

## NOTES

# Structure of African Swine Fever Virus Late Promoters: Requirement of a TATA Sequence at the Initiation Region

RAMÓN GARCÍA-ESCUADERO\* AND ELADIO VIÑUELA

Centro de Biología Molecular “Severo Ochoa” (CSIC-UAM), Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

Received 16 March 2000/Accepted 30 May 2000

**A number of mutations, including deletions, linker scan substitutions, and point mutations, were performed in the promoter of the late African swine fever virus (ASFV) gene coding for the capsid protein p72. The consequences of the mutations in terms of promoter activity were analyzed by luciferase assays using plasmids transfected into infected cells. The results showed that the promoter function is contained between nucleotides –36 and +5 relative to the transcription initiation site. Moreover, two major essential regions for promoter activity, centered at positions –13 and +3, were located along the 41-bp sequence, the latter mapping in the transcription start site. Sequence alignment with other ASFV late promoters showed homology in the region of transcriptional initiation, where the presence of the sequence TATA was observed in most of the promoters. Substitution of these four residues in three other late viral promoters strongly reduced their respective activities. These results show that *cis*-acting control elements of ASFV p72 gene transcription are restricted to a short sequence of about 40 bp and suggest that transcription of late genes is initiated around a TATA sequence that would function as an initiator element.**

African swine fever virus (ASFV) is a large enveloped virus with icosahedral morphology whose genome is a double-stranded DNA molecule of about 170 kbp. Although the early events in the replication of the viral DNA occur in the nucleus of the host cells (17), ASFV is mainly a cytoplasmic virus. Virion assembly takes place in discrete cytoplasmic areas called viral factories, where the virus becomes engulfed by a two-membraned collapsed cisterna derived from the endoplasmic reticulum (6). Mature particles contain the enzymes required for early RNA synthesis and processing (19, 30, 31), most of which are known to be encoded by the virus genome (34). ASFV gene expression appears to follow a cascade mechanism similar to that described for poxviruses (22). Thus, four temporal classes of ASFV genes have been described. The expression of immediate-early and early genes begins immediately after infection, the latter being silenced at the time of maximal DNA replication (3). Intermediate gene expression is dependent on the synthesis of viral DNA, decreasing sharply when late mRNA synthesis becomes maximal (29).

At present, very little is known about the molecular mechanisms controlling ASFV gene transcription. Individual genes are tightly regulated independent transcription units whose mRNA start sites are located at a short distance from the corresponding translation initiation codon (3, 29). However, the structure of the promoters and the identity of transcription factors which regulate the expression of viral genes are unknown. In order to investigate the structural requirements for late promoter activity, we have obtained a number of mutants of the capsid protein p72 gene promoter and analyzed their transcriptional activities when directing the expression of the

luciferase gene in transient-expression assays with ASFV-infected cells. The results obtained have allowed us to localize *cis*-acting elements required for late gene transcription and suggest the existence of sequence specificity in the region of transcriptional initiation.

**Transient expression of luciferase gene regulated by the p72 promoter.** To study the promoter sequences required for transcription of ASFV late genes, we decided to analyze the 5' flanking region of the gene encoding protein p72. This protein is the major component of the icosahedral capsid of the virus (10, 18) and constitutes about 35% of the virion mass (7). Therefore, we assumed that the promoter that controls its expression is one of the strongest late-functioning promoters and could thus be very useful for our purposes. Previously, we demonstrated that a DNA fragment including the first eight codons of the p72 gene and the 197-bp upstream sequence was capable of directing the expression of the *lacZ* gene at late times postinfection, either from a recombinant virus (v72-βgal) or from a plasmid transfected into ASFV-infected cells (27). We therefore cloned the firefly luciferase gene downstream of the 197-bp upstream sequence and the first A of the translation initiation codon of the p72 gene. A 218-bp PCR fragment was obtained using ASFV DNA as a template and oligonucleotides 5'-A and 3'-A (Table 1), which were cut with restriction enzymes *Bam*HI and *Bgl*II and cloned into *Bam*HI-linearized pUC118, generating plasmid pp72. The luciferase gene of *Photinus pyralis* was then inserted as a 1.4-kbp *Bam*HI fragment from pKluc (16) into the pp72 plasmid that had been linearized with *Bam*HI enzyme, thus obtaining the pp72.luc plasmid, in which the reporter is placed downstream of the 5' flanking sequence of the p72 gene. Plasmid pp72.luc was capable of inducing the synthesis of high levels of luciferase at late times postinfection when transfected into ASFV-infected Vero cells (Fig. 1). A kinetic analysis of transient expression of luciferase showed that maximal levels of reporter activity were obtained

\* Corresponding author. Mailing address: Centro de Biología Molecular “Severo Ochoa” (CSIC-UAM), Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain. Phone: 34-91-3978438. Fax: 34-91-3974799. E-mail: rgescudero@cbm.uam.es.

TABLE 1. Names and sequences of the oligonucleotides used for the generation of promoter mutants

Oligonucleotide name <sup>a</sup>	Sequence <sup>b</sup>	Promoter(s)
<b>5' and 3' deletions</b>		
5'-A	<u>cgcgagatc</u> GGGTCGCCGAGAAAAGTCAAAAGG	<i>p72</i> , <i>p72.1</i> , <i>p72.2</i>
5'-B	<u>cgcgagat</u> CTAGCAAGAGCGTGTCAATAATT	<i>p72.3</i> , <i>p72.5</i>
5'-C	<u>cgcgaga</u> TCITTTGTTATTATCAAGATCC	<i>p72.4</i>
3'-A	<u>gcgcgatcc</u> TATATAATGTTATAAAAAATAATT	<i>p72</i> , <i>p72.3</i> , <i>p72.4</i>
3'-B	<u>gcgcgatcc</u> AATTATTGACACGCTCTTGCT	<i>p72.1</i>
3'-C	<u>gcgcgatcc</u> GGATCTTGATAATAACAAAGA	<i>p72.2</i> , <i>p72.5</i>
6A	<u>gatc</u> TATTTAATAAAAAACAATAAAATATTTTTATAACATTATATAg	
6B	<u>gatcc</u> TATATAATGTTATAAAAAATAAATTATTGTTTTTATAAAATA	<i>p72.6</i>
7A	<u>gatc</u> TATTTTTATAACATTATATAg	
7B	<u>gatcc</u> TATATAATGTTATAAAAAATA	<i>p72.7</i>
<b>Linker scan substitutions</b>		
5'-8	<u>gcggtaccgcg</u> gcATAAAAACAATAAAATTATTTTTATAACATTATATAggatccccccc	<i>p72.8</i>
5'-9	<u>gcggtacc</u> TATTTA <u>gcg</u> gcAACAATAAAATTATTTTTATAACATTATATAggatccccccc	<i>p72.9</i>
5'-10	<u>gcggtacc</u> TATTTAATAAA <u>gcg</u> gcTAAATTATTTTTATAACATTATATAggatccccccc	<i>p72.10</i>
5'-11	<u>gcggtacc</u> TATTTAATAAAAAACA <u>gcg</u> gcTATTTTTATAACATTATATAggatccccccc	<i>p72.11</i>
5'-12	<u>gcggtacc</u> TATTTAATAAAAAACAATAAA <u>gcg</u> gcTTATAACATTATATAggatccccccc	<i>p72.12</i>
5'-13	<u>gcggtacc</u> TATTTAATAAAAAACAATAAAATTATTT <u>gcg</u> gcACATTATATAggatccccccc	<i>p72.13</i>
5'-14	<u>gcggtacc</u> TATTTAATAAAAAACAATAAAATTATTTTTATA <u>gcg</u> gcATATAggatccccccc	<i>p72.14</i>
5'-15	<u>gcggtacc</u> TATTTAATAAAAAACAATAAAATTATTTTTATAACATT <u>gcg</u> cgatccccccc	<i>p72.15</i>
3'-G	<u>gggggggggatcc</u>	From <i>p72.8</i> to <i>p72.15</i>
<b>Single-nucleotide substitutions<sup>c</sup></b>		
5'-D	<u>cgcggtacc</u> TATTTAATAAAAAACAATAAA <u>T</u> A	From <i>p72.16A</i> to <i>p72.21T</i>
3'-16	<u>gcgcgatcc</u> TATAT <u>B</u> ATGTTATAAAAAATAA <u>T</u> TATTGTTTT	<i>p72.16A</i> , <i>p72.16C</i> , <i>p72.16G</i>
3'-17	<u>gcgcgatcc</u> TATAVAATGTTATAAAAAATAA <u>T</u> TATTGTTTT	<i>p72.17C</i> , <i>p72.17G</i> , <i>p72.17T</i>
3'-18	<u>gcgcgatcc</u> TAT <u>B</u> TAAATGTTATAAAAAATAA <u>T</u> TATTGTTTT	<i>p72.18A</i> , <i>p72.18C</i> , <i>p72.18G</i>
3'-19	<u>gcgcgatcc</u> TAVATAATGTTATAAAAAATAA <u>T</u> TATTGTTTT	<i>p72.19C</i> , <i>p72.19G</i> , <i>p72.19T</i>
3'-20	<u>gcgcgatcc</u> <u>T</u> BTATAATGTTATAAAAAATAA <u>T</u> TATTGTTTT	<i>p72.20A</i> , <i>p72.20C</i> , <i>p72.20G</i>
3'-21	<u>gcgcgatcc</u> <u>V</u> ATATAATGTTATAAAAAATAA <u>T</u> TATTGTTTT	<i>p72.21C</i> , <i>p72.21G</i> , <i>p72.21T</i>
<b>TATA substitutions in <i>p10</i>, <i>p11.5</i>, and <i>p54</i> promoters</b>		
5'-p11.5	<u>cgcggtacc</u> TTTAAAATAAGCCATTTAAAGATTTAGAATTTA	<i>p11.5</i> , <i>p11.5mut</i>
3'-p11.5	<u>gcgcgatcc</u> TATACATATAAAATCTAAATCTT	<i>p11.5</i>
3'-p11.5mut	<u>gcgcgatcc</u> gcgcCATATAAAATCTAAATCTT	<i>p11.5mut</i>
5'-p10	<u>cgcggtacc</u> ATAATAAAAAATATTTTTACTTTTTTTCTTC	<i>p10</i> , <i>p10mut</i>
3'-p10	<u>gcgcgatcc</u> TATATTATGAAGAAAAAAAAGTA	<i>p10</i>
3'-p10mut	<u>gcgcgatcc</u> gcgcTTATGAAGAAAAAAAAGTA	<i>p10mut</i>
5'-p54	<u>cgcggtacc</u> GTAATTTTCATTGCGCCACAACATTTTATATAT	<i>p54</i> , <i>p54mut</i>
3'-p54	<u>gcgcgatcc</u> TATAAATAATATATAAAAAATGTT	<i>p54</i>
3'-p54mut	<u>gcgcgatcc</u> gcgcAATAATATATAAAAAATGTT	<i>p54mut</i>

<sup>a</sup> Oligonucleotides are listed according to the type of mutagenesis for which each oligonucleotide was used.

<sup>b</sup> Nonviral sequences are in lowercase letters, and recognition sequences for restriction enzymes *Bgl*II, *Bam*HI, and *Kpn*I are underlined.

<sup>c</sup> Positions indicated as **B** are C, G, or T residues, and positions indicated as **V** are A, C, or G residues.

at 21 h postinfection. Moreover, when the cells were infected in the presence of AraC, a specific inhibitor of the replication of ASFV DNA and therefore of late transcription, the luciferase activity was almost undetectable. Therefore, we conclude that the 198-bp DNA fragment cloned contains the promoter elements of the *p72* gene.

**Location and length of the *p72* promoter.** In order to determine the location of essential residues within the *p72* promoter sequence, a deletion analysis was first undertaken. A series of luc plasmids containing 5', 3', or 5' and 3' promoter sequence deletions was constructed. Five mutants (*p72.1* to *p72.5*) were obtained by PCR from pp72.luc by using the oligonucleotides 5'-A, 5'-B, 5'-C, 3'-A, 3'-B, and 3'-C (Table 1). The amplified fragments were cloned into pUC118 as described for plasmid pp72. Mutants *p72.6* and *p72.7* were obtained by hybridization of oligonucleotides 6A with 6B and 7A with 7B, respectively

(Table 1). The resulting hybrids contained one end that was compatible with the *Bam*HI enzyme in such a way that the cloning of one copy of the fragments into a *Bam*HI site resulted in the retention of a single *Bam*HI site in the derived plasmids. After annealing, the DNA fragments were incubated with T4 polynucleotide kinase and ATP and were inserted into the plasmid pUC118, which was linearized with *Bam*HI endonuclease. The luc gene was cloned in the plasmids containing mutants *p72.1* to *p72.7* as in the case of pp72.luc. A control plasmid, p72GAL10T, in which the *p72* promoter and nucleotides up to the eighth codon of the *p72* gene are fused to the *lacZ* gene (16), was cotransfected with luc plasmids for 1 h into Vero cells just prior to infection with ASFV. Luciferase activities were normalized to that of  $\beta$ -galactosidase.

Analysis of the mutants showed that 3' deletions completely abolished luciferase activity (*p72.1*, *p72.2*, or *p72.5*) (Fig. 2). By

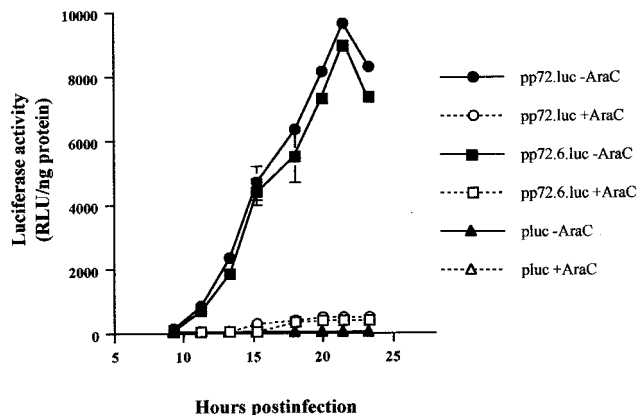


FIG. 1. Regulated expression of the luciferase gene by the *p72* late promoter. Preconfluent monolayers of Vero cells ( $1.5 \times 10^5$ ) were transfected for 1 h with 16  $\mu$ g of Lipofectamine reagent (Gibco BRL) and 1  $\mu$ g of pluc, pp72.luc, or pp72.6.luc plasmids. Transfected cells were then infected with the BA71V strain of ASFV (13) at a multiplicity of infection of 5 PFU per cell in the absence or presence of AraC (40  $\mu$ g/ml). Cells were washed and harvested at different times postinfection, and the luciferase activity was measured as previously described (16). The data are represented as the average  $\pm$  the standard deviation of three assays. Luciferase values were normalized to the protein content.

contrast, deletion of 5' sequences in mutants *p72.3*, *p72.4*, and *p72.6* did not affect promoter activity. The results indicate that the promoter elements are located within the 41 bp at the 3' end of the sequence. More importantly, this region (*p72.6* promoter) contains all the essential residues which drive late expression of a reporter gene, as shown by time course transient-expression assays and the inhibitory effect of AraC in infected Vero cells (Fig. 1). Moreover, the sites for transcription initiation of the luc gene under the control of this promoter in transfection assays are the same as those of the *p72* gene in the viral genome. Thus, as can be seen in Fig. 3, the 5' ends of the reporter gene mRNA map at residues  $-5$  to  $-2$  relative to the translation initiation codon, as did those described for the *p72* gene mRNA (26). We will designate in this work the second residue ( $-4$ ) used for initiation as the  $+1$  position.

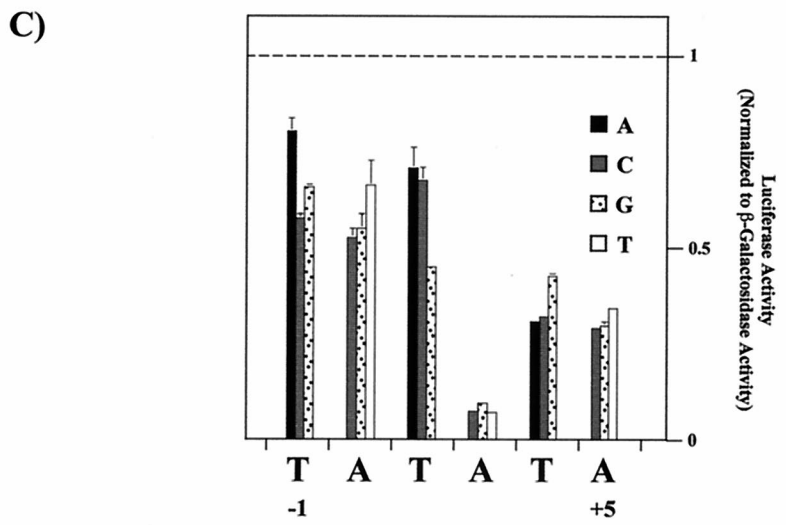
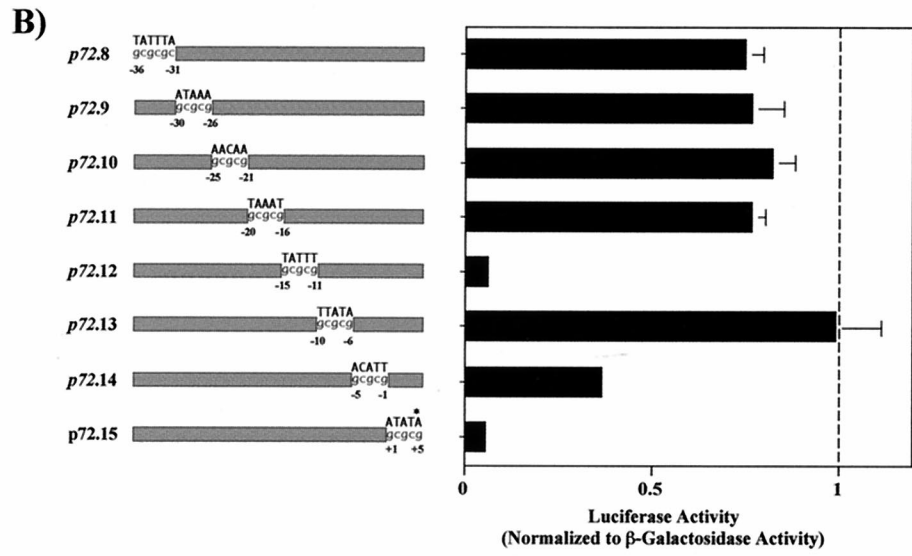
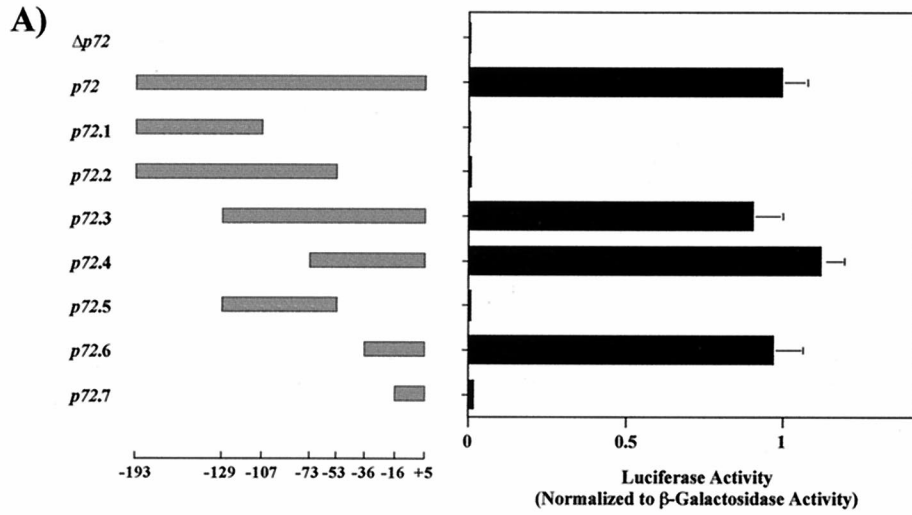
On the other hand, complete loss of reporter activity occurred when the length of the promoter fragment in the transfected plasmid was reduced from 41 to 21 bp (*p72.7*) (Fig. 2A), suggesting that ASFV sequences between  $-36$  and  $-16$  are critical for activity. However, the possibility that transcription of the reporter gene was inhibited by bringing vector sequences closer to the mRNA start site cannot be ruled out at present.

**Linker scan substitutions along the *p72* promoter.** To localize the critical residues within the 41-bp promoter, linker scan substitutions were performed. A set of eight mutants (*p72.8* to *p72.15*) containing GCGCGC or GCGCG substitutions along the 41-bp sequence was constructed as follows. The nontem-

plate strands of the promoters were synthesized as oligonucleotides which contained *Kpn*I and *Bam*HI sites at their 5' and 3' ends, respectively (Table 1, oligonucleotides 5'-8 to 5'-15). In addition, a tail of seven C residues was included at their 3' ends in such a way that the *Bam*HI site and the C tail together would create a 13-bp sequence common to all of them. The single-stranded oligonucleotides were annealed to a 13-bp primer complementary to this 3' end common sequence (Table 1, oligonucleotide 3'-G). Extension of the second strand was performed with DNA polymerase I (Klenow fragment) and all four deoxynucleotides. The double-stranded products were then digested with *Kpn*I and *Bam*HI and cloned into pUC118 linearized with the same endonucleases, and the luciferase gene was inserted downstream of the promoters. The plasmids thus obtained were cotransfected with control plasmid p72GAL10T into infected cells, and the normalized luciferase activity was determined. The results are summarized in Fig. 2B. Two regions, from  $-15$  to  $-11$  (*p72.12*) and from  $+1$  to  $+5$  (*p72.15*), in which substitution of the original sequence for a GCGCG sequence greatly decreased luciferase levels were found. The activity of these mutants was reduced to  $<6\%$  of the wild-type value. On the other hand, substitutions from  $-5$  to  $-1$  (*p72.4*) significantly inhibited reporter activity (by approximately 60%), while mutations in the region from  $-36$  to  $-16$  reduced this activity to a much lesser extent. The deletion of this latter region, however, considerably diminished luciferase levels (Fig. 2A, *p72.7* construct), suggesting that sequences with a length of more than five or six bases are involved in the promoter function of this region.

**Single-nucleotide substitutions at the mRNA initiation region of the *p72* promoter.** Mutants *p72.14* and *p72.15* have substitutions in residues which are sites for mRNA initiation. To better analyze the effect of changes in these and surrounding nucleotides, we obtained a complete set of all three possible single-nucleotide substitutions in positions from  $-1$  to  $+5$ . Mutant promoters (*p72.16A* to *p72.21T*) were generated using mutually priming long oligonucleotides (8). Oligonucleotide 5'-D, which contains the 23-bp 5' sequence of the promoter nontemplate strand, was independently mixed with oligonucleotides 3'-16 to 3'-21, which contain the 33-bp 5' sequence of the promoter template strand mutated at different residues (Table 1). These pairs of oligonucleotides annealed at a 15-bp segment at their 3' ends and acted as both template and primer. The extension of the oligonucleotides was performed with Sequenase (Pharmacia Biotech) in the presence of all four deoxynucleotides. The double-stranded DNA mutant promoters obtained in this way had a structure similar to that of linker scan mutants, with *Kpn*I and *Bam*HI sites at their ends, allowing their cloning into pUC118 plasmid. Plasmids containing point mutations were transfected together with control plasmid p72GAL10T, and the luciferase activity was measured and normalized to  $\beta$ -galactosidase activity. The results show that position  $+3$  is notable for its intolerance to any deviation from the wild-type sequence (Fig. 2C). Changes at nucleotides  $+4$

FIG. 2. (A) Mapping of the *p72* late promoter. Luciferase activity was measured in ASFV-infected Vero cells previously transfected with plasmids containing deletion fragments of the *p72* promoter and was normalized to  $\beta$ -galactosidase activity produced from cotransfected control plasmid p72GAL10T. Transfections were carried out as described in the legend to Fig. 1 by using 12 ng of luciferase plasmids and 1  $\mu$ g of p72GAL10T plasmid per  $1.5 \times 10^5$  cells. Cells were infected with the BA71V strain of ASFV at a multiplicity of infection of 5 PFU per cell, and they were washed and harvested at 18 h postinfection. Luciferase and  $\beta$ -galactosidase activities were determined using the Dual-Light kit from Tropix Inc. The data are represented as the average plus the standard deviation of three assays and are normalized to *p72* activity. The positions of the deletion fragment ends with respect to the transcriptional start point ( $+1$ ) are indicated below the diagram.  $\Delta p72$  corresponds to the luciferase gene without the *p72* promoter. (B) Linker scanning analysis of the *p72* promoter. The substituted sequence is indicated above each promoter, and the positions of mutated residues are indicated. Activities were determined as in panel A. All values were normalized to *p72.6* activity, indicated with a vertical dashed line. The asterisk indicates the position of the first nucleotide of the *p72* gene translation start codon. (C) Effect of single-nucleotide substitutions at or near the transcriptional start site. Activities were determined as for panel A. All values were normalized to *p72.6* activity, indicated with a horizontal dashed line. Symbols for base substitutions are shown.



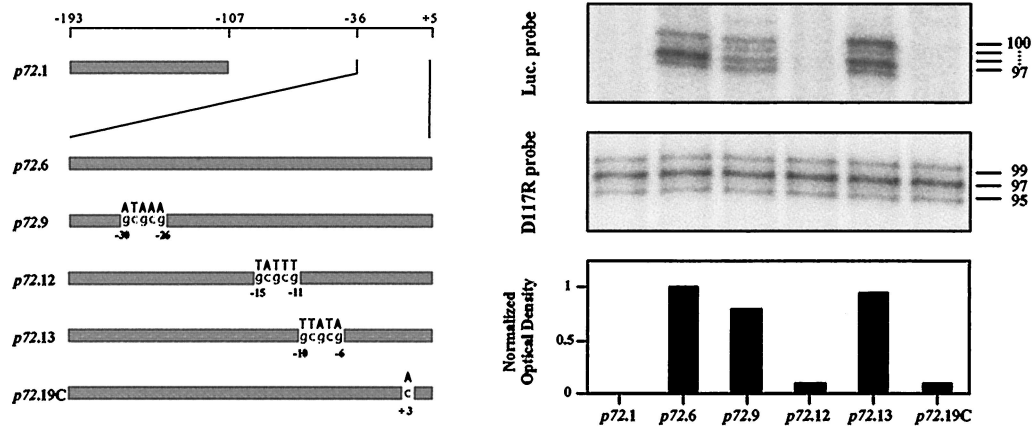


FIG. 3. Analysis of RNA synthesized from *p72* promoter mutants. Preconfluent monolayers of Vero cells ( $1.5 \times 10^5$ ) were transfected with 16  $\mu$ g of Lipofectamine reagent and 1  $\mu$ g of plasmids containing different mutations of the *p72* promoter regulating the luciferase gene. The primer for the luciferase gene was complementary to nucleotides between bp 35 and 64 of the noncoding strand of the gene. The primer for the D117R gene has been previously described (32). Total RNA was obtained at 18 h postinfection, hybridized with end-labeled primer, and extended with avian myeloblastosis virus reverse transcriptase. The sizes (in nucleotides) of the major bands, calculated by using products of an irrelevant DNA sequencing reaction for markers, are indicated. The lower panel shows the quantitated values of the transcripts produced by each mutant normalized first to the internal D117R transcript value and then to the *p72.6* value. A diagram showing the relative promoter lengths and sequences of the promoters used is shown on the left. Mutant *p72.19C* has a point mutation in which the A residue at position +3 was replaced by C.

and +5 reduced the activity of the promoter about threefold, while only a small effect was observed when positions  $-1$  to  $+2$  were replaced.

Compilation of the linker scanning and point mutation data defined two critical regions within the *p72* late promoter: an upstream region located around positions  $-15$  to  $-11$  and a downstream region encompassing the transcription initiation site.

**Effects of *p72* promoter mutations on transcription.** As described above, the relative strengths of the *p72* promoter and promoter mutants were examined by measuring luciferase activity, which provided a convenient and precise method of assessing the promoter activity. The enzymatic activity is assumed to be an accurate reflection of the amount of mRNA synthesized. This was confirmed by direct analysis of the RNA isolated from infected cells that had been transfected with the wild type and representative mutations of the *p72* promoter. As an internal control, transcripts from the D117R gene of the viral genome, which encodes the major structural protein p17 (32), were analyzed in parallel. As probes for this analysis, we used oligonucleotides that contained sequences of the coding strands of the luciferase and D117R genes close to their 5' ends. These primers were end labeled, hybridized with total RNA, and elongated with avian myeloblastosis virus reverse transcriptase as described elsewhere (29). The elongated products were electrophoresed on a polyacrylamide gel. Quantitation of the luciferase transcripts (Fig. 3) demonstrated a good correlation between the RNA synthesized and the activity measured (Fig. 2).

**A TATA sequence is required for efficient late viral transcription.** An analysis of the 5'-flanking sequences of ASFV late genes whose transcription initiation site has been mapped showed that late transcription starts very frequently at A or T residues. A comparison of these sequences showed a conservation around the mRNA start site (Fig. 4). Thus, the sequence TATA of the coding strand is located at or near the start point in 14 out of 17 late genes. No clear regions of similarity were observed in upstream sequences, although tracts of T or A residues were present in most cases.

As determined above, the corresponding TATA sequence in the *p72* promoter is a critical region. To assess if the TATA

sequence acts as a general essential element for initiation of late transcription, we cloned the putative promoter sequences of three other late genes. The 41-bp DNA fragments containing the putative TATA elements and upstream sequences of the late genes encoding the p11.5, p10, and p54 major structural proteins were cloned. A mutant version of each promoter in which the TATA sequence was replaced by GCGC was also cloned. The cloning strategy was similar to that described for single-nucleotide mutagenesis of the *p72* promoter, using the oligonucleotides shown in Table 1. Plasmids containing the promoters and the luciferase gene were cotransfected with *p72GAL10T*, and the normalized luciferase levels were determined. The activity of the wild type versus the TATA mutant promoter for each case is shown in Fig. 5. As can be seen, the percentage of mutant promoter activities ranged from less than 2% (*p11.5* promoter) to about 6% (*p10* and *p54* promoters). These results show that a TATA sequence located at the transcription start point is essential for the activity of four late promoters and suggest that ASFV late transcription is dependent on the presence of this sequence at or near the region of transcriptional initiation.

**A model for regulation of ASFV late gene expression.** Expression of ASFV late genes takes place after the onset of viral DNA replication and is dependent on it. It has been proposed that viral *trans*-acting factors regulating late transcription are expressed from newly synthesized viral genomes (29). These factors may be involved in the transactivation of the luciferase gene under the control of the late viral promoters, which was observed in the infected cells. The luciferase activities found under these conditions are correlative to mRNA levels, as demonstrated by primer extension assays of the accumulated messengers during ASFV infection (Fig. 2 and 3). The results show that the system is very useful for analyzing the structure of *cis*-acting late control elements using mutated promoters.

Two regions essential for *p72* promoter activity have been detected. One region is located at positions  $-15$  to  $-11$ , as replacement of these nucleotides by a GCGCG sequence abolished the promoter activity. Whether this effect is a consequence of the deletion of some essential sequence motif is unknown. No evident nucleotide conservation in that region of other late promoters is observed. On the other hand, A and T

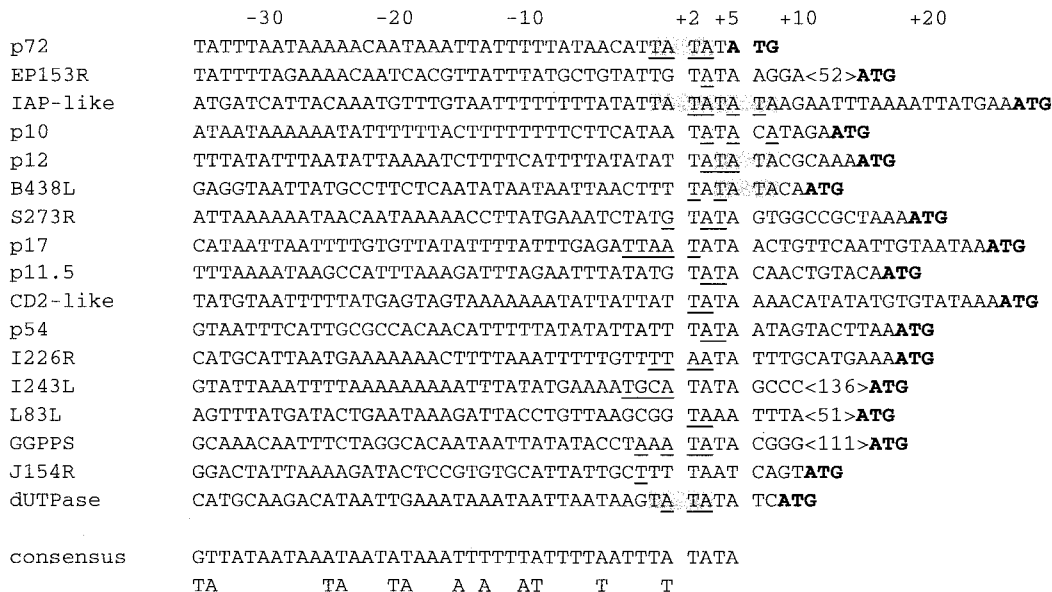


FIG. 4. Alignment of 5' flanking sequences of ASFV late genes. Sequences were aligned by maximizing the identities around the transcriptional start site. Initiation sites (underlined) mapped within or close to a TATA sequence. The translation initiation codons are indicated in bold. References are as follows: p72 (26), p10 (J. M. Rodríguez, unpublished data), EP153R (14), IAP-like (11), p12 (4), B438L (15), S273R (A. Alejo, unpublished data), L83L (1), p17 (32), p11.5 (data not shown), CD2-like (28), p54 (25), I226R and I243L (29), Geranylgeranyl pyrophosphate synthase (GGPPS) (2), J154R (5), and dUTPase (24). The consensus sequence shows the most common nucleotide(s) at each position upstream of residue +5. Also listed is the next most common nucleotide that differs by only one from the most common residue. Alternative TATA sequences are indicated over a gray background.

tracts are frequently found along ASFV late promoters, but the significance of this characteristic in transcription activation remains to be determined. In relation to this, it has been suggested that T-rich sequences upstream from a number of eukaryotic genes activate RNA polymerase II transcription (21). Similarly, tracts of A or T residues activate late promoter function in vaccinia virus (12).

The second essential region of the p72 promoter is localized at positions +1 to +5 and includes sites for mRNA initiation. When the 5' sequences of 17 late genes were aligned by maximizing the identities around the transcription start region, we found some similarity. Thus, TATA sequences are present in

14 out of 17 promoters. As for the p72 promoter, replacement of this sequence in three other late gene promoters was also deleterious for activity, suggesting that the TATA sequence could be a motif for late promoter function. Sequence motifs which map around the site for transcription initiation and are necessary for transcription have been described for other systems. Thus, initiator elements of eukaryotes are common for many promoters of protein-encoding genes and have been grouped into families (33). Similar elements are present in poxvirus and baculovirus promoters. Intermediate and late vaccinia virus transcriptions start in the essential sequences TAAA and TAAAT (23), respectively. In the case of baculovirus, transcription of late and very late genes initiates at a TAAG sequence which is an essential element for both classes of promoters (20). Transcription factors which associate with these initiator elements have been described for eukaryotes (33) and vaccinia virus late promoters (9). It is likely that ASFV late transcription will depend on late transcription factors with affinity for the TATA sequence, an affinity which could be weaker or stronger, depending on the surrounding sequences. On the other hand, the essentiality of the sequence could be related to the mechanism of transcriptional initiation. Thus, an essential step in this mechanism is the melting at the site of initiation, which requires the unwinding of the double helix, a process which is energy dependent. It is possible that the A/T-rich sequence at the transcription initiation site could play a role in facilitating the unwinding of the double-stranded DNA and, consequently, the initiation of transