DNA Sequences That Regulate Expression of a Vaccinia Virus Late Gene (L65) and Interact with a DNA-Binding Protein from Infected Cells†

JEFFREY N. MINER AND DENNIS E. HRUBY*

Center for Gene Research, Department of Microbiology, Oregon State University, Corvallis, Oregon 97331-3804

Received 13 June 1988/Accepted 28 February 1989

To be efficiently expressed in vivo, the vaccinia virus late gene, L65, requires 5'-proximal *cis*-acting elements which bind a factor from infected cells. Deletion mutagenesis and vaccinia virus helper-dependent transient expression procedures were used to demonstrate that two distinct late promoter elements direct transcription from two different start sites (proximal [+1] and distal [-92]). The -128 to -112 region was essential for L65 distal promoter function, while sequences between -59 and +50 were sufficient for L65 proximal promoter function. The proximal DNA sequences interact with a protein, binding factor I (BF-I), which was isolated and partially purified from vaccinia virus-infected cells at late times postinfection. This activity is not detectable in uninfected cells or in purified virions. This factor binds specifically to two different sites within the proximal promoter, one 5' and one 3' to the transcription start site, but does not bind to the distal promoter element.

Within infected cells, the expression of vaccinia virus (VV) genes is regulated in a distinct temporal fashion (16, 32). During replication, VV proceeds through three general phases of gene expression: immediate-early, delayed-early, and late. Since the replication cycle of VV occurs exclusively in the cytoplasm of infected cells, many of the regulatory functions affecting viral gene expression may be encoded by the virus (25). The lack of any apparent eucaryotic consensus signals within the promoter regions of VV genes tends to support this possibility (19). Understanding the mechanisms responsible for the regulation of VV gene expression is of particular importance not only as a model system for a relatively complex genetic regulatory pathway but also for facilitating the continued development of VV as a eucaryotic expression vector.

Experiments directed towards the identification and analysis of the *cis* and *trans* elements responsible for the regulated expression of VV genes in vivo are under way in a number of laboratories. The information accumulated to date indicates that VV early promoter elements are quite small, requiring only about 30 base pairs (bp) of upstream sequence to achieve regulated expression of a downstream reporter gene (7, 30). Similarly, protein factors which bind to these early VV promoters have been identified in virion extracts, partially purified (6, 31, 33), and shown to enhance transcription of cloned early genes in a fractionated cell-free transcription system prepared from VV virions (5, 6). However, in both cases the precise molecular identification of the *cis* and *trans* elements responsible for transcriptional enhancement has not yet been achieved.

Although relatively little is known regarding the signals which regulate the expression of VV early genes, even less is known concerning late-gene expression, the study of which is complicated by two unusual transcriptional properties. First, VV late-gene transcription is characterized by a lack of discrete 3' termination, generating long (>10-kilobase) transcripts traversing through other downstream late or

early genes (29). Second, recent reports have indicated that VV late messages contain a 5' poly(A) head or leader sequence of variable size (3, 17, 23).

In previous work on late-gene promoters, putative late-promoter regions have been abutted to reporter genes and assayed either by VV helper-dependent transient expression procedures or after recombination of the chimeric genes back into the viral genome (2, 7, 10, 15, 30). Both of these approaches have given equivalent results (7, 30). Several studies have provided evidence that a TAAAT(G) motif, located at or near the transcription start site, is important for late-gene expression (2, 10). However, this motif alone is apparently insufficient to direct high-level late-gene transcription, as a number of VV late genes have been shown to require additional short upstream elements for full promoter activity, including the L5 (28K) and A1 genes (15, 30).

In contrast to previously mentioned VV late-gene promoters, the regulatory sequences directing the expression of the L65 (VV D13) gene appear to be somewhat more complex. By utilizing two transcriptional start sites (distal [-92] and proximal [+1]), the L65 promoter governs the expression of a transcript encoding an abundant late VV protein of 65 kilodaltons (L65) to which rifampin resistance has been mapped (1, 24). The L65 gene is located within the *Hind*III D/A region of the VV genome (open reading frame [ORF] D13) (28) and lies within a cluster of ORFs which are noncoordinately expressed at late times postinfection (26, 27). The L65 promoter has been shown to contain two separate transcriptional elements, one between -128 and -112 (element I) and another between -235 and -175 (element II) (15).

In the present work, we have used internally controlled S1 nuclease analysis of RNA from transiently expressed L65 promoter/chloramphenicol acetyltransferase (CAT) gene constructs to analyze the sequences required for transcription from the proximal site. Signals sufficient for expression from the proximal promoter lie between -59 and +50. To investigate whether virus-encoded regulatory factors may interact with elements in the L65 promoter, cytoplasmic extracts were made from uninfected and VV-infected HeLa cells and assayed for the presence of DNA-binding proteins.

* Corresponding author.

† Technical paper 8851 from the Oregon State University Agricultural Station, Corvallis, Ore.

DNA fragments corresponding to the L65 promoter region were isolated, labeled, and used as probes in gel retention assays (8) and were found to specifically bind a relatively heat-stable, proteinase K-sensitive factor (binding factor I [BF-I]) from infected-cell extracts. After partial purification, we were able to demonstrate by competition that this factor interacts specifically with regions of the proximal L65 promoter found to be essential for expression. To map the binding sites more closely, deleted versions of the promoter were used in both competition and direct-binding studies, revealing the presence of two binding sites within the promoter.

MATERIALS AND METHODS

Plasmid construction. Recombinant plasmids were constructed, propagated, and analyzed essentially as described by Maniatis et al. (14). The pPD13:CAT plasmid contains 300 bp of upstream sequence from the L65 ORF (distal L65 promoter) fused to the bacterial CAT gene (15). pPD13L:CAT contains 440 bp of upstream sequence and 50 bp of downstream sequence from the distal and proximal L65 promoters fused to the CAT gene (Fig 1). This 490-bp *FokI-PstI* promoter fragment was trimmed with S1 nuclease and ligated to the *SmaI* site of pUC13:CAT, generating pPD13L:CAT. Plasmid pPD13L:CAT served as the starting material for the deletion series. Digestion with *KpnI* and *XbaI* followed by limited exonuclease III-S1 nuclease treatment allowed unidirectional 5' deletion into the promoter fragment. Dideoxy DNA sequencing of the plasmids confirmed the extent of the deletions (22).

