JOURNAL OF VIROLOGY, Dec. 1993, p. 7201–7214
0022-538X/93/127201-14$02.00/0

Vol. 67, No. 12

Mutually Exclusive Binding of Two Cellular Factors within a Critical Promoter Region of the Gene for the IE110k Protein of Herpes Simplex Virus

D. O'ROURKE AND P. O'HARE*

Marie Curie Research Institute, The Chart, Oxted, Surrey, England RH8 0TL

Received 28 May 1993/Accepted 17 August 1993

We have examined the *cis*- and *trans*-acting factors involved in constitutive transcription of the promoter for the IE110k protein of herpes simplex virus type 1. Our results indicate that while the IE110k gene is activated by Vmw65, it also exhibits very efficient constitutive expression approximating that from the simian virus 40 early enhancer-promoter region. We show that despite the presence of multiple copies of the octamer consensus site which mediate Oct-1 binding and subsequent Vmw65 activation, these upstream sequences have a minor effect on constitutive transcription. By progressive exonuclease digestion and subsequent site-directed mutagenesis of the promoter, we have identified a 15-bp region (termed the EC region), from position −89 to −74, which is required for efficient constitutive expression from the IE110k promoter. We demonstrate that two cellular proteins interact with this region and, by competition and methylation interference analyses, show they have distinct but overlapping sequence requirements for binding. One of these proteins is identified as NF-Y, a CCAAT box-binding factor, which binds an inverted CCAAT box located between positions −71 and −75. The second cellular factor, F2, appears to be novel and binds a region with the sequence CGCGCGGC CAT which overlaps the 3′ end of the CCAAT box. The terminal AT of the recognition site for F2 represents, on the opposite strand, the terminal AT of the CCAAT box, and these and adjacent bases are critically required for the binding of both factors. These results together with further competition analysis indicate that these factors bind in a mutually exclusive manner to the EC region. The implications of these results for regulation of expression of the IE110k gene are discussed.

There is abundant evidence that transcriptional regulation is the main level of control of the coordinate induction of the differentially regulated immediate-early (IE), delayed-early (DE), and late groups of genes of herpes simplex virus (HSV) (for reviews, see references 14, 27, and 66). It is also clear that expression of the viral gene products is controlled by additional complex processes at other levels (37, 38, 56, 68). Of the five IE proteins encoded by HSV, two, IE175k (ICP4) and IE63k (ICP27), are essential for virus replication (11, 47, 54). The IE175k protein is necessary for the transition to the DE and late phases of virus gene expression and acts by facilitating the initiation of transcription at the corresponding promoters (9, 45, 67), although its precise mechanism of action is unclear. Results to date indicate that IE63k is a multifunctional protein involved in transcriptional and posttranscriptional regulation of virus gene expression (34, 51, 60). Of the three other IE proteins, the IE110k protein (ICP0) has been the most intensively studied. This protein nonspecifically activates gene expression in transient systems, and although viruses with conditional lethal mutations of this gene have not been isolated, variants containing deletions within the IE110k gene exhibit substantial impairment in infectivity, particularly at low multiplicities of infection (3, 6, 13, 18, 43, 55, 61). However, while IE110k is distinct from IE175k and IE63k in that it is not absolutely essential for virus replication (55, 61), it is nonetheless required for normal levels of expression of all classes of virus genes (3, 6, 15). Moreover, unlike the other IE proteins, IE110k may have a specialized role in facilitating the reactivation of HSV from a latent infection. Viruses containing deletions in the gene for the IE110k protein exhibit decreased virulence and are impaired to various degrees in the ability to reactivate from a latent infection (31, 61). Additionally, studies on an in vitro tissue culture latency system have demonstrated that infection with recombinant adenoviruses expressing the IE110k protein could reactive latent HSV, while wild-type adenovirus or adenovirus expressing the IE175k protein could not (53, 70). For a review of the functions of the IE110k protein in HSV replication, see reference 16. Definitive roles for the IE12k (ICP47) and the IE68k (ICP22) proteins have yet to be established, although it has been reported that a mutant virus with a deletion within the IE68k gene was debilitated for growth in a cell-type-dependent manner (57).

The critical requirement for IE proteins in the proper orchestration of virus gene expression and replication is reflected in the operation of a complex mechanism to boost transcription of the IE genes themselves (46). This process is mediated by a structural component of HSV, Vmw65 (VP16), which interacts with at least two host cell factors to induce high levels of IE transcription. However, while a considerable amount of information has been obtained on the identification of the participating proteins and the mechanism involved in Vmw65-mediated induction of IE transcription (for a review, see reference 40), much less is known about the requirements for IE transcription in the absence of the Vmw65 transactivation pathway. It is likely that such information may be relevant to an understanding of IE expression in cells where Vmw65 is unable to function during virus infection and in particular to the elucidation of the transcriptional requirements for IE gene expression during the initial stages of reactivation from a latent infection where Vmw65 is absent. To this end we have examined the *cis*- and *trans*-acting factors involved in constitutive transcription of the promoter for the IE110k protein, since several lines of evidence indicated that this protein may have a

* Corresponding author.

qualitatively distinct role in virus reactivation. We show that the IE110k promoter is very efficient, but the presence of multiple copies of the octamer consensus site, which mediate Oct-1 binding and subsequent Vmw65 activation, have a minor effect on constitutive transcription. By progressive exonuclease digestion and subsequent site-directed mutagenesis, we have identified a region which is essential for the efficient constitutive expression from the IE110k promoter. We demonstrate that two cellular proteins bind with overlapping sequence requirements to this region. One of these proteins is identified as NF-Y, a CCAAT box-binding factor, which binds to an inverted CCAAT box located between positions −71 and −75. The second, presently unidentified, cellular factor binds to a region with the sequence CGCGCGGCCAT which overlaps the 3′ end of the CCAAT box. The terminal AT of the recognition site for the second cellular factor represents, on the opposite strand, the terminal AT of the CCAAT box. These and adjacent bases are critically required for the binding of both factors.

MATERIALS AND METHODS

Plasmid constructs. The plasmid pAB2 was constructed by insertion of a *Sma*I fragment from pPO49, which contains the complete promoter-regulatory region of the IE110k gene from HSV-1 strain KOS, into the *Sma*I site of pUC19 as described previously (41). The IE110k promoter fragment in pAB2 extends from position −130 to +150, lacks upstream TAAT GARAT elements, and is unresponsive to Vmw65 (41).

A series of 5′ deletions of the IE110K promoter was constructed by digestion of pAB2 with the restriction enzyme *Pst*I at a unique site in the polylinker at the 5′ end of the promoter. *Bal* 31 exonuclease was used to degrade both 5′ and 3′ termini of the cleaved plasmid at approximately 20 bp/min/U. At specified times, *Bal* 31 exonuclease digestion was stopped with EGTA (ethylene glycol-bis(β-aminoethylether)-*N*,*N*,*N*′,*N*′-tetraacetic acid), and single-stranded DNA was removed by mung bean nuclease treatment. *Hin*dIII linkers were added, and the vectors were religated and transformed into *Escherichia coli* HB101. A series of constructs named pDR1 to pDR6 was selected on the basis of restriction analysis of plasmid DNAs from individual colonies, and the endpoints of the deletions were subsequently mapped by dideoxy sequencing of purified DNA with the Sequenase version 2 kit (U.S. Biochemical Corp.). In case variability in the vector sequences 5′ of the endpoint of the IE110k promoter might have any influence on the results, the *Hin*dIII-*Bam*HI fragment containing each of the deletion mutants was subcloned into *Hin*dIII-*Bam*HI-cleaved pAB2, making the vector sequence upstream of each construct identical in all cases. Equivalent results were obtained in both the original and vector-controlled settings. Other promoter-chloramphenicol acetyltransferase (CAT) constructs included for comparative purposes in this work were pVP5CAT, which contains the HSV late promoter for the major capsid gene; pTKCAT and p38kCAT, which contain the HSV DE promoters for thymidine kinase and the small subunit of ribonucleotide reductase, respectively; pSV2CAT, which contains the simian virus 40 (SV40) early enhancer-promoter region; and pWT760CAT, which contains the powerful major IE enhancer-promoter region from human cytomegalovirus.

Cells, transfection procedures, and CAT assays. HeLa cells and Vero cells were grown in Dulbecco modified Eagle medium with 10% newborn calf serum. Cells were plated the day before transfection into cluster dishes (6 by 35 mm) at 10^6 cells per well. DNA transfections were carried out by the calcium phosphate precipitation method modified by the use of BES [*N*,*N*-bis(2-hydroxethyl)-2-amino-ethanesulfonic acid]-buffered saline (pH 6.95) in place of HEPES (hydroxyethylpiperzine-*N*′-2-ethanesulfonic acid)-buffered saline (5, 42). Extracts were made 40 to 48 h after transfection, and CAT assays were performed and quantitated as described previously (42). Cells were cotransfected with the indicated amount of CAT reporter vectors, and the total amount of DNA was equalized to 2 μg in all cases by the addition of pUC19 carrier DNA.

Electrophoretic mobility shift assay. Large-scale preparation of HeLa cell nuclear extracts (from approximately 5×10^9 cells) was performed according to the method of Dignam et al. (10). The assay mixtures (routinely in a final volume of 20 μl) to examine DNA-binding activity contained 25 mM HEPES-KOH (pH 7.9), 50 mM NaCl, 10% glycerol, 0.05% Nonidet P-40, 5 mM dithiothreitol, 1 mM EDTA, 5.0 mg of bovine serum albumin per ml, and 1 to 2 μg of poly(dI · dC)-poly(dI · dC). After 10 min of incubation at room temperature, end-labelled oligonucleotide probes (1 ng) were added, and incubation was continued for a further 20 min. Samples were then loaded onto 4 or 8% nondenaturing polyacrylamide gels (acrylamide-bisacrylamide, 19:1). Electrophoresis was performed for 2.5 to 3 h at 200 V in a buffer containing 0.5× Tris-borate-EDTA (17). Complexes were visualized by autoradiography of the dried gels. For quantitative measurements, the areas of interest were cut out of dried gels and Cerenkov radiation of ^{32}P was measured with an efficiency of approximately 30%. For competition analysis, unlabelled probes were added at the indicated amounts and incubated with the extract for 10 to 15 min prior to the addition of the radiolabelled probe. In all cases, unless otherwise stated, the probe was present in excess amounts but was frequently run off the bottom of the gel to facilitate separation of the binding complexes. Oligonucleotides were synthesized by using an Applied Biosystems DNA synthesizer. Complementary strands were annealed and radiolabelled by using the Klenow fragment of DNA polymerase I to fill in designed 5′ overhangs. Double-stranded probes were purified by electrophoresis in polyacrylamide gels, elution, and precipitation or by gel filtration using Nuc-Trap columns (Stratagene).

