizman.** 1987. Host cell proteins bind to the *cis*-acting site required for virion-mediated induction of herpes simplex virus 1 α genes. Proc. Natl. Acad. Sci. USA **84**:71–75.
27. **Laimins, A. L., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA **79**:6453–6457.
28. **Mackem, S., and B. Roizman.** 1980. Regulation of herpesvirus macromolecular synthesis: transcription-initiation sites and domains of α genes. Proc. Natl. Acad. Sci. USA **77**:7122–7126.
29. **Mackem, S., and B. Roizman.** 1982. Regulation of α genes of herpes simplex virus: the α 27 promoter-thymidine kinase chimeric gene is positively regulated in converted L cells. J. Virol. **43**:1015–1023.
30. **Mackem, S., and B. Roizman.** 1982. Structural features of the α gene 4, 0, and 27 promoter-regulatory sequences with confer α regulation on chimeric thymidine kinase gene. J. Virol. **44**:939–949.
31. **Mackem, S., and B. Roizman.** 1982. Differentiation between α promoter and regulator regions of herpes simplex virus. I. The functional domains and sequence of a movable α regulator. Proc. Natl. Acad. Sci. USA **79**:4917–4921.
32. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
33. **McKnight, J. L. C., T. M. Kristie, and B. Roizman.** 1987. The binding of the virion protein mediating α gene induction in herpes simplex virus 1 infected cells to its *cis* site requires cellular proteins. Proc. Natl. Acad. Sci. USA **84**:7061–7065.
34. **McKnight, J. L. C., T. M. Kristie, S. Silver, P. E. Pellett, P. Mavromara-Nazos, G. Campadelli-Fiume, M. Arsenakis, and B. Roizman.** 1986. Regulation of herpes simplex virus 1 gene expression: the effect of genomic environments and its implication for model systems, p. 163–173. *In* Cancer Cell, vol. IV. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science **217**:316–324.
36. **Morse, L. S., L. Pereira, B. Roizman, and P. A. Schaffer.** 1978. Anatomy of HSV DNA. XI. Mapping of viral genes by analysis of polypeptides and functions specified by HSV-1 × HSV-2 recombinants. J. Virol. **26**:389–410.
37. **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.
38. **O'Hare, P., and G. S. Hayward.** 1985. Three *trans*-acting regulatory proteins of herpes simplex virus modulate immediate early-early gene expression in a pathway involving positive and negative feedback regulation. J. Virol. **56**:723–733.
39. **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.
40. **Pellett, P. E., J. L. C. McKnight, F. J. Jenkins, and B. Roizman.** 1985. Nucleotide sequence and predicted amino acid sequence of a protein encoded in a small herpes simplex virus DNA fragment capable of trans-inducing α genes. Proc. Natl. Acad. Sci. USA **82**:5870–5874.
41. **Pizer, L. I., D. C. Tedder, J. L. Betz, K. W. Wilcox, and P.**

Beard. 1986. Regulation of transcription in vitro from herpes simplex virus genes. J. Virol. **60**:950–959.

42. **Post, L. E., S. Mackem, and B. Roizman.** 1981. The 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.
43. **Post, L. E., B. Norrild, T. Simpson, and B. Roizman.** 1982. Chick ovalbumin gene fused to a herpes simplex virus α promoter and linked to a thymidine kinase gene is regulated like a viral gene. Mol. Cell. Biol. **2**:233–240.
44. **Preston, C. M.** 1979. Control of herpes simplex virus 1 mRNA synthesis in cells infected with wildtype virus or the temperature-sensitive mutant tsK. J. Virol. **29**:275–284.
45. **Preston, C. M., M. G. Cordingley, and N. D. Stow.** 1984. Analysis of DNA sequences which regulate transcription of a herpes simplex virus immediate early gene. J. Virol. **50**:708–716.
46. **Pruijn, G. J. M., W. Van Driel, and P. C. Van der Vliet.** 1986. Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. Nature (London) **322**:656–659.
47. **Roizman, B.** 1979. The structure and isomerization of herpes simplex virus genomes. Cell **16**:481–494.
48. **Roizman, B., and W. Batterson.** 1985. The replication of herpesviruses, p. 497–526. *In* B. Fields (ed.), Virology. Raven Press, New York.
49. **Roizman, B., T. M. Kristie, W. Batterson, and S. Mackem.** 1984. The regulation of α genes of herpes simplex virus 1, p. 227–238. *In* H. Vogel and H. S. Ginsberg (ed.), Proceedings of the Physicians and Surgeons Biomedical Sciences Symposium on Transfer and Expression of Eukaryotic Genes. Academic Press, Inc., New York.
50. **Roizman, B., and A. E. Sears.** 1987. An inquiry into the mechanism of herpes simplex virus latency. Annu. Rev. Microbiol. **41**:543–571.
51. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46**:705–716.
52. **Shih, M.-F., M. Arsenakis, P. Tiollais, and B. Roizman.** 1984. Expression of hepatitis B virus S gene by herpes simplex virus 1 vectors carrying α and β regulated gene chimeras. Proc. Natl. Acad. Sci. USA **81**:5867–5870.
53. **Siebenlist, U., and W. Gilbert.** 1980. Contacts between E. coli RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77**:122–126.
54. **Singh, H., R. Sen, D. Baltimore, and P. A. Sharp.** 1986. A nuclear factor that binds a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) **319**:154–158.
55. **Strauss, F., and A. Varshavsky.** 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell **37**:889–901.
56. **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.
57. **Watson, R. J., and G. F. Vande Woude.** 1982. DNA sequence of an immediate early gene (IEmRNA-5) of herpes simplex virus 1. Nucleic Acids Res. **10**:979–991.