Cells and virus. Growth, purification, and plaque titration of VV (strain WR) was carried out as previously described (12). BSC-40 and L cells were maintained in Earle minimum essential medium (GIBCO Laboratories) supplemented with 10% heat-inactivated fetal bovine serum, 50 µg of gentamicin sulfate per ml, and 2 mM L-glutamine at 37°C in 5% CO₂ and 95% humidity (15).

Transient expression. Transient expression assays were carried out as previously described (7, 15). Briefly, recombinant plasmids (5 µg of internal control plasmid and 10 µg of test plasmid) were coprecipitated with carrier salmon sperm DNA by using the calcium phosphate method (9). Precipitated DNA was layered onto monolayers of L cells. The monolayers were treated 3 h later with glycerol at 37°C. After a 1-h recovery period, the cells were superinfected with VV at a multiplicity of 30 PFU per cell. The infected cells were incubated at 37°C for 24 h prior to RNA extraction or CAT enzyme extract preparation.

CAT extracts and assays. Infected cells were harvested at 24 h postinfection, washed twice with ice-cold phosphate-buffered saline, centrifuged, suspended in 1 ml of CAT assay buffer (40 mM Tris chloride [pH 7.8], 1 mM EDTA, 150 mM NaCl), and incubated at 25°C for 5 min. The cells were centrifuged again and suspended in 0.1 ml of 0.25 M Tris chloride (pH 7.8). After the cells were subjected to three freeze-thaw cycles, the nuclei were removed and the extract was stored at -20°C. CAT assays were performed as previously described (15).

RNA preparation and S1 nuclease analysis. After transient expression, total cytoplasmic RNA was isolated from monolayers of transfected L cells. Cells were washed twice with phosphate-buffered saline, lysed with 2 ml of guanidine isothiocyanate solution (4 M guanidine isothiocyanate, 25 mM sodium citrate [pH 7.0], 0.5% lauryl sarcosine, 0.1 M β-mercaptoethanol), sheared by six passages through a

23-gauge needle, layered onto a 3-ml CsCl cushion (5.7 M CsCl, 0.1 M EDTA [pH 7.0]), and centrifuged for 12 h at 15°C and 35,000 rpm in an SW50.1 rotor. RNA pellets were rinsed two times with ice-cold 70% ethanol, suspended, and ethanol precipitated twice. DNA probes for the S1 nuclease analyses were isolated from agarose gels by electroelution of DNA fragments onto Whatman DE 81 paper after digestion of plasmid DNA with the appropriate restriction enzymes and agarose gel electrophoresis. The 5' ends were radiolabeled with [γ -³²P]ATP by using polynucleotide kinase. Probes to be single-end labeled were asymmetrically cut with restriction enzymes and then reisolated. Probes were denatured, annealed to 10 µg of infected-cell cytoplasmic RNA, and digested with S1 nuclease as described previously (27). Reaction products were subjected to electrophoresis in 8 M urea-8% polyacrylamide gels. Enzymes and radionucleotides were obtained from commercial suppliers (Bethesda Research Laboratories, Inc.; New England BioLabs, Inc.; Boehringer Mannheim Biochemicals; and Dupont, NEN Research Products) and used according to the instructions of the manufacturers.

Cell extract and protein fractionation. Growth and infection of HeLa cells was done by using previously described protocols (12). At either 3 or 8 h postinfection, 2 liters of infected cells was centrifuged in a Beckman TJ-6 centrifuge (2,000 rpm, 10 min), washed twice in buffer W (137 mM NaCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, 2.7 mM KCl, 0.5 mM MgCl₂) (20), and then snap frozen in liquid nitrogen and stored at -70°C. Cells prepared in this manner were stable for several months. To prepare cell extracts, twice the pellet volume of lysis buffer (0.5% Nonidet P-40, 20 mM HEPES [N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid] [pH 7.6], 2 mM magnesium acetate, 1 mM dithiothreitol [DTT]) was added and the suspension was mixed, incubated on ice for 5 min, and remixed. Nuclei and cell debris were removed by centrifugation at 10,000 rpm for 2 min in a microcentrifuge. One volume of buffer A (200 mM Tris chloride [pH 8.0], 20 mM DTT, 500 mM KCl, 0.2 mM EDTA) was added to the supernatant fluid, which was then applied to a 1-ml DEAE-cellulose (DE-52, Whatman, Inc.) column which had been equilibrated with buffer A (100 mM Tris chloride [pH 8.0], 10 mM DTT, 250 mM KCl, 0.1 mM EDTA). With buffer A used as an eluant, 12 100-µl fractions were collected and the protein concentration of each fraction was determined by using a Bio-Rad protein assay kit. Peak fractions were pooled and dialyzed for 3 h against several changes of TGED buffer (50 mM Tris chloride [pH 8.0], 1 mM EDTA, 1 mM DTT, 0.1% Nonidet P-40, 20% glycerol) containing 0.1 M KCl. A 1-ml phosphocellulose column (Bio-Rad Cellex-P) pre-equilibrated with TGED-0.1 M KCl was used to fractionate the dialyzed extract. Extract was added to the column and then step eluted with successive 1.0-ml aliquots of TGED containing 0.1, 0.35, and 0.6 M KCl followed by 4 ml of TGED containing 1 M KCl. Fractions (500 µl) were collected and assayed for protein concentration and specific binding activity (data not shown). Alternatively, the phosphocellulose column was developed with a linear gradient from 0.1 to 1 M KCl. Fractions which demonstrated the ability to bind specifically to L65 promoter sequences (BF-I) were pooled (usually fractions containing 0.8 M KCl). These fractions were dialyzed against buffer Z (25 mM HEPES [pH 7.8], 12 mM MgCl₂, 1 mM DTT, 0.1% Nonidet P-40, 20% glycerol) containing 10 mM KCl (13) and were further purified by using a salmon sperm DNA affinity column which consisted of sonicated salmon sperm DNA coupled to CNBr-activated Sepharose 4B (Pharmacia) (4). Flowthrough

FIG. 1. Plasmids and probes derived from the *Hind*III D/A region of the VV genome. The locations of the L65 (orf D13) gene and promoter and direction of transcription are shown. ▼, Previously mapped start sites at +1 (proximal) and -92 (distal) on the L65 promoter (25). Two VV L65 promoter/CAT constructs (pPD13:CAT and pPD13L:CAT) used in these experiments are drawn to scale. The probes utilized in various S1 nuclease protection experiments are shown (*, positions of radiolabeled nucleotides). Probe D13 was isolated from pPD13:CAT as a *Pvu*II fragment, labeled, digested with *Sst*I, and purified, generating a single-end-labeled fragment of 450 bp. Probe D13L was isolated from pPD13L:CAT by precisely the same protocol, generating a single-end-labeled fragment of 607 bp. The 680-bp D13Q probe was isolated from pPD13L:CAT by digestion with *Hpa*II (which cuts within the CAT gene just 5' to the *Pvu*II site), end-labeled, digested with *Eco*RI, and reisolated.

fractions were collected, and the column was step eluted with four 1-ml aliquots of buffer Z-10 mM KCl and then with two 1-ml aliquots each of buffer Z containing 0.1, 0.35, 0.6, and 1.0 M KCl. After glycerol was added to a concentration of 30%, protein concentrations were established and fractions were assayed for BF-I activity. BF-I-containing fractions were pooled, concentrated by dialysis against dry Aquacide III (Calbiochem-Behring) at 4°C, and stored at -20°C.

Gel retention assays. Binding reactions contained the protein fraction being tested, 1 µg of poly(I)-poly(C), 4 µl of buffer B (10 mM HEPES [pH 8.0], 10% glycerol, 0.5 mM EDTA, 0.5 mM DTT, 15 mM KCl), and the indicated amount of unlabeled competitor fragment in a total volume of 20 µl. Heat stability of BF-I was tested by heating the extracts prior to addition to the reaction. Sensitivity to sodium dodecyl sulfate was tested by addition of sodium dodecyl sulfate to a final concentration of 0.01%. The reactions were incubated on ice for 10 min prior to addition of 1 µl of radiolabeled DNA probe fragment (20 fmol [~10,000 cpm]). Reactions were incubated for an additional 5 min on ice and subsequently transferred to room temperature for 30 min (21). DNA-protein complexes were electrophoretically separated on low-ionic-strength, prewarmed 7.5% polyacrylamide gels (1.7 mM Tris chloride [pH 8.0], 0.37 mM sodium acetate, 1 mM EDTA) at a constant current of 20 mA with recirculated buffer. After approximately 3 h, gels were dried and radioactivity was detected by exposure

to Kodak XAR-5 film (Eastman Kodak Co., Rochester, N.Y.) with an intensifying screen at -70°C.

RESULTS

Transcription from the L65 promoter. Previous analysis of the L65 promoter has shown that transcriptional initiation from the distal site is governed by two upstream signals, one between -123 and -106 and the other between -235 and -170 (15). The experiments described below were done to delineate the sequences required for proximal start site function and determine whether these sequences specifically bind *trans*-acting factors from infected cells.

To facilitate these experiments, plasmids were assembled which contained various portions of the region upstream from the L65 ORF abutted to the CAT gene. The first plasmid (pPD13:CAT) has been described previously (15) and consists of the L65 upstream region from +1 to -300 adjacent to the CAT gene within a pUC13 vector (Fig. 1). The second plasmid (pPD13L:CAT) contained additional 3' and 5' sequences which spanned from +50 to -440 (Fig. 1). The probe fragments used to detect transcripts from each of these constructions by S1 nuclease analysis are shown below each construction (Fig. 1).

S1 nuclease protection assays using single-end-labeled probe D13 and cytoplasmic RNA generated by VV helper-dependent transient expression of plasmid pPD13:CAT produced two bands, one which corresponded to full-length

FIG. 2. S1 nuclease analysis of RNA derived from transient expression of pPD13:CAT and pD13L:CAT. S1 nuclease analysis was performed as described in Materials and Methods. (A) Lanes: 1, probe D13 digested with S1 nuclease in the presence of calf liver tRNA only; 2, reaction with probe D13 and RNA isolated from VV-infected cells transiently expressing the pPD13:CAT construct; M, size marker DNA. Numbers indicate length in bases. (B) Lane 1 is a control lane containing probe D13X hybridized to 10 μ g of calf liver tRNA and digested with S1 nuclease under conditions identical to those used for lanes 3 and 4; lane 2 is identical to lane 1 except that the probe is D13L. Lanes 3 and 4 are identical to lanes 1 and 2 except that RNA which had been isolated at 24 h postinfection from cells transiently expressing plasmids -235:D13:CAT and D13L:CAT, respectively, was used. Lane M and numbers are as in panel A.

protection of the probe at 450 nucleotides (nt) and another, at 258 nt, mapping to the distal transcriptional start site of the L65 promoter (Fig. 2A). A protected fragment of 166 nt would have been expected if transcription from the proximal start site had occurred. Likewise, with probe D13X and the RNA transcribed from pPD13:CAT (-235) (a plasmid with a small 70-bp 5' deletion [15] [Fig. 2B, lane 3]), a 258-nt band which corresponded to transcription from the distal site was protected. CAT gene expression driven by the -235 deletion mutant was identical to the full-length construction as measured by the CAT assay. An additional band was also visible at 390 nt which corresponded to transcripts reading through the 5' end of the L65 promoter and protecting the D13 probe up to the -235 position (Fig. 2B, lane 3; see below). No 166-nt band which would correspond to transcription from the proximal start site was detected. Therefore, in both the full-length pPD13:CAT construction (Fig. 2A) and the -235 deletion mutant (Fig. 2B), transcription from the proximal start site was virtually undetectable. The lack of transcripts mapping to the proximal start site suggested either that the proximal transcriptional start site had been destroyed by truncation at +1 or that additional sequences 3' or 5' to +1 were required for full proximal promoter function. Thus, the plasmid pPD13L:CAT, which contains an additional 50 bp of 3' sequence in addition to 440 bp of 5' sequence was assembled and used in similar transient expression assays.

FIG. 3. Effect of 5' deletions into the full-length L65 promoter (pPD13L:CAT) on CAT gene expression. (A) Map showing the position of each deletion tested relative to the CAT reading frame and S1 nuclease-mapped start sites. (B) CAT activity generated from the full-length construct (-440) compared with CAT activity generated from various 5' deletions. All values were normalized to CAT activity generated by the full-length promoter (-440), and the standard deviations (bars) were calculated after normalization of each of three independent trials. Column C is the CAT activity produced by the identical vector without VV sequences (pUC:CAT).

S1 nuclease analysis with RNA derived from this construct and probe D13L showed transcriptional activity from both sites (Fig. 2B, lane 4). The three protected fragments were 607, 309, and 217 nt in size, corresponding to full-length probe protection and distal and proximal start sites, respectively. This result indicates the importance of the sequences between +1 and +50 in pPD13L:CAT. Deletion of the upstream sequences to -128 did not affect the levels of transcription from the proximal start site, which strongly suggests that it is the additional sequences 3' to +1 that are required for proximal start site function (see Fig. 4, lane -128). A TAAAT motif implicated in the regulation of other late genes is located between +2 and -3 and is restored by the additional sequences.

Specific sequences involved in L65 transcriptional regulation. To further dissect the L65 promoter element, a series of 5' deletions of pPD13L:CAT were prepared which spanned both the upstream elements responsible for the transcriptional activity of the distal start site and the region immediately upstream of the proximal site (Fig. 3A). These deletion mutants were used in concert with transient expression assays. Relative expression of these mutants was measured by CAT enzyme assays (Fig. 3B) and quantitative S1 nuclease analysis (Fig. 4).

Figure 3B shows that deletion from -440 to -128 had no reproducible effect on CAT activity from pPD13L:CAT. Further deletion of 16 bp to nt -112 resulted in a twofold drop in CAT gene expression to 60% of that with the full-length promoter. This deleted region corresponds to the distal promoter element I previously discussed (15). Deletion of sequences between -59 and -35 resulted in an additional threefold decrease in CAT enzyme produced (to 20% of that made with the full-length promoter); removal of the next 26

FIG. 4. S1 nuclease analysis of RNA derived from transient expression of 5' deletion constructs with the D13L probe. (A) Each lane is designated by the deletion construct used in transient expression (top, -440 to -9). Large arrowheads mark bands of interest. FL 680, Full-length probe protection and probe rehybridization; D 368, distal site; P 274, proximal site; IC 222, internal control signal. Smaller arrowheads indicate bands generated by readthrough into the deleted promoter constructs. Size markers (in bases) are shown at the right of the panel. The tRNA lane is the identical reaction except that tRNA was substituted for VV RNA. (B) Quantitation of transcription of S1 nuclease signals from the L65 promoter/CAT deletion series. Laser densitometry of appropriate exposures of the gel shown in panel A provided data regarding transcription from both the distal and the proximal start sites. These data are shown in graph form after normalization to the internal control. The deletion mutants analyzed are shown on the x axis. The readthrough data are also included.

base pairs to -9 decreased activity fivefold below the -35 deletion signal (to 4% of that made with the full-length promoter), effectively eliminating all CAT expression above background levels. When taken together with the results of the experiments discussed earlier, this indicates that the elements which are responsible for expression of CAT from pPD13L:CAT appear to be located between -128 and -9.

To examine the effects of the deletions at the level of RNA synthesis, internally controlled S1 nuclease analysis was used on RNA generated from transient expression of the deletion series. Cells were transfected with 10 μ g of a given deletion plasmid plus 5 μ g of pPA1:CAT (pPA1:CAT is a previously described plasmid [15] which contains the promoter for ORF A1 abutted to the CAT gene). The ORF A1 promoter is a strong late promoter with a single start site. The transfected cells were infected with VV and then cytoplasmic RNA was harvested at 24 h postinfection. Probes which detected transcription from pPD13L:CAT also protected a fragment of 222 nt due to transcription from pPA1:CAT allowing internal control of transfection efficiency and RNA extraction procedures (Fig. 4). Titration of

pPA1:CAT and pPD13L:CAT plasmids shows that at the concentrations used, no competitive effects were detected which affected RNA synthesis or expression from either plasmid during transient expression (data not shown). Figure 4A shows the results of this S1 nuclease analysis with the D13Q probe (Fig. 1); protected fragments at P (274 nt) and D (368 nt) are signals from the proximal and distal start sites of parental pPD13L:CAT, respectively (Fig. 1 and 3). The internal control migrated at 222 nt and was designated IC 222. Bands marked by small arrowheads in Fig. 4 correspond to protection of the probe fragment which maps to the 5' end of the particular deletion used, suggesting that transcriptional readthrough from the circular vector into each of these deletion mutants from the 5' side occurs (Fig. 2B, lane 3 and Fig. 4A). Densitometric scanning of appropriate exposures and normalization to the internal control enabled quantitation of relative transcriptional signal strengths (summarized in Fig. 4B). Figure 4 shows that, in agreement with the CAT expression data (Fig. 3) as well as with previous work (15), this assay detected a positive regulatory region acting on transcription from the distal start site which is located

FIG. 5. Schematic representation of various probes and competitors used in gel retention assays. The L65 promoter and part of the L65 ORF are shown. Each probe is named by its endpoints on the L65 promoter. Each of these probes contains flanking pUC sequences to correct for differences in size. +50/-59, 180 bp; +1/-88, 120 bp; -181/-300, 120 bp; -99/-181, 100 bp; +50/+9, 150 bp; +1/-13, 110 bp.

between -128 and -112 and two other positive regulatory regions acting on the proximal start site located between -59 and -9. Sequences between -128 and -440 had only minor effects on the distal site in this context, in contrast to earlier work showing a twofold decrease due to these sequences in the pPD13:CAT deletion mutants (15). Deletion of the distal elements had little effect on the transcriptional activity of the proximal site. Deletion of the region between -128 and -112 reduced the signal from the distal start site by 65-fold; the proximal site was relatively unaffected. A 2.5-fold reduction in transcription from the proximal start site was seen when sequences from -59 to -35 were removed, and a further 25-fold decrease occurred upon deletion to -9. This region between -59 and -9 does not contain any previously described regulatory signals. The TAAAT motif implicated in the regulation of several late genes (19) is found at the proximal start site. This TAAAT motif is also found within the -128 to -112 region governing activity of the distal promoter (15).

VV promoter-specific binding factor. It is reasonable to predict that both the distal and proximal L65 promoter elements function *in vivo* through interaction with *trans*-activating factors that stimulate or repress transcription. Using a gel retention assay, previous workers have isolated putative transcription factors from VV virions which interact with VV early promoters (5, 6, 33). We have used a similar strategy to search for factors from VV-infected cells that bind specifically to the proximal L65 promoter region. The probes and competitors used in these experiments are shown in Fig. 5. DNA-protein complexes were resolved from free DNA by native gel electrophoresis (8). The crude cytoplasmic extract had to be subjected to both DEAE-cellulose and phosphocellulose ion-exchange chromatography steps before individual DNA-protein complexes could be resolved (see Materials and Methods). The protein eluted from the phosphocellulose column at different KCl concentrations contained a number of DNA-binding activities which, when incubated with a labeled probe containing the proximal promoter region from +50 to -59 (Fig. 5), generated a number of DNA-protein complexes that migrated more slowly than free probe in native gels. However, only the protein in the eluate from the 1 M KCl wash generated a DNA-protein complex which could be eliminated by prior incubation with a 10- to 20-fold molar excess of unlabeled +50/-59 L65 promoter fragment but not by preincubation

with the same molar ratio of a similarly sized unlabeled pUC fragment (Fig. 6A; data not shown). This activity was designated BF-I. The specificity of BF-I binding for the proximal promoter was confirmed by competition with a 10- to 20-fold molar excess of a fragment containing -99/-181 or a 10- to 20-fold molar excess of a fragment containing -181/-300 from the distal promoter (Fig. 6A). At the concentrations tested, neither of these elements affected formation of the complex. Competitor (+50/+9 and -99/-181)-to-probe (+50/-59) molar ratios of 1, 10, and 100 were used to further demonstrate the specificity of the interaction (Fig. 6B). The fragment containing the 3' portion of the L65 proximal promoter competed efficiently for binding at low molar ratios, while the competitor from the distal promoter competed for binding only above a molar ratio of 100. BF-I activity appears to be resistant to mild heating, because incubation at 65°C for 10 min did not affect complex formation, whereas incubation of the extracts at 90°C for 10 min eliminated complex formation. In addition, the BF-I complex was sensitive to either 0.01% sodium dodecyl sulfate or proteinase K in the reactions (data not shown). As a nonspecific competitor, double-stranded polyribonucleotide poly(I)-poly(C) was used. Poly(dI)-poly(dC) or poly(dA)-poly(dT) generally produced less distinct retarded bands.

By using promoter fragments derived from the pPD13L:CAT and pPD13:CAT deletion families, the binding sites could be positioned on the promoter fragment either by testing for the ability of a given deletion to compete for complex formation (Fig. 6) or by testing for binding of BF-I to each deletion mutant (Fig. 7). Competition with fragments from either 5' or 3' regions relative to the proximal start site demonstrated the existence of two separable binding sites (Fig. 5 and 6). A fragment containing sequences between +1 and -88 (1/-88) competed equally as well for binding as a fragment containing sequences between +50 and +9 (+50/+9), although neither fragment had the competition activity of the entire +50 to -59 (+50/-59) fragment (Fig. 6A). Fragments containing sequences between +1 and -13 (+1/-13) did not compete for binding. These results were confirmed by using these fragments as labeled probes (Fig. 7A, B, and C, lanes 1 and 2). Fragments containing either the +1/-88 or +50/+9 region formed complexes similar to -59/+50 (Fig. 7A and B, lanes 1 and 2), whereas a fragment with the +1/-13 region did not (Fig. 7C, lane 1 and 2). All of

FIG. 6. Gel retention assays establishing the specificity of BF-I binding to the L65 proximal promoter region. All reactions contained radioactively labeled +50/-59 probe incubated with BF-I-containing fractions from the salmon sperm DNA column under the conditions described in Materials and Methods. (A) Lane 1 contains no competitor; lanes 2 through 13 are identical except that unlabeled competitor DNA is included in the reaction. Lanes: 2 and 3, +50/-59 (proximal promoter) 10- and 20-fold molar excess, respectively; 4 and 5, +1/-88 (5' portion of proximal promoter) 10- and 20-fold molar excess, respectively; 6 and 7, -99/-181 (region upstream from distal promoter [element II]) 10- and 20-fold molar excess, respectively; 8 and 9, -99/-181 (distal promoter [element I]) 10- and 20-fold molar excess, respectively; 10 and 11, +50/+9 (3' portion of proximal promoter) 10- and 20-fold molar excess, respectively; 12 and 13, +1/-13 (proximal start site only) 10- and 20-fold molar excess, respectively; 14, probe without protein. (B) Lanes (from left to right): 1, no protein; 2, no competitor; 3 through 5, 1-, 10-, and 100-fold molar excesses of the +50/+9 competitor, respectively; 6 through 8, 1-, 10-, and 100-fold molar excesses of the -181/-300 competitor, respectively.

the fragments in the binding and competition studies contained 88 bp of flanking sequences of the pUC vector from which they were derived; however, these flanking sequences were not involved in the binding, because the fragment +1/-13, which contains these same pUC sequences, did not form a complex itself or competitively inhibit +50/-59 complex formation.

Because the data suggest the presence of two BF-I-binding sites within the L65 proximal promoter and because fractions containing BF-I were not highly purified, it was of

interest to determine whether one binding site could compete effectively for BF-I binding to the other site. The results of this so-called cross-competition experiment are shown in Fig. 7A and B. In both cases, each binding site was capable of competing for BF-I bound to the other site, which suggests that the same factor interacted with each fragment.

Increasing the amount of BF-I protein in a reaction with the +50/-59 probe generated what appeared to be a multimeric form of the probe-BF-I complex (Fig. 8A). This result would be expected if the +50/-59 fragment contained two

FIG. 7. Specific binding of BF-I to two sites near the L65 promoter and cross-competition between them. (A) Standard binding reaction including probe +1/-88. Lanes (from left to right): 1 and 2, no BF-I protein and no competitor, respectively; 3 through 5, results of binding reactions which contained 1-, 10-, and 100-fold molar excesses of +1/-88 unlabeled competitor fragment, respectively; 6 through 8, competition with 1-, 10-, and 100-fold molar excesses of the unlabeled +50/+9 fragment, respectively. (B) Binding reactions which contain labeled probe +50/+9. Lanes (from left to right): 1 and 2, controls without and with BF-I protein, respectively; 3 through 6, binding reactions which contain 1-, 10- and 100-fold molar excesses of the +50/+9 fragment, respectively; 7 to 9, binding reactions which contain 1-, 10-, and 100-fold molar excesses of +1/-88 unlabeled competitor fragment, respectively. (C) Binding reaction with probe +1/-13. Lane 1 and 2 are without and with BF-I protein, respectively.

binding sites which were both bound by BF-I at a high protein-to-probe ratio. Competition experiments showed that both complexes were subject to competition by fragments from the proximal promoter region. However, complex II was affected at a lower concentration of competitor compared with complex I, which is consistent with this model (Fig. 6A; data not shown). It follows that a fragment containing only a single binding site (+1/-88) should only generate a single bound complex, as shown in Fig. 8B.

DISCUSSION

To investigate *cis*-acting signals affecting this promoter, S1 nuclease analysis of RNA generated from transiently expressed promoter-CAT plasmids was used to monitor the effects of various promoter deletion mutations on transcription efficiency. The pPD13:CAT plasmid initially used in these studies showed undetectable levels of transcription from the proximal site. In contrast, high levels of genomic L65-specific RNA from the proximal site were generated

FIG. 8. Effect of increasing protein concentration on retarded species. BF-I was added to binding reactions in concentrations of 0, 50, 100, 200, and 400 ng with both +50/-59 and +1/-88 probes as indicated above the lanes. Complex I corresponds to the usual BF-I complex, and complex II is the additional species formed at high protein concentrations, with the probe containing two binding sites. In the 400-ng lane, the +1/-88 probe was retained at the origin of the gel (not shown), similar to that seen in the 400-ng lane with probe +50/-59.

during infection. To analyze sequences required for transcription from the proximal site, additional VV sequences, both 5' and 3' to those already present in pPD13:CAT, were included in another plasmid (pPD13L:CAT). S1 nuclease analysis of RNA from this plasmid showed that the additional sequences activated the proximal site without affecting the distal site. Deletion of the additional 5' sequences -440 to -128 had no effect on the level of transcription from the proximal site, strongly suggesting that the additional 50 bp of 3' sequence contained additional information required for transcription from the proximal site. These sequences contain part of a TAAAT motif found at the proximal start site, and it is possible that restoration of this signal was necessary for the transcriptional activity of the proximal promoter.

To dissect functional regions within the proximal promoter, deletions that removed the distal start site and the region up to and including the proximal start site were made in pPD13L:CAT. These plasmids were used in transient expression assays, and the level of CAT enzyme produced from each construct was measured. Cotransfection of a control plasmid in these experiments allowed quantitative S1 nuclease analysis of the RNA generated by these plasmids. Results obtained from the S1 nuclease analysis and CAT assays generally agreed, although the CAT assays presumably measured expression from both sites, whereas the S1 nuclease analysis differentiated between the two. Deletion of

the region from -440 to -128 has been previously shown to cause a twofold decrease in CAT gene expression from pPD13:CAT. We were unable to fully confirm this result by using the longer pPD13L:CAT construct. In this case, CAT activity was not affected significantly when these sequences were deleted, and S1 nuclease analysis showed only a moderate 1.4-fold decrease in transcription from the proximal start site. The distal start site was also relatively unaffected by this deletion. Removal of the next 16 bp to -112 effectively eliminated the signal from the distal start site, causing a 60-fold reduction in its intensity. These results confirm and refine previous work showing the presence of a distal promoter regulatory element between -128 and -106. Taken together, these deletions map the 5' end of a distal promoter regulatory element to a 12-bp segment between -128 and -112, the presence of which caused a 60-fold increase in the level of transcription from a site located 30 bp downstream. The start site affected by this element maps to a TAAAG at -92 by S1 nuclease analysis of both genomic RNA (27) and RNA from transient expression of pPD13:CAT and pPD13L:CAT (Fig. 4). The TAAAG start site is downstream from a TAAAT motif (at -113 to -117) which has been shown to be important for the expression of several other late genes (2, 10, 19). In these previously characterized promoters, the TAAAT signal is located at or within a few base pairs of the apparent start site. In the context of the L65 distal promoter, the TAAAT motif may

be functioning in another role since it lies within an element shown to be crucial to a transcription start site 30 bp downstream. It is also possible that in this context that TAAAT motif is nonessential and that other signals within this short segment constitute the distal promoter.

The sequences responsible for the transcriptional activity of the proximal promoter are located further downstream from the distal promoter, between -59 and -9. Sequences between -59 and -35 were responsible for a 2.5-fold effect on transcription from the proximal start site. When the remaining sequences between -35 and -9 were deleted, an additional 20-fold decrease in activity occurred, essentially eliminating the signal and thus mapping the 5' end of the element to this region. Sequence similarity searches of this segment have revealed none of the previously characterized VV regulatory elements.

In addition to the S1 nuclease signals generated by the two start sites and the internal control, a signal was present in each lane in Fig. 4 and in lane 3 in Fig. 2B, which corresponds to readthrough transcripts from each of the deleted plasmids. The 680-bp probe was complementary to the deletions up to the deletion endpoint; any RNA generated which read through the junction was detected as a higher-molecular-weight species which migrated according to the extent of the deletion. The presence of readthrough transcripts is explained by the fact that transcripts from the L65 promoter do not terminate but read around the circular plasmid and back into the promoter region. The S1 nuclease-mapped RNA is in fact bona fide mature VV late RNA, since both the distal and proximal start site-derived transcripts contained a 5' poly(A) leader sequence (J. Miner, unpublished results).

The *trans*-activation of these promoters during transient expression in VV-infected cells implies that diffusible factors are involved in recognition of and transcription from these sequences. A factor (BF-I) that specifically binds to the proximal promoter region shown to be involved in regulation of the proximal site has been detected in infected-cell extracts at late times postinfection. This activity was not detectable in uninfected cells, virion extracts, or infected cells at early times postinfection (data not shown). A specific complex was formed in gel retention assays which was subject to competition only with fragments containing parts of the proximal promoter. Fragments containing the distal promoter (-181/-300) did not compete for complex formation, nor did fragments containing 120 bp of sequence upstream from the distal promoter (-99/-181) or fragments from pUC19 (Fig. 6A and B). This suggests that BF-I binding is specific for the proximal promoter region. BF-I is sensitive to proteinase K digestion, heat, and sodium dodecyl sulfate, which strongly suggests that BF-I is a protein.

Using appropriate fragments from the deletion constructs described above, we have mapped what appear to be two BF-I-binding sites between -13 and -59 and between +50 and +9 on the proximal promoter. Nonoverlapping fragments from the proximal promoter (+50/+9 and +1/-88) formed specific complexes with BF-I, while fragments corresponding to other regions did not (Fig. 7A to C). That both sites can be filled at once in vitro was implied by results shown in Fig. 8. When a probe with two putative binding sites (+50/-59) was used, another more slowly migrating complex was generated at high BF-I-to-probe ratios. Filling of these two sites does not appear to be highly cooperative. These two sites presumably bind identical BF-I molecules; this presumption is based on the fact that cross-competition experiments show that fragments containing one binding site

can compete for binding of BF-I to the other and vice versa (Fig. 7A and B). It should be noted, however, that in several experiments, the +50/+9 fragment competed for BF-I binding less well than did the +1/-88 fragment, implying a qualitative difference in BF-I affinity for the two sites (Fig. 6B and 7A and B). Sequences 3' to the proximal start site bound to BF-I, but it is not yet clear which portion of the 3' element is responsible for activating the proximal promoter in pPD13L:CAT. The TAAAT motif that is located between +2 and -3 is likely to be important for proximal promoter function, which leaves open the importance of the 3' BF-I-binding site, especially in light of its reduced affinity for this site. The same proximal DNA sequences which are responsible for transcriptional activation of the proximal site also bound avidly to BF-I. The association of BF-I with transcriptionally important DNA sequences has several interesting implications. First, it is possible that BF-I functions to modify the structure of the DNA which lies between the two proteins, allowing an entry point for VV RNA polymerase. Second, it is possible that BF-I interacts directly with VV RNA polymerase to provide a surface that positions the polymerase at the start site. The fact that BF-I binds only to the proximal promoter and not to the distal promoter may in part explain the differential regulation of these two start sites (27).

In addition, we have preliminary data indicating that this factor also binds to the VV thymidine kinase (TK) early promoter (L. Wilson and J. Miner, unpublished results). It may be relevant that TK RNA synthesis is negatively regulated at 2 to 3 h postinfection, which is the same time that transcription from the L65 promoter is beginning (11). The sequence AATTA(T/A)A is conserved between the 5' and 3' portions of the L65 proximal promoter and is found in the inverted orientation at -43 within the TK promoter. The assessment of the importance of this sequence will await DNase I footprinting of BF-I on the L65 and TK promoters and demonstration that BF-I occupancy of these sites correlates with transcriptional activation in vitro and in vivo. Thus far, footprinting of DNA-binding proteins on VV DNA has proved to be problematic. This may be in part due to the extreme A/T richness of the VV promoter regions or the relatively low affinity of the binding factors or both. In either case it is likely that it will be necessary to purify large quantities of biologically active BF-I before such studies can be undertaken. Perhaps of more immediate importance will be to make use of recently described VV cell-free transcription systems (18, 23) which are responsive to exogenously added VV late genes to determine whether BF-I has a positive or negative effect on transcription from the L65 proximal promoter element.

ACKNOWLEDGMENTS

We thank L. Wilson and S. Strain for their assistance with several of these experiments. We also thank S. Weinrich, L. Wilson, J. Carrington, W. Dougherty, and the Hruby laboratory for helpful discussions and encouragement.

These studies were supported by Public Health Service grant AI-20563 and Research Career Development Award AI-0666 to D.E.H. from the National Institutes of Health. J.N.M. was supported by Public Health Service Molecular and Cellular Biology Training Grant GM07774-07 from the National Institutes of Health.

LITERATURE CITED

1. Baldick, C., and B. Moss. 1987. Resistance of vaccinia virus to rifampicin conferred by a single nucleotide substitution near the predicted NH₂ terminus of a gene encoding a M_r 62,000 polypeptide. *Virology* 156:138-147.

2. Bertholet, C., P. Stocco, E. Van Meir, and R. Wittek. 1986. Functional analysis of the 5' flanking sequence of a vaccinia virus late gene. *EMBO J.* 5:1951-1957.
3. Bertholet, C., E. Van Meir, B. ten Heggeler-Bordier, and R. Wittek. 1987. Vaccinia virus produces late mRNAs by discontinuous synthesis. *Cell* 50:153-162.
4. Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* 235:47-52.
5. Broyles, S., and B. Moss. 1988. DNA-dependent ATPase activity associated with vaccinia virus early transcription factor. *J. Biol. Chem.* 263:10761-10766.
6. Broyles, S., L. Yuen, S. Shuman, and B. Moss. 1988. Purification of a factor required for transcription of vaccinia virus early genes. *J. Biol. Chem.* 263:10754-10760.
7. Cochran, M. A., M. Mackett, and B. Moss. 1985. Eucaryotic transient expression system dependent on transcription factors and regulatory DNA sequences of vaccinia virus. *Proc. Natl. Acad. Sci. USA* 82:19-23.
8. Fried, M., and D. M. Crothers. 1981. Equilibrium and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* 9:6505-6525.
9. Graham, F., and A. J. Van Der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52:456-467.
10. Hanggi, M., M. Bannwarth, and H. G. Stunnenberg. 1986. Conserved TAAAT motif in vaccinia virus late promoters: overlapping TATA box and site of transcription initiation. *EMBO J.* 5:1071-1076.
11. Hruby, D. E., and A. Ball. 1981. Control of expression of the vaccinia virus thymidine kinase gene. *J. Virol.* 40:456-464.
12. Hruby, D. E., L. A. Guarino, and J. R. Kates. 1979. Vaccinia virus replication. I. Requirement for the host-cell nucleus. *J. Virol.* 29:705-715.
13. Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* 83:5889-5893.
14. Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual, p. 90-93. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. Miner, J. N., S. L. Weinrich, and D. E. Hruby. 1988. Molecular dissection of *cis*-acting regulatory elements from 5'-proximal regions of a vaccinia virus late gene cluster. *J. Virol.* 62:297-304.
16. Moss, B. 1985. Replication of poxviruses, p. 685-803. In B. N. Fields, D. M. Knipe, and R. M. Chanock (ed.), *Virology*. Raven Press, N.Y.
17. Patel, D., and D. Pickup. 1987. Messenger RNA's of a strongly-expressed late gene of cowpox virus contain 5' terminal poly(A) sequences. *EMBO J.* 6:3787-3794.
18. Puckett, C., and B. Moss. 1983. Selective transcription of vaccinia virus genes in template dependent soluble extracts of infected cells. *Cell* 35:441-448.
19. Rosel, J. L., P. L. Earl, J. P. Weir, and B. Moss. 1986. Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the *HindIII* H genome fragment. *J. Virol.* 60:436-449.
20. Rosenfeld, P. J., and T. J. Kelly. 1986. Purification of nuclear factor I by DNA recognition site affinity chromatography. *J. Biol. Chem.* 261:1398-1408.
21. Samuels, M., A. Fire, and P. A. Sharp. 1982. Separation and characterization of factors mediating accurate transcription by RNA polymerase II. *J. Biol. Chem.* 257:14419-14427.
22. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
23. Schwer, B., J. Visca, C. Vos, and H. G. Stunnenberg. 1987. Discontinuous transcription or RNA processing of vaccinia virus late messengers results in a 5' poly(A) leader. *Cell* 50:163-169.
24. Tartaglia, J., A. Piccini, and E. Paoletti. 1986. Vaccinia virus rifampicin-resistance locus specifies a late 63,000 Da gene product. *Virology* 105:45-54.
25. Villarreal, E. C., N. A. Roseman, and D. E. Hruby. 1984. Isolation of vaccinia virus mutants capable of replicating independently of the host cell nucleus. *J. Virol.* 51:359-366.
26. Weinrich, S. L., and D. E. Hruby. 1986. A tandemly-oriented late gene cluster within the vaccinia virus genome. *Nucleic Acids Res.* 14:3003-3015.
27. Weinrich, S. L., and D. E. Hruby. 1987. Noncoordinate regulation of a vaccinia virus late gene cluster. *J. Virol.* 61:639-645.
28. Weinrich, S. L., E. Niles, and D. E. Hruby. 1985. Transcriptional and translational analysis of the vaccinia virus late gene L65. *J. Virol.* 55:450-457.
29. Weir, J. P., and B. Moss. 1984. Regulation of expression and nucleotide sequence of a late vaccinia virus gene. *J. Virol.* 51:662-669.
30. Weir, J. P., and B. Moss. 1987. Determination of the transcriptional regulatory region of a vaccinia virus late gene. *J. Virol.* 61:75-80.
31. Wilson, E. M., C. Edbauer, and D. Hruby. 1988. Characterization of a binding factor that interacts with the sequences upstream of the vaccinia virus thymidine kinase gene. *Virus Genes* 2:31-48.
32. Wittek, R. 1982. Organization and expression of the poxvirus genome. *Experientia* 38:285-410.
33. Yuen L., A. Davison, and B. Moss. 1987. Early promoter binding factor from vaccinia virions. *Proc. Natl. Acad. Sci. USA* 84:6069-6073.