Methylation interference. Oligonucleotides used for methylation interference (59) were purified by UV shadowing and radiolabelled with [γ-^{32}P]ATP and polynucleotide kinase. Radiolabelled single-stranded oligonucleotides were then annealed to give the appropriately end-labelled double-stranded probes, which were then methylated with dimethyl sulfate. A larger, 265-bp *Hin*dIII-*Sal*I fragment isolated from pDR2 (see Results) was also used for methylation interference studies. The noncoding strand was labelled by cutting pDR2 with *Hin*dIII, filling in the recessed 3′ end with the Klenow fragment of DNA polymerase I and [^{32}P]dATP, and then digesting with *Sal*I. The coding strand was specifically labelled by digestion with *Hin*dIII and treatment with calf intestinal alkaline phosphatase and then polynucleotide kinase to phosphorylate the 5′ end, and the appropriately end-labelled fragment was then produced by *Sal*I digestion. The end-labelled 262-bp fragment probes were isolated from a 12% polyacrylamide gel. The probes (100 ng) were mixed with 500 ng of sonicated salmon sperm DNA and incubated with 1 μl of dimethyl sulfate for 1 min in 200 μl of a buffer containing 50 mM sodium cacodylate (pH 8.0) and 0.1 mM EDTA. Methylation was terminated by the addition of 50 μl of stop buffer (1.5 M sodium acetate, 1.0 M Tris-acetate [pH 7.5], 1.0 M β-mercaptoethanol, 1 mM EDTA) and 750 μl of cold ethanol and incubation on ice for 2 min. The probe was then pelleted,

resuspended, reprecipitated, washed with 70% ethanol, and dried.

Approximately 1.5×10^6 cpm of methylated probes was used in each reaction mixture consisting of 10 μl of HeLa cell nuclear extract and 10 μg of poly(dI · dC)-poly(dI · dC)in 25 mM HEPES-KOH (pH 7.9)–50 mM NaCl–10% glycerol–0.05% Nonidet P-40–5 mM dithiothreitol–1 mM EDTA–5.0 mg of bovine serum albumin per ml in a final volume of 50 μl. The amount of competitor was increased to 75 μg and the total volume was increased to 200 μl when the larger fragment probe was used. Following incubation, bound and free probes were separated by electrophoresis in a 4% polyacrylamide gel, visualized by exposure of the wet gel to X-ray film for 2 to 4 h, excised, and incubated overnight at 37°C in 500 μl of extraction buffer consisting of 10 mM magnesium acetate, 1 mM EDTA, 500 mM ammonium acetate, and 0.1% sodium dodecyl sulfate. Probe was recovered by precipitation, washed in 70% ethanol, and dried before cleavage in 100 μl of freshly diluted 10% piperidine at 90°C for 30 min. The reaction was quenched by incubation in a dry ice-ethanol bath. Piperidine was subsequently removed by drying and repeated washes with 100 μl of distilled water. Following the final drying, the pellet was resuspended in 10 μl of formamide loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) and applied to an 8% denaturing urea-polyacrylamide gel. Maxam-Gilbert sequencing reactions (33) were used to prepare G and G+A tracks for electrophoresis in parallel.

RESULTS

Efficient constitutive promoter activity of the gene for the IE110k protein. Although several previous reports on IE regulation by Vmw65 have examined the basal activity of the IE110k promoter, there has been no systematic analysis of its relative efficiency or any identification of the relevant cellular proteins. To examine this question we constructed a chimeric CAT target gene (pPO49) which contained the promoter-regulatory region of the IE110k gene spanning from approximately position −800 to +150 relative to the transcriptional site (see Fig. 2b). Expression from this construct was compared with that from a series of other CAT target genes which contained different test promoters. These included two typical HSV DE promoters, those for the small subunit of ribonucleotide reductase (UL40) and thymidine kinase (UL23) (p38kCAT and pTKCAT respectively), the DE-late promoter for major capsid protein (UL19) (pVP5CAT), the enhancer-promoter region from SV40 (pSV2CAT), and the cytomegalovirus IE enhancer-promoter (pWT760CAT). The various constructs were transfected in dose-response experiments into HeLa or Vero cells, two cell lines routinely used for HSV replication. Typical results from transfections into HeLa cells are shown in Fig. 1. Significant CAT activity was obtained with transfection of only 50 ng of the IE110kCAT construct (pPO49) (lane 1), pWT760CAT (lane 5), and pSV2CAT (lane 6), while activity from p38kCAT, pTKCAT, and pVP5CAT (lanes 2 to 4) was undetectable. Quantitative analysis of expression over the course of this work demonstrated that constitutive activity of the IE110k promoter is only twofold lower than that of the SV40 enhancer-promoter and approximately 10-fold lower than that of the very powerful cytomegalovirus IE enhancer-promoter region. The difference between the activity of the IE110k promoter and that of the HSV DE and DE-late promoters is highlighted by the comparison at input doses of 250 ng (Fig. 1, right). In this case activity from the IE110k promoter was underestimated because of substrate

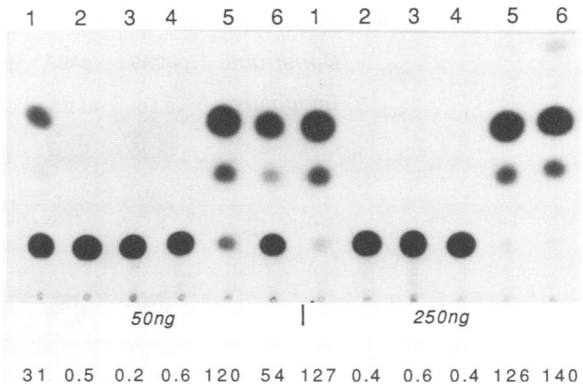

FIG. 1. Comparison of IE110k basal promoter activity with that of various viral promoters. HeLa cells were transfected with 50 or 250 ng of each of the promoter-CAT constructs, harvested 40 h after transfection, and assayed for CAT activity. Lanes: 1, pPO49; 2, pVP5CAT; 3, p38kCAT; 4, pTKCAT; 5, pWT760CAT; 6, pSV2CAT. The input of radiolabelled substrate in each assay was approximately 140×10^3 cpm. CAT activities (absolute counts appearing in the acetylated products) are indicated below the autoradiogram; note that while the assays for pPO49 and pSV2CAT are not substrate limited at the 50-ng dose, they are at the 250-ng dose. CAT activity for pWT760CAT is substrate limited even at the 50-ng dose.

limitation in the assay, while activity from pVP5CAT, p38kCAT, and pTKCAT was not significantly above background. (Note that expression from these same constructs can be efficiently activated in the presence of IE proteins.) The results of several independent experiments indicate that constitutive activity from the IE110k promoter was 20- to 40-fold higher than that from the DE and DE-late promoters. Similar results were obtained with Vero cells (data not shown). Thus, even though transcription from the IE110k promoter-regulatory region responds strongly to activation by Vmw65 (32, 41), constitutive activity of the promoter is comparatively very efficient, on par with that of the early SV40 enhancer-promoter region.

The upstream Vmw65 response elements are not required for efficient constitutive activity. The sequence of the upstream region of the IE110k gene has been determined (32, 44) and contains at least six elements which conform well to the core consensus binding site for the cellular transcription factor Oct-1 (Fig. 2b, solid boxes). The motif nearest the transcription initiation site is located at position −165. Several less-well-conserved motifs are also present on the opposite strand (Fig. 2b, stippled boxes). Although not all of these regions may represent Vmw65 response elements, several of the motifs individually support Oct-1 binding and contain overlapping signals necessary for induction by Vmw65 (1, 32, 41). To determine whether these sites or any other elements present in the upstream regulatory region were required for the efficient constitutive activity of the promoter, a series of deletions was constructed by progressive *Bal* 31 exonuclease digestion of sequences from the 5′ end. Internal deletion of sequences from the upstream region was also performed by restriction enzyme digestion and religation. Endpoints were determined by sequence analysis, and appropriate constructs spanning the promoter-regulatory region were compared for expression efficiency. Typical results are illustrated in Fig. 2a. Strikingly, deletion of an approximately 700-bp region spanning from position −810 to −125 and including all of the consensus and putative Oct-1-binding sites had at most a marginal effect on promoter activity, with the deletion construct pAB2 exhibiting

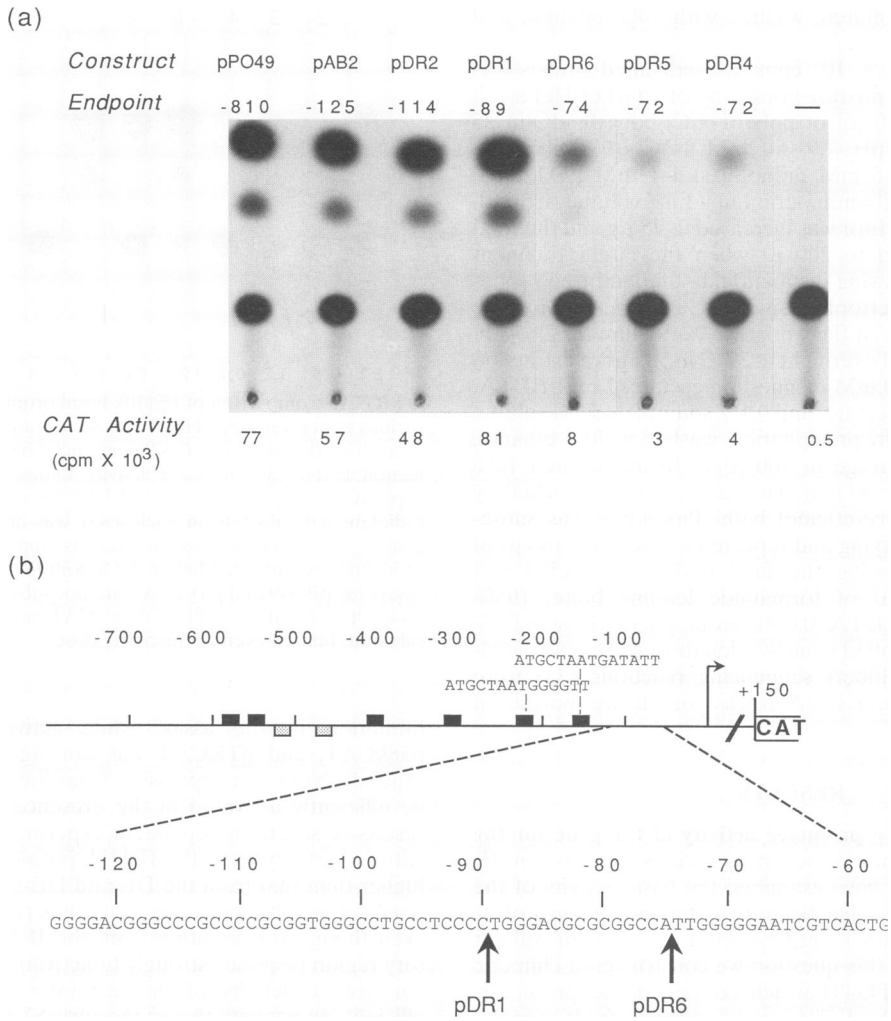

FIG. 2. (a) 5′ deletions identify an IE110k promoter region required for high-level basal promoter activity. A typical CAT assay showing the results of transfection of a series of 5′ deletions of pPO49 in HeLa cells is shown; 200 ng of pPO49 and each of the deletion variants were transfected into HeLa cells, which were harvested 40 h after transfection and assayed for CAT activity. Absolute counts appearing in the acetylated products are indicated below the autoradiogram. (b) Schematic diagram of the IE110k promoter. Upstream consensus and near-consensus octamer sites are indicated by solid and stippled rectangles, respectively. Sequence data are from references 32 and 44 and this work. The start site of transcription is indicated by an arrow. Endpoints of the 5′ deletions which delineate the IE110k basal promoter element are shown on the primary sequence.

a less-than-twofold reduction in activity compared with that of the parental plasmid pPO49. By comparison we have previously shown that pAB2 has completely lost the ability to respond to Vmw65 (41). Further deletion from −125 to −114 (pDR2) or −89 (pDR1), including a consensus SP1-binding site located around position −113, had little significant effect on activity (Fig. 2a). Similarly, internal deletion from position −400 to −125 had no detectable effect (data not shown). There was on average less than twofold variation, to which it is difficult to attribute significance, between any of the constructs to endpoint −89, and in fact pDR1 exhibited activity very similar to that of the parent construct. However, deletion of the 15 bp from position −89 to −74 resulted in a dramatic reduction in activity, with the deletion construct pDR6 exhibiting approximately 8% of the activity of pDR1. The deletion constructs pDR4 and pDR5, independent isolates with endpoints at position −72, exhibited similarly weak activity. The unexpected conclusion from this series of experiments is that neither the multiple Oct-1-binding motifs nor any other element in the upstream regulatory region makes a major contribution to the strong constitutive activity of the IE110k promoter. While other elements may affect activity to a minor degree, our results indicate that a major determinant of IE110k promoter efficiency resides within or overlaps the 15-bp region between −89 and −74, which has been termed the EC (efficient constitutive activity) region.

Two cellular proteins bind to the EC region. The most likely explanation for the involvement of a specific upstream region in determining promoter efficiency is that the region binds one or more transcription factors which directly or indirectly influence transcriptional initiation. To examine the interaction of cellular proteins with the EC region, a series of overlapping probes which spanned the region from approximately position −100 to −60 was constructed and analyzed for binding to proteins from nuclear extracts of HeLa cells by a standard electrophoretic mobility shift assay. The locations of the

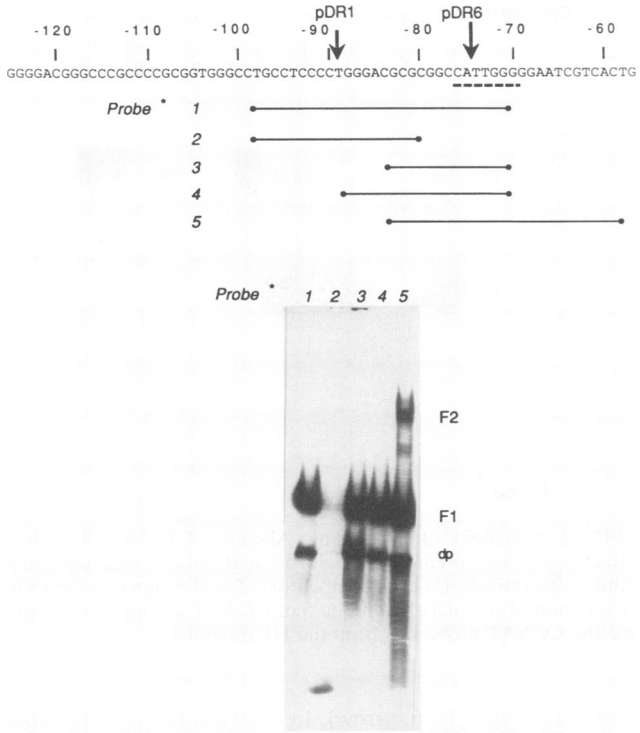

FIG. 3. Demonstration of two complexes binding to the EC region. A series of overlapping oligonucleotides which spanned the EC region is illustrated with the positions of the endpoints of each of the probes and the promoter deletions that delineate the EC region. The position of an inverted CCAAT box is indicated by a dashed line. The lower panel shows the results of binding of equal aliquots (2 μg) of a HeLa cell nuclear extract to each of the probes in the presence of nonspecific competitor poly(dI · dC)-poly(dI · dC) as indicated in Materials and Methods. In this and certain other assays, free probe was electrophoresed off the end of the gel to aid in separation but was normally present in excess (see Fig. 6 and 11). Nuclear proteins binding to the oligonucleotide probes were designated factor 1 (F1) and factor 2 (F2). A DNA-binding degradation product of F1 is also indicated (dp).

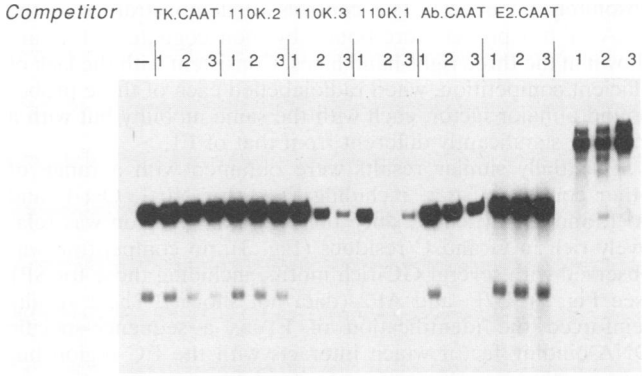

FIG. 4. Specificity of F1 binding. Oligonucleotide probe 1 was radiolabelled and used in competition analysis to determine the specificity of F1 binding. Cold competitors were used at doses of 10-, 100-, and 200-fold excesses over radiolabelled probe and were preincubated with HeLa cell nuclear extract, binding buffer, and poly(dI · dC) prior to the addition of radiolabelled probe 1. The cold competitors used included the HSV type 1 thymidine kinase CAAT box (TK.CAAT), the rat albumin CAAT box (Ab.CAAT), and the adenovirus E2 CAAT box (E2.CAAT). The right-hand three lanes show the controls in which each of the competitors was used as a radiolabelled probe to demonstrate their binding activities. Each of the probes bound an identical doublet consistent with CAAT box-binding activity. Lanes: 1, TK.CAAT; 2, Ab.CAAT; 3, E2.CAAT.

probes, together with the endpoints of pDR1 and pDR6, are summarized in Fig. 3. Typical results for each of the probes are illustrated in the lower panel (Fig. 3). Probes 1, 3, and 4, which had the same 3′ endpoint at −70 but different 5′ boundaries, all bound a single factor termed F1. A second product (dp) was observed binding to the same probes as those bound by F1. This product behaved identically to F1 in all binding and competition analyses (see, e.g., Fig. 4), varied in abundance relative to that of F1 in different preparations, and was completely absent in certain preparations made in the presence of high levels of multiple protease inhibitors. We therefore consider it likely that dp represents a degradation product of F1 which retains DNA-binding activity. Probes 1 and 4, which extended further 5′ than probe 3, bound F1 only marginally better (Fig. 3; see Fig. 4), and probe 2, which contains the 5′ section of probe 1 and only partly overlaps probe 3, bound neither F1 nor any other component. (Note that minor amounts of F1 seen in Fig. 3 with probe 2 are due to leakage from lanes 1 and 3). These results indicate that the main determinants for F1 binding are largely contained within the 14-bp sequence of probe 3 from position −84 to −70. However, probe 5, which contains the same 5′ boundary as probe 3 but extends further in the 3′ direction, bound a factor, F2, in addition to F1. Both F1 and F2 specifically bound to probe 5 (see below), and no other complexes were observed under a variety of assay conditions in which salt concentration, pH, and competitor type were systematically altered (data not shown). In addition, F1 and F2 were the only complexes observed on a larger, 45-bp probe which spanned the entire region from −100 to −54 (results not shown). The results indicate that two cellular proteins bound to sequences within or overlapping the EC region defined by the deletion between pDR1 and pDR6: F1, for which sequences in probe 3 suffice, and F2, which requires additional sequences extending further 3′.

To examine further the specificity of binding of the F1 factor and to attempt to identify candidate components, we performed a series of experiments in which we examined the abilities of a variety of known consensus motifs to compete for F1 binding. Figure 4 shows typical results from such assays, which in this case examined a series of CCAAT box motifs, including those from the HSV thymidine kinase gene, the adenovirus E2 late gene, and the rat albumin promoter. In this experiment a probe (probe 1) which bound only F1 was used as the target probe, but similar results for the specificity of F1 were obtained whether or not the probe bound F2 (but see below). Each of the competitors was preincubated with the nuclear extract at 10-, 100-, and 200-fold excesses over the radiolabelled probe (Fig. 4, lanes 1, 2 and 3, respectively, for each competitor). As expected, probe 1 (110k.1, the homologous probe) and probe 3 (110k.3) competed for F1 binding, with probe 1 exhibiting slightly stronger competition than probe 3. Probe 2 (110k.2), corresponding to the 5′ two-thirds of probe 1, as predicted from the binding profile (Fig. 3) failed to compete at any dose tested. With the various CCAAT box probes, although some reduction in F1 binding was observed at the higher doses, this effect was minor compared with that of the homologous competitor (see also Fig. 5, in which it is shown that the homologous competitor virtually eliminates F1 binding at doses at which the CCAAT box competitor has

absolutely no effect). The controls to demonstrate that the CCAAT box probes were bound by their cognate factors are shown in the three right-hand lanes. Consistent with the lack of efficient competition, when radiolabelled each of these probes bound a major factor, each with the same mobility but with a mobility significantly different from that of F1.

Essentially similar results were obtained with a range of other consensus sites, including those for NF-1, Oct-1, and AP1, and in particular, since the F1-binding region was relatively rich in G and C residues (Fig. 3), no competition was observed with several GC-rich motifs, including those for SP1 (see Fig. 5), E2F, and AP2 (data not shown). These results reinforced the identification of F1 as a sequence-specific DNA-binding factor which interacts with the EC region but gave no obvious indication as to its identity.

Identification of F2. We noted that probe 5, which bound the additional factor F2, contained an inverted CCAAT box (Fig. 3, underlined). The core CCAAT sequence would have been placed in close proximity to the ends of probes 1, 3, and 4, and it was possible that F2 binding to probe 5 represented a CCAAT box-binding factor which failed to bind to the other probes because of a specific or nonspecific requirement for sequences flanking the 5′ side of the core CCAAT box (as read on the opposite strand in the IE110k promoter sequence). Moreover, in the controls for the competition analysis of F1 binding shown in Fig. 4, the CCAAT box probes bound complexes with mobilities similar to that of F2. To examine this question directly, we compared F2 with the CCAAT box factors both by comparative mobilities when used as radiolabelled probes in parallel and by cross competition analysis. In this experiment the rat albumin CCAAT box motif was used, although similar results were obtained with each of the other CCAAT motifs which bound complexes of identical mobility (Fig. 4 and data not shown). The results show that F2 comigrates with the CCAAT box-binding factor, although F2 clearly did not bind as efficiently to probe 5 as the CCAAT factor did to the rat albumin CCAAT motif. Furthermore, although not as efficient as the homologous competitor, probe 5 competed for binding of the CCAAT box factor (Fig. 5, right panel), and in the reciprocal experiment the CCAAT box motif competed efficiently for F2 (Fig. 5, left panel). This latter experiment also demonstrated specificity in competition, since the CCAAT box motif competed for F2 but not F1 binding to probe 5, while the homologous competitor competed efficiently for both. As expected, the control SP1 motif competed for neither F1 nor F2.

Although there has been some ambiguity in the identification of CCAAT box-binding factors, it is now clear that a number of the factors identified in different studies are the same and are identical to NF-Y, a transcription factor originally identified from analysis of major histocompatibility complex gene expression (see Discussion). To provide further evidence for the identity of F2, we examined the effect of antibody to NF-Y, which has been previously shown also to bind to the CCAAT box motifs employed in this work. In this type of experiment the identity of a DNA-binding complex is inferred from the ability of an antibody to selectively inhibit its formation or to further retard its mobility in an electrophoretic mobility shift assay. Probe 5 was incubated with equal aliquots of nuclear extract in the presence of preimmune or immune serum to NF-Y, and complexes were separated by gel retardation assay (Fig. 6). The control preimmune serum had little significant effect on any of the probe 5-binding complexes (Fig. 6, left lanes). However, after incubation with the immune serum, the F2 complex disappeared and there was the concomitant appearance of a large complex which barely migrated from the origin (Fig. 6, arrow). In contrast, in the same lanes no effect on F1 binding was observed. These results indicate that the F2 complex is selectively retarded by antibody to NF-Y, and together with the nature of the primary sequence itself, the similar mobilities of F2 and the CCAAT box-binding factors, and the results of reciprocal cross competition analysis, the data present a strong case that the F2 complex is the CCAAT box-binding factor NF-Y.

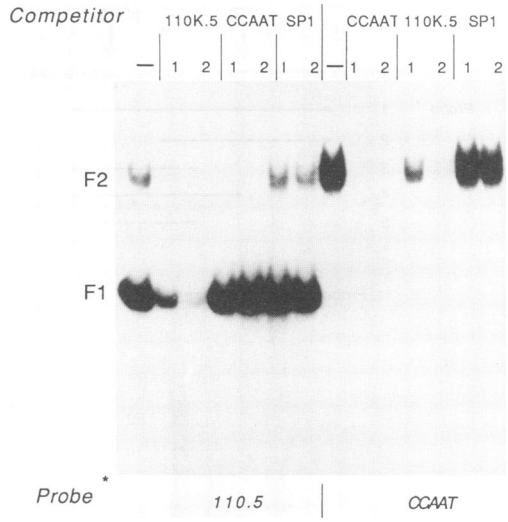

FIG. 5. Specificity of F2 binding. Aliquots of HeLa cell nuclear extract (2 μg) were preincubated with 50- and 250-fold excesses (lanes 1 and 2, respectively, for each competitor) of cold competitors prior to the addition of radiolabelled probe 5 or CCAAT as indicated. In this case the CCAAT probe was from the H1 promoter.

Separation of requirements for binding of F1 and F2. The

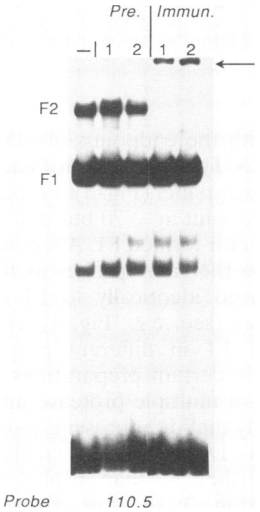

FIG. 6. Identification of F2 as an NF-Y-type CAAT box-binding protein. Serum I320 raised against the N terminus of NFY-B, a component of the heterodimeric NF-Y factor which binds to the CCAAT box of major histocompatibility complex class II genes, was preincubated with 2 μg of HeLa cell nuclear extract for 25 min; this was followed by addition of radiolabelled IE110k promoter probe 5. A control preimmune serum (Pre.) was tested in parallel. The two lanes represent 2- and 4-μl amounts for each serum. The shift in F2 binding in the presence of the immune serum is indicated by an arrow.

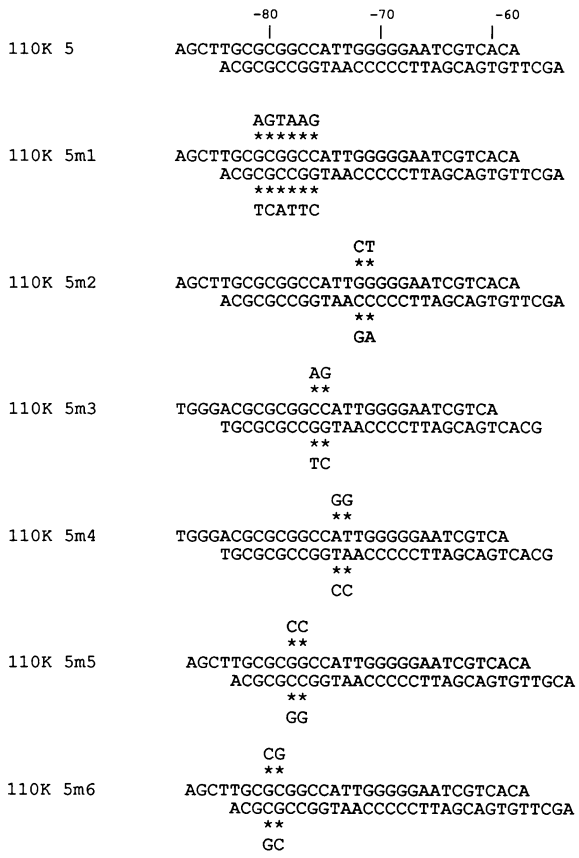

FIG. 7. Summary of wild-type and mutant probe 5 sequences. Substituted bases are indicated by asterisks, with the new bases in the mutant variants given above and below each strand.

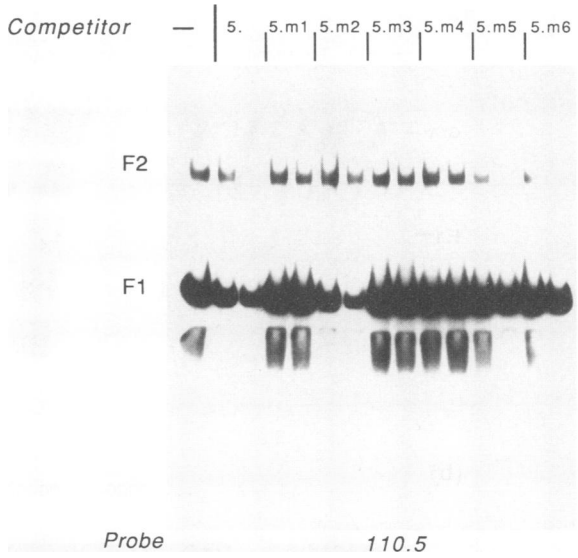

FIG. 8. Competition analysis identifies selective and shared requirements for F1 and F2 binding. Aliquots (2 μg) of HeLa cell nuclear extract were preincubated for 10 min at 20°C with 10- or 50-fold excesses (the two lanes, respectively, for each competitor) of unlabelled wild-type or variant competitors in the presence of poly(dI · dC) and binding buffer. Radiolabelled probe 5 was added, and incubation proceeded for 20 min at 20°C. Samples were then run on a 4% nondenaturing polyacrylamide gel. F1 and F2 binding in the absence of competitor is also indicated (—).

results of the binding profiles of probes 1 and 3 versus that of probe 5 indicated that F1 bound to the 5′ GC-rich segment and that F2 bound to the inverted CCAAT box with some requirement for adjacent sequences (Fig. 3). In an attempt to separate binding requirements within probe 5 for F1 and F2, a series of mutant probes which contained 2-base substitutions across the likely F1-binding and CCAAT box regions (Fig. 7, 110K 5m2 to 5m6) was constructed. By using radiolabelled parent probe 5, each of the mutant probes was examined in gel retardation assays for the ability to compete for F1 and F2 (Fig. 8; the two lanes for each competitor represent 10- and 50-fold excesses over the radiolabelled probe). The results show that several of the mutations affected the binding of both F1 and F2. Thus, probe 5m4, which contained substitutions at the terminal AT of the CCAAT box, had reduced ability to compete for F2 as predicted but also showed reduced competition for F1 compared with the wild-type probe 5. Similarly, substitution of the two C bases immediately 5′ but outside the core CCAAT box (5m3) caused reduced competition not only for F1 but also for F2. Consistent with this latter result, the probe containing multiple mutations within the GC-rich 5′ section, but retaining the CCAAT box (5m1), failed to compete for both F1 and F2. On the other hand, certain of the mutations clearly had a selective effect on the binding of either F1 or F2. Substitution of the two 5′ C bases of the inverted CCAAT box (5m2) affected the ability to compete for F2 without having any significant effect on competition for F1. Conversely, substitution of the GC or GG bases within the GC-rich section (5m6 and 5m5, respectively) selectively reduced competition for F1 without having an effect on competition for F2. These results indicated that F1 and F2 had selective binding requirements within the GC-rich region around −80 and at the 5′ end of the CCAAT box (read on the opposite strand) around position −71 but that they shared binding requirements within the central region between −74 and −77. In general, the gel retardation competition analysis gave a consistent picture of the binding requirements for F1 and F2. However, it was difficult to obtain a refined delineation of the respective interactions by this method, since while certain of the mutant probes were clearly deficient in binding one or the other factor, they would nonetheless compete to some degree at higher doses. To examine further the recognition sites and identify specific bases which were exclusively required for one or the other factor, and in particular to determine whether any individual bases represented important contact points for both factors, we performed methylation interference analysis of each of the isolated complexes. In this method a radiolabelled probe is partially methylated by dimethyl sulfate at the N-7 position of guanine residues (in the major groove) and, much less efficiently, at the N-3 position of adenine residues (in the minor groove). After protein binding, free probe is separated from bound complexes by nondenaturing gel retardation analysis, the DNA is then isolated and cleaved with piperidine, and the products are separated by electrophoresis in denaturing sequencing gels. If methylation at a particular purine interferes with binding of a factor, then the product of cleavage at that purine will be missing or underrepresented in the bound fraction corresponding to that factor.

We first analyzed methylation interference of F1 binding by using a 265-bp probe produced by restriction digestion and

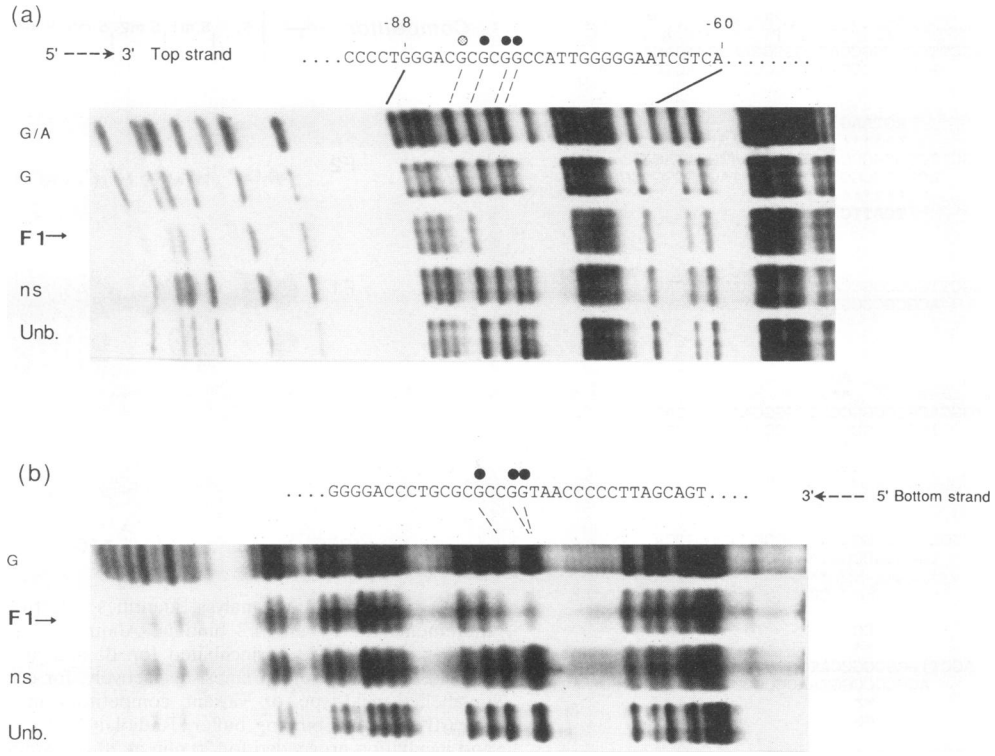

FIG. 9. Methylation interference data for F1 binding on the IE110k promoter element. A *Hin*dIII-*Sal*I fragment from pDR2 (positions −114 to +151) was radiolabelled at the 5′ end of the top strand (a) and at the 3′ end of the bottom strand (b) for methylation interference analysis. The radiolabelled fragment was partially methylated with dimethyl sulfate and used in a large-scale binding reaction. Bound and free probes were isolated from a 4% nondenaturing polyacrylamide gel, recovered, and cleaved with piperidine. Equalized counts of bound probe (F1), unbound probe (Unb), a nonspecific control (ns), and a Maxam-Gilbert sequencing reaction (G and G/A) were loaded onto an 8% sequencing gel. The sequence of this region of the promoter is also shown, along with the methylation interference data for F1 on the upper and lower strands. Methylation interference data for F1 are indicated as circles, with the extent of shading reflecting the extent to which methylation of bases inhibits binding. Solid circles represent bases with the strongest inhibitory effect on binding.

spanning from position −114 to +151 of the IE110k promoter. The methylated fragment was radiolabelled on either strand and incubated with the protein extract, and the F1 complex was identified by specific competition analysis. Unbound probe and probe from the F1 complex was isolated and electrophoresed in parallel with G and G+A sequencing lanes (Fig. 9, Unb, F1, G, and G+A, respectively). A complex which nonspecifically bound to the probe was also analyzed as an additional control (Fig. 9, ns). The results are illustrated together with the sequence of the appropriate region in the 5′-to-3′ direction for each strand (Fig. 9). For F1, strong contacts were observed for the G residues at positions −78, −79, and −81 and a weaker contact was observed at position −83 (Fig. 9a, solid and shaded circles respectively). No contacts were observed at G residues further 5′, i.e., from position −86 onward, or from position −72 onward on the 3′ side (the next 3′ G after the protected base at −78). In comparison, no contacts were observed for the nonspecific control complex (Fig. 9, cf. F1, Unb, and ns). These results reinforce those from the competition analysis (Fig. 8), in which substitution of residues −78 and −79 (5m5) or −80 and −81 (5m6) affected F1 binding while substitutions at positions −72 and −73 (5m2) did not. Similarly, analysis of the opposite strand indicated strong contacts for F1 at the G residues −76, −77, and −80 (Fig. 9b, cf. F1 and Unb), consistent with the competition analysis, in which probe 5m3, substituted at positions −76 and −77, and probe 5m6, substituted at positions −80 and −81, failed to compete for F1 (Fig. 8).

To identify selective requirements for F1 and F2 binding and to address the possibility of the F1-binding site overlapping the adjacent F2-binding site at the CCAAT box, we compared methylation interference patterns for each of these factors on the top strand of a shorter oligonucleotide probe spanning sequences from position −100 to −54 (Fig. 10). As expected, for F1 the pattern of contacts around the G residues at −78 to −83 (Fig. 10a) was similar to that obtained on the longer probe. By comparison, none of these residues exhibited any detectable contacts with F2. The difference in F1 and F2 contacts within this region was particularly striking for bases −78, −79 (migrating as a doublet in this gel), and −81 (Fig. 10, cf. F1 and F2). Conversely, while no contacts were observed for the F1 factor within the string of G residues from −72 onwards, contacts at positions −72 and −71 were observed for F2 (Fig. 10, cf. F1, circles, and F2, solid triangles). Weak contacts for F2 were also indicated in several experiments at purines further 3′ (Fig. 10b, stippled triangles). Importantly, methylation of A residues and their subsequent detection after cleavage were much more efficient with this shorter probe. For example, the A residues at positions −66, −67, and −60 flanking the 3′ side of the string of Gs were easily detected with the unbound oligonucleotide probe (Fig. 10a) but were absent in the profile of the unbound fragment probe (Fig. 9a; cf. the G+A and G sequencing lanes with the unbound methylated probe). Critically, this feature rendered the A residues informative for contact analysis and allowed the identification of the A residue at position −75, which corresponds to the 3′-

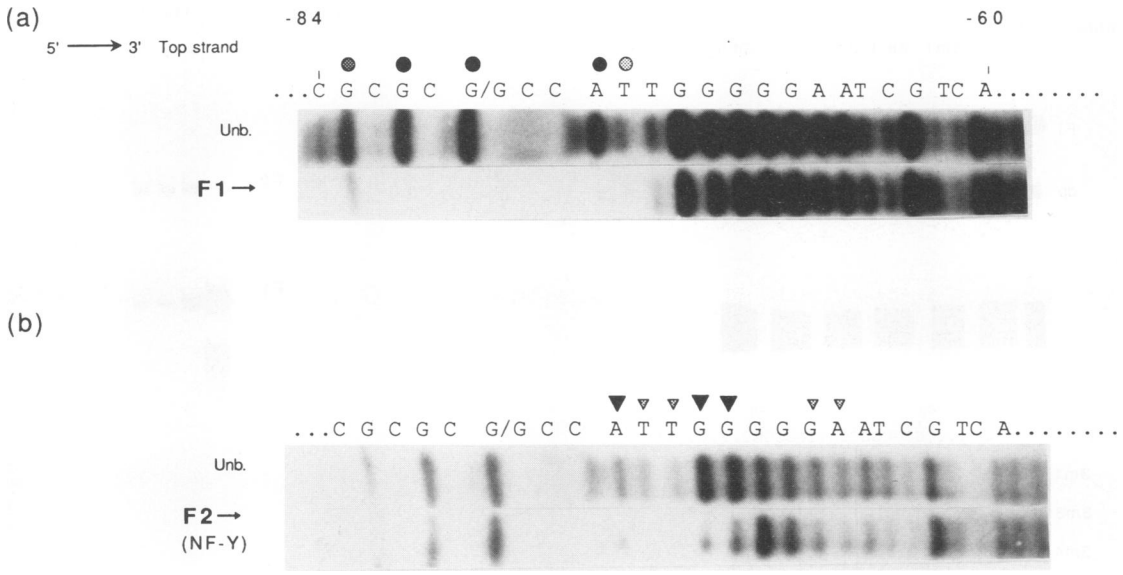

FIG. 10. Comparative methylation interference data for F1 and F2 on the IE110k promoter element. The top strand of a 45-bp oligonucleotide corresponding to positions −100 to −54 of the IE110k promoter was radiolabelled at the 5′ end and partially methylated with dimethyl sulfate. After incubation with HeLa cell nuclear extract, F1- and F2-bound complexes were separated from free probe by electrophoresis in an 8% nondenaturing polyacrylamide gel, and the DNA was recovered and cleaved with piperidine. F1 (a)- and F2 (b)-bound DNA were then analyzed on an 8% sequencing gel in parallel with the unbound probe (Unb), equalized with respect to counts per minute for either F1 or F2. The sequence of the F1/F2 region is shown above the methylation interference gel. Data for F1 and F2 are indicated by circles and triangles, respectively, with the extent of the shading reflecting the extent to which methylation of these bases inhibits binding.

terminal T residue of the inverted CCAAT box, as a critical contact point for the F1 factor. Methylation at this residue strongly interfered with F1 binding, while, as indicated above, methylation of the G residues 2 bases to the 3′ side had no effect (Fig. 10a). In the parallel analysis methylation at this same residue also interfered with F2 binding (Fig. 10b). These results therefore strongly support and provide an explanation for the results of the competition analysis, in which mutation of A at position −75 and T at −74 (probe 5m4) inhibited binding of both F1 and F2 (Fig. 8). In addition, we performed further gel retardation analysis to examine the effect of single mutations at the A at −75 or the T at −74 in the context of 110k probe 3, which bound only F1 (Fig. 3). The results are shown in Fig. 11; the homologous wild-type probe 3 or a series of mutant versions were used as competitors at 10- and 100-fold excesses over the radiolabelled probe. While the wild-type probe competed effectively for F1, the single substitution at the A at −75 (3m3) virtually eliminated binding, with little competition being observed even at the highest dose of competitor. Substitution of the adjacent T residue at position −74 (3m4) had an equally strong deleterious effect on F1 binding, with each of these mutant probes being as defective as the mutant probe containing a double substitution at the C residues −76 and −77 (3m5). In contrast, substitution (3m6) of the T residue at position −73 together with the G at position −72 (which from the methylation interference analysis did not make any detectable contact with F1 [Fig. 9 and 10]) had little effect on F1 binding.

A summary of the results of F1 and F2 binding, together with the positions of the deletion endpoints which defined the EC region, is illustrated in Fig. 12. Together, the results from competition analysis (indicated by hatched bars) and from methylation interference analysis (indicated by circles and triangles for F1 and F2, respectively) indicate the presence of overlapping binding sites for the two factors. While each has distinct and selective requirements, within the 5′ GC-rich region for F1 and within the 3′ region corresponding to the CCA of the inverted CCAAT box for F2, both factors require the central region, including the T at −74, the A at −75, and most likely the C at −76 and −77.

The proximity of the binding sites for the two factors and in particular the demonstration of shared requirements and contact points strongly indicate that F1 and F2 may bind in a mutually exclusive manner (although there are several precedents for the simultaneous binding of proteins which share contact points [see Discussion]). To test this proposal we performed a competition experiment using radiolabelled probe 5, which bound F1 and F2, and unlabelled probe 3, which bound only F1. If F1 and F2 had bound to probe 5 in a mutually exclusive manner (with F1 generally being detected in greater amounts because of either its greater abundance or its greater binding affinity), then competition with a probe which could bind only F1 should lead to increased binding of F2 on the radiolabelled probe. Alternatively, if F1 binding had no effect on (or even stabilized) F2 binding, then reducing F1 by competition should have no effect on (or even reduce) F2 binding to the radiolabelled probe. In this experiment we employed amounts of the radiolabelled probe which were subsaturating as determined from dose-response experiments. The results show that with increasing amounts of competitor probe 3, F1 binding is decreased but F2 binding is concomitantly increased (Fig. 13). In control experiments addition of competitor probe 3 had no effect on F2 binding to a radiolabelled probe which contained only the CCAAT box site. In addition, the amount of F2 binding to probe 5 when F1 was reduced by competition approached that observed with high-affinity CCAAT box sites such as that in the rat albumin promoter (data not shown). Thus, with respect to a probe

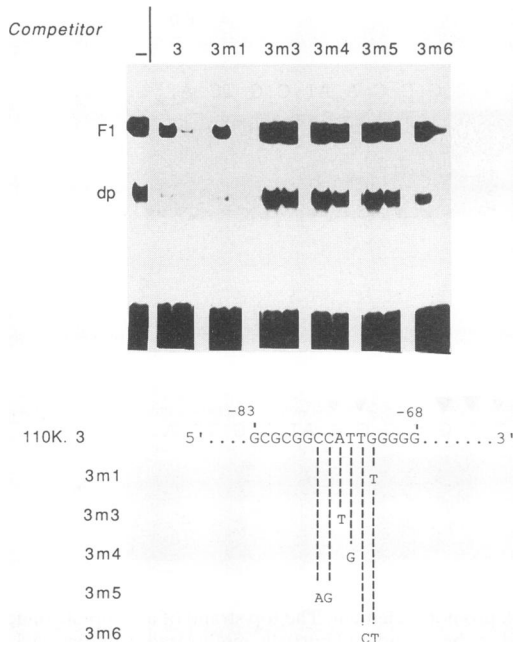

FIG. 11. F1 binding requires bases within the conserved CCAAT box. Aliquots (2 μg) of HeLa cell nuclear extract were preincubated for 10 min at 20°C with a 10- or 50-fold excess (the two lanes, respectively, for each competitor) of unlabelled wild-type or variant competitors in the presence of poly(dI·dC) and binding buffer. Radiolabelled probe 3 was added, and incubation proceeded for 20 min at 20°C. Samples were then run on a 4% nondenaturing polyacrylamide gel. F1 binding in the absence of competitor is indicated (—). The degradation product of F1 is indicated (dp) and behaves identically with respect to each of the competitors. The lower panel shows the extent of the IE110k promoter sequences in probe 3 together with the base substitutions for each variant.

spanning the EC region, although F1 binding generally predominates, the inverted CCAAT box can represent a strong site for its cognate factor. Furthermore, it is noteworthy that while both factors can bind to probe 5, we have never observed a specific complex of lower mobility, which would have been indicative of simultaneous binding. This was also the case for large 45-bp probes spanning the region. These results, together

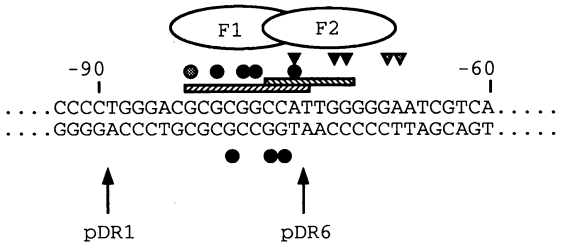

FIG. 12. Summary of the binding of F1 and F2 to the EC region of the IE110k promoter. The endpoints of the deletion constructs which define the EC region are shown below the sequence. Hatched bars indicate the extent of binding requirements determined by comparison of binding to radiolabelled probes with different endpoints together with the results of competition analysis with probes containing single- and double-base substitutions. Circles and triangles indicate bases which make contacts with F1 and F2, respectively, as determined by methylation interference analysis as discussed in the text.

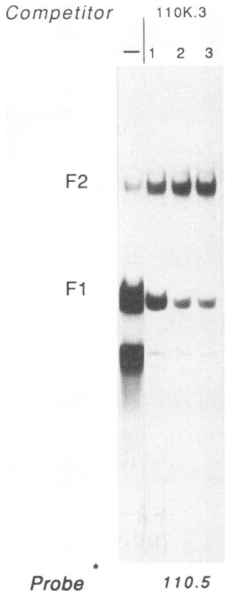

FIG. 13. Increased binding of F2 after competition with a probe which binds only F1. Aliquots of HeLa cell extract were preincubated for 10 min with increasing amounts (10, 100, and 500 ng) of unlabelled probe 3 before the addition of radiolabelled probe 5. In this case the amount of probe 5 added was limiting with respect to F2 formation, as determined in parallel probe titration experiments. Samples were incubated for a further 20 min, and F1 and F2 complexes were then separated on a 4% nondenaturing polyacrylamide gel. The control sample with no preincubation with probe 3 is also indicated (—).

with those on the shared requirements within the region from position −77 to −74, provide strong support for the proposal that F1 and F2 bind to the EC region in a mutually exclusive manner.

DISCUSSION

In this work we have examined requirements for constitutive promoter activity of the gene encoding the IE110k protein (ICP0) of HSV type 1, since a number of lines of evidence (see below) indicate that IE110k has distinctive features with respect to both transcriptional regulation and function during HSV replication. Such information is likely to be relevant to the understanding of the requirements for IE transcription in situations in which, for any of a number of reasons, Vmw65-mediated induction of IE gene expression may be inoperative. We show that while IE110k transcription can be induced by Vmw65, constitutive activity of the promoter is also extremely strong, equivalent to that of the SV40 early enhancer-promoter region. Despite the presence of multiple copies of the octamer consensus site which mediate Oct-1 binding and subsequent Vmw65 activation, the upstream region containing these motifs and other potential transcription factor-binding sites has only a minor effect on constitutive transcription. By progressive exonuclease digestion of the promoter we have identified a 15-bp region, from position −89 to −74, which is required for efficient constitutive expression from the IE110k promoter. We demonstrate that two cellular proteins bind to this region with overlapping sequence requirements. Our results indicate that one of these proteins, termed F2 in this study, is NF-Y, a CCAAT box-binding factor, which binds to the inverted CCAAT box located between positions −71 and −75. The second factor, F1, appears to be novel and binds to a region

with the sequence CGCGCGGCCAT which overlaps the 3′ end of the CCAAT box. The terminal AT of the recognition site for the second cellular factor represents, on the opposite strand, the terminal AT of the CCAAT box. These two bases and adjacent bases are critically required for the binding of both factors, indicating, together with other results from gel retardation analysis, that F1 and F2 may bind in a mutually exclusive manner.

There have been few previous reports of a detailed dissection of determinants required for HSV IE promoter activity in the absence of Vmw65 and none on the IE110k promoter. However, earlier work focusing on virion-mediated stimulation of transcription of the IE175k gene indicated that sequences spanning approximately 200 bp upstream of the IE175k promoter from −330 to −110 and containing the Vmw65 response elements could act as an enhancer-type element and increase basal expression in the absence of Vmw65 (8, 29, 30, 48). However, there are multiple GC elements with complete homology to SP1 motifs in the IE175k upstream regulatory region, and it has subsequently been shown that these sites can bind SP1, with a corresponding stimulatory effect on expression (23). Furthermore, results from studies on deletion variants of the IE175k promoter showed a correlation between the effects of progressive removal of the SP1 sites and basal expression (2). Conversely, while a recent analysis of the octamer-binding proteins indicates that the HeLa cell Oct-1 protein can independently activate transcription (24), nonetheless Oct-1 is a much weaker transactivator than, e.g., Oct-2, and the presence or addition of octamer-binding sites has comparatively little effect on basal transcription of the IE genes (41; this work). Thus, it is likely that the major determinants of the enhancer-type activity of the IE175k upstream region are the multiple SP1-binding sites. Despite this, as we have previously shown, constitutive activity of the IE175k promoter is less efficient than that of the IE110k promoter (52), and a deleted variant of the IE110k promoter which lacks all of the upstream sequences but contains the promoter-proximal F1- and F2-binding sites still exhibits 3- to 5-fold-greater levels of basal expression than the complete IE175k promoter-regulatory region and approximately 50-fold greater levels than the equivalent IE175k promoter-proximal region (39). Thus, the determinants of IE110k promoter activity differ from those of the IE175k promoter; the region upstream of the IE110k promoter, while GC rich, contains comparatively few consensus SP1-binding sites, and this region, including the multiple high-affinity Oct-1-binding sites, has little influence on constitutive transcription. The pseudorabies virus IE promoter also is more similar to the HSV IE175k promoter than to the IE110k promoter in that a series of upstream reiterated motifs containing binding sites for several factors are the main determinants of constitutive activity (4, 28). While CCAAT box-like motifs have been noted within these upstream repeats, they have been shown to bind a factor distinct from the prototypical CCAAT box factor NF-Y (28).

The core CCAAT box is involved in binding diverse factors, which led to some ambiguity in the early identification of the specific factors involved in promoter activity of a number of genes. Thus, NF-1, C/EBP, NF-Y, CBF, and CP1 each bound to elements containing the CCAAT box, (for reviews, see references 12, 22, and 50). It is now known that NF-1, C/EBP, and NF-Y are unrelated proteins whose recognition sites, while being clearly distinct and exhibiting qualitatively different contact points, nonetheless exhibit similarities or overlap in primary sequence. In addition, it is clear that CBF and almost certainly CP1 are in fact identical to NF-Y, a factor which recognizes the Y box in the promoters of major histocompatibility genes and is a ubiquitous factor involved in the expression of a number of diverse genes (7, 19, 64, 65). Thus, while there are diverse factors which can recognize the CCAAT box, except for NF-1 and C/EBP a number of these factors may be identical to NF-Y. Several lines of evidence from our results, e.g., the absence of cross competition for NF-1, the difference in methylation interference analysis between F2 and that reported for C/EBP (21) versus the similarity to that reported for NF-Y (12), the effect of flanking mutations, and recognition by an antipeptide antibody to the NF-Y site, together indicate that F2 is indeed NF-Y. The accumulative evidence from other studies, including deletion or site-directed mutagenesis of the binding site in several different promoters and in vitro transcription studies with the purified protein, associates NF-Y binding with transcriptional stimulation. In addition, $NF\text{-}Y_A$ and $NF\text{-}Y_B$, the subunits of NF-Y, contain regions of strong homology to HAP2 and HAP3, two yeast proteins which in combination bind specific UAS motifs and stimulate transcription (22, 64). It would seem reasonable to ascribe the strong constitutive activity of the IE110k promoter to NF-Y binding within the −70 region, and there are precedents for strong constitutive activity of a promoter, e.g., that of the cellular β-actin gene, being determined largely by a single NF-Y-binding site (49). However, a number of points argue that this may be an overly simplistic interpretation. Constitutive activity from the IE110k promoter, whose main determinant is within the EC region, is almost on par with that from the SV40 early enhancer-promoter region, and NF-Y has not been shown to promote such high levels of expression. Furthermore, several other promoters, including the HSV thymidine kinase promoter and adenovirus EII late promoter, contain NF-Y-binding sites which are functional (19, 69) and possess an affinity for NF-Y similar to that within the EC region, yet in our comparative analysis these promoters are considerably weaker than the IE110k promoter, exhibiting 20- to 50-fold-lower constitutive activity in parallel experiments (Fig. 1 and data not shown). It has also been shown that while the varicella-zoster virus IE promoter contains a CCAAT box motif (35) which binds a cellular factor (36), constitutive activity of this promoter nonetheless is more than 100-fold weaker than that of, e.g., the HSV IE175k promoter (35). In addition, while mutation of a CCAAT box can have a pronounced effect on certain promoters, in others, particularly those which contain adjacent or additional factor-binding sites, disruption of the CCAAT box may have a relatively minor effect. It is therefore possible that F1 contributes to the strong activity of the IE110k promoter either alone or in combination with NF-Y. Although there are precedents for the simultaneous binding of two factors to directly adjacent sites with an involvement of a particular central base for both factors, all our results indicate that F1 and NF-Y bind in a mutually exclusive manner, and it would seem unlikely that F1 could somehow act in combination with NF-Y. Thus, binding of F1 could be the major determinant in the constitutive activity of the IE110k promoter. To answer the question of the relative contribution of each factor, we are currently constructing base substitution mutants which, based on the methylation interference and binding competition studies reported here, should be selectively affected in the binding of either F1 or NF-Y.

If F1 and NF-Y are both positively acting transcription factors with mutually exclusive binding, it is possible that they are differentially utilized for IE110k promoter activity in different cell types or cell states, depending on their relative activity or abundance. The NF-Y factor at least has been shown to be involved in the expression of a variety of genes, both in ubiquitous expression and in cell-type-specific expres-

sion, and to cooperate with several different transcription factors, and the CCAAT box sequence is present in a number of cell cycle-regulated genes with altered binding of a complex to, e.g., the human thymidine kinase CCAAT box, correlating with its pattern of transcription during the cell cycle (20, 26). We are currently examining whether there is any difference in the relative abundances of F1 and NF-Y in other cell types compared with that observed here in HeLa and Vero cells, whether there are any differences induced by cell cycle progression or messenger signalling pathways, and whether these differences, if observed, correlate with IE110k promoter activity.

In conclusion, the stimulus for this study stems from the possibility that IE transcription may take place or be regulated by mechanisms independent of the Vmw65 transactivation pathway. Particularly with reference to latency it is clear, since productive cycle gene transcription is effectively repressed, that transcription of the IE110k gene is a prerequisite to synthesis of the IE110k protein and its subsequent involvement in facilitating virus gene expression. But the initial phases of IE110k transcription must take place in the absence of the virus-encoded IE transactivator Vmw65, and it is therefore relevant to examine whether the determinants involved in transcription under such circumstances are the same as the well-characterized determinants required for optimal transcription in the presence of Vmw65. To date much interest has focused on the possibility that the octamer-binding proteins play a pivotal role in IE transcription during latency and reactivation (25, 58, 62, 63), mainly because through the examination of the mechanism of action of Vmw65, the octamer-binding protein Oct-1 is the best-characterized protein involved in IE transcription. However, our results show that the presence of multiple Oct-1-binding sites has relatively little effect on the strong constitutive activity of the IE110k promoter. This does not rule out a role for octamer-binding activity in the initial phases of IE transcription, and indeed it has been shown that ocular scarification, a stimulus which reactivates latent HSV from the trigeminal ganglia, also activates Oct-1 expression (together with fos and jun) in the trigeminal neurons (62). For example, it could be that Oct-1 together with the required cellular factor CFF and a cellular equivalent of Vmw65 could mediate IE transcription. However, it may be more likely that if the Vmw65 activation pathway does play a role in reactivation, it would be later in the process, and in the face of evidence that Oct-1 binding has little influence on the IE110k promoter and that the EC region is the main determinant of activity, F1 and/or NF-Y represents a strong candidate for an important role in the initial phase of IE110k transcription. We have been able to demonstrate both F1 and NF-Y binding activity in extracts of trigeminal ganglia from latently infected animals and are currently examining whether there is any relationship between virus reactivation after ganglionic explantation and F1 or NF-Y abundance.

ACKNOWLEDGMENTS

We gratefully acknowledge the supply of antibody from D. Mathis and thank C. Goding for advice on the identification of CCAAT box promoters.

This work is funded by the Marie Curie Cancer Foundation.

REFERENCES

1. **apRhys, C. M., D. M. Ciufo, E. A. O'Neill, T. J. Kelly, and G. S. Hayward.** 1989. Overlapping octamer and TAATGARAT motifs in the VF65-response elements in herpes simplex virus immediate-early promoters represent independent binding sites for cellular nuclear factor III. J. Virol. **63:**2798–2812.
2. **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.
3. **Cai, W., and P. A. Schaffer.** 1992. Herpes simplex virus type 1 ICP0 regulates expression of immediate-early, early, and late genes in productively infected cells. J. Virol. **66:**2904–2915.
4. **Campbell, M. E., and C. M. Preston.** 1987. DNA sequences which regulate the expression of the pseudorabies virus major immediate early gene. Virology **157:**307–316.
5. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7:**2745–2752.
6. **Chen, J., and S. Silverstein.** 1992. Herpes simplex viruses with mutations in the gene encoding ICP0 are defective in gene expression. J. Virol. **66:**2916–2927.
7. **Chodosh, L. A., J. Olesen, S. Hahn, A. S. Baldwin, L. Guarente, and P. A. Sharp.** 1988. A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally interchangeable. Cell **53:**25–35.
8. **Cordingley, M. G., M. E. 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. **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.
10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.
11. **Dixon, R. A., and P. A. Schaffer.** 1980. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. J. Virol. **36:**189–203.
12. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CCAAT box-binding proteins. Cell **50:**863–872.
13. **Everett, R. D.** 1986. The products of herpes simplex virus type 1 (HSV-1) immediate early genes 1, 2 and 3 can activate HSV-1 gene expression in trans. J. Gen. Virol. **67:**2507–2513.
14. **Everett, R. D.** 1987. The regulation of transcription of viral and cellular genes by herpesvirus immediate-early gene products. Anticancer Res. **7:**589–604.
15. **Everett, R. D.** 1989. Construction and characterisation of herpes simplex virus type 1 mutants with defined lesions in immediate early gene 1. J. Gen. Virol. **70:**1185–1202.
16. **Everett, R. D., C. M. Preston, and N. D. Stow.** 1990. Functional and genetic analysis of the role of Vmw110 in herpes simplex virus replication, p. 49–76. *In* E. K. Wagner (ed.), Herpesvirus transcription and its regulation. CRC Press, Boca Raton, Fla.
17. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9:**6505–6525.
18. **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.
19. **Goding, C. R., S. M. Temperley, and F. Fisher.** 1987. Multiple transcription factors interact with the adenovirus-2 EII-late promoter: evidence for a novel CCAAT recognition factor. Nucleic Acids Res. **15:**7761–7780.
20. **Hinkley, C., and M. Perry.** 1992. Histone H2B gene transcription during *Xenopus* early development requires functional cooperation between proteins bound to the CCAAT and octamer motifs. Mol. Cell. Biol. **12:**4400–4411.
21. **Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. L. McKnight.** 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes Dev. **1:**133–146.
22. **Johnson, P. F., and S. L. McKnight.** 1989. Eucaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58:**799–841.
23. **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:**179–182.
24. **Kemler, I., E. Bucher, K. Seipel, M. Muller-Immergluck, and W. Schaffner.** 1991. Promoters with the octamer DNA motif (ATG CAAAT) can be ubiquitous or cell type-specific depending on binding affinity of the octamer site and Oct-factor concentration. Nucleic Acids Res. **19:**237–242.
25. **Kemp, L. M., C. L. Dent, and D. S. Latchman.** 1990. Octamer motif mediates transcriptional repression of HSV immediate-early genes and octamer-containing cellular promoters in neuronal cells. Neuron **4:**215–222.
26. **Knight, G. B., J. M. Gudas, and A. B. Pardee.** 1987. Cell-cycle-specific interaction of nuclear DNA-binding proteins with a CCAAT sequence from the human thymidine kinase gene. Proc. Natl. Acad. Sci. USA **84:**8350–8354.
27. **Knipe, D. M.** 1989. The role of viral and cellular nuclear proteins in herpes simplex virus replication. Adv. Virus Res. **37:**85–123.
28. **Kozmik, Z., L. Arnold, and V. Paces.** 1991. Multiple sets of adjacent mu E1 and oct-1 binding sites upstream of the pseudorabies virus immediate-early gene promoter. Virology **182:**239–249.
29. **Kristie, T. M., and B. Roizman.** 1984. Separation of sequences defining basal expression from those conferring alpha gene recognition within the regulatory domains of herpes simplex virus 1 alpha genes. Proc. Natl. Acad. Sci. USA **81:**4065–4069.
30. **Lang, J. C., D. A. Spandidos, and N. M. Wilkie.** 1984. Transcriptional regulation of a herpes simplex virus immediate early gene is mediated through an enhancer-type sequence. EMBO J. **3:**389–395.
31. **Leib, D. A., D. M. Coen, C. L. Bogard, K. A. Hicks, D. R. Yager, D. M. Knipe, K. L. Tyler, and P. A. Schaffer.** 1989. Immediate-early regulatory gene mutants define different stages in the establishment and reactivation of herpes simplex virus latency. J. Virol. **63:**759–768.
32. **Mackem, S., and B. Roizman.** 1982. Structural features of the herpes simplex virus alpha gene 4, 0, and 27 promoter-regulatory sequences which confer alpha regulation on chimeric thymidine kinase genes. J. Virol. **44:**939–949.
33. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74:**560–563.
34. **McCarthy, A. M., L. McMahan, and P. A. Schaffer.** 1989. Herpes simplex virus type 1 ICP27 deletion mutants exhibit altered patterns of transcription and are DNA deficient. J. Virol. **63:**18–27.
35. **McKee, T. A., G. H. Disney, R. D. Everett, and C. M. Preston.** 1990. Control of expression of the varicella-zoster virus major immediate early gene. J. Gen. Virol. **71:**897–906.
36. **McKee, T. A., and C. M. Preston.** 1991. Identification of two protein binding sites within the varicella-zoster virus major immediate early gene promoter. Virus Res. **20:**59–69.
37. **McLauchlan, J., S. Simpson, and J. B. Clements.** 1989. Herpes simplex virus induces a processing factor that stimulates poly(A) site usage. Cell **59:**1093–1105.
38. **Mosca, J. D., P. Pitha, and G. S. Hayward.** 1992. Herpes simplex virus infection selectively stimulates accumulation of beta interferon reporter gene mRNA by a posttranscriptional mechanism. J. Virol. **66:**3811–3822.
39. **O'Hare, P.** 1993. Unpublished data.
40. **O'Hare, P.** 1993. The virion transactivator of herpes simplex virus, p. 145–155. *In* A. J. Davison (ed.), The alpha-herpesviruses, vol. 3. Saunders Scientific Publications, Philadelphia.
41. **O'Hare, P., and C. R. Goding.** 1988. Herpes simplex virus regulatory elements and the immunoglobulin octamer domain bind a common factor and are both targets for virion transactivation. Cell **52:**435–445.
42. **O'Hare, P., and G. S. Hayward.** 1984. Expression of recombinant genes containing herpes simplex virus delayed-early and immediate-early regulatory regions and trans activation by herpesvirus infection. J. Virol. **52:**522–531.
43. **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.
44. **Perry, L. J., and D. J. McGeoch.** 1988. The DNA sequences of the long repeat region and adjoining parts of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69:**2831–2846.
45. **Persson, R. H., S. Bacchetti, and J. R. Smiley.** 1985. Cells that constitutively express the herpes simplex virus immediate-early protein ICP4 allow efficient activation of viral delayed-early genes in *trans*. J. Virol. **54:**414–421.
46. **Post, L. E., S. Mackem, and B. Roizman.** 1981. Regulation of alpha genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with alpha gene promoters. Cell **24:**555–565.
47. **Preston, C. M.** 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK. J. Virol. **29:**275–284.
48. **Preston, C. M., and D. Tannahill.** 1984. Effects of orientation and position on the activity of a herpes simplex virus immediate early gene far-upstream region. Virology **137:**439–444.
49. **Quitschke, W. W., Z. Y. Lin, L. DePonti-Zilli, and B. M. Paterson.** 1989. The beta actin promoter. High levels of transcription depend upon a CCAAT binding factor. J. Biol. Chem. **264:**9539–9546.
50. **Raymondjean, M., S. Cereghini, and M. Yaniv.** 1988. Several distinct "CCAAT" box binding proteins coexist in eukaryotic cells. Proc. Natl. Acad. Sci. USA **85:**757–761.
51. **Rice, S. A., and D. M. Knipe.** 1988. Genetic evidence for two distinct transactivation functions of the herpes simplex virus alpha protein ICP27. J. Virol. **64:**1704–1715.
52. **Roberts, M. S., A. Boundy, P. O'Hare, M. C. Pizzorno, D. M. Ciufo, and G. S. Hayward.** 1988. Direct correlation between a negative autoregulatory response element at the cap site of the herpes simplex virus type 1 IE175 (alpha 4) promoter and a specific binding site for the IE175 (ICP4) protein. J. Virol. **62:**4307–4320.
53. **Russell, J., E. C. Stow, N. D. Stow, and C. M. Preston.** 1987. Herpes simplex virus genes involved in latency in vitro. J. Gen. Virol. **68:**3009–3019.
54. **Sacks, W. R., C. C. Greene, D. P. Aschman, and P. A. Schaffer.** 1985. Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol. **55:**796–805.
55. **Sacks, W. R., and P. A. Schaffer.** 1987. Deletion mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein ICP0 exhibit impaired growth in cell culture. J. Virol. **61:**829–839.
56. **Sandri-Golden, R. M., and G. E. Mendoza.** 1992. A herpesvirus regulatory protein appears to act post-transcriptionally by affecting mRNA processing. Genes Dev. **6:**848–863.
57. **Sears, A. E., I. W. Halliburton, B. Meignier, S. Silver, and B. Roizman.** 1985. Herpes simplex virus 1 mutant deleted in the alpha 22 gene: growth and gene expression in permissive and restrictive cells and establishment of latency in mice. J. Virol. **55:**338–346.
58. **Sears, A. E., V. Hukkanen, M. A. Labow, A. J. Levine, and B. Roizman.** 1991. Expression of the herpes simplex virus 1 alpha transinducing factor (VP16) does not induce reactivation of latent virus or prevent the establishment of latency in mice. J. Virol. **65:**2929–2935.
59. **Siebenlist, U., and W. Gilbert.** 1980. Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77:**122–126.
60. **Smith, I. L., M. A. Hardwicke, and R. M. Sandri-Golden.** 1992. Evidence that the herpes simplex virus immediate early protein ICP27 acts post-transcriptionally during infection to regulate gene expression. Virology **186:**74–86.
61. **Stow, N. D., and E. C. Stow.** 1986. Isolation and characterization of a herpes simplex virus type 1 mutant containing a deletion within the gene encoding the immediate early polypeptide Vmw110. J. Gen. Virol. **67:**2571–2585.
62. **Valyi, N. T., S. Deshmane, A. Dillner, and N. W. Fraser.** 1991. Induction of cellular transcription factors in trigeminal ganglia of mice by corneal scarification, herpes simplex virus type 1 infection, and explanation of trigeminal ganglia. J. Virol. **65:**4142–4152.
63. **Valyi, N. T., S. L. Deshmane, J. G. Spivack, I. Steiner, C. I. Ace, C. M. Preston, and N. W. Fraser.** 1991. Investigation of herpes

simplex virus type 1 (HSV-1) gene expression and DNA synthesis during the establishment of latent infection by an HSV-1 mutant, in1814, that does not replicate in mouse trigeminal ganglia. J. Gen. Virol. **72:**641–649.

64. **van Huijsduijnen, R. H., X. Y. Li, D. Black, H. Matthes, C. Benoist, and D. Mathis.** 1990. Co-evolution from yeast to mouse: cDNA cloning of the two NF-Y (CP-1/CBF) subunits. EMBO J. **9:**3119–3127.
65. **Vuorio, T., S. N. Maity, and B. deCrombrugghe.** 1990. Purification and molecular cloning of the "A" chain of a rat heteromeric CCAAT-binding protein. Sequence identity with the yeast HAP3 transcription factor. J. Biol. Chem. **265:**22480–22486.
66. **Wagner, E.** 1990. Herpesvirus transcription—general aspects, p. 1–17. *In* E. K. Wagner (ed.), Herpesvirus transcription and its regulation. CRC Press, Boca Raton, Fla.
67. **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.
68. **Weinheimer, S. P., B. A. Boyd, S. K. Durham, J. L. Resnick, and D. R. O'Boyle II.** 1992. Deletion of the Vp16 open reading frame of herpes simplex virus type 1. J. Virol. **66:**258–269.
69. **Zeleznik-Le, N. J., J. C. Azizkhan, and J. P. Ting.** 1991. Affinity-purified CCAAT-box-binding protein (YEBP) functionally regulates expression of a human class II major histocompatibility complex gene and the herpes simplex virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA **88:**1873–1877.
70. **Zhu, X., J. Chen, C. S. H. Young, and S. Silverstein.** 1990. Reactivation of latent herpes simplex virus by adenovirus recombinants encoding IE-0 gene products. J. Virol. **64:**4489–4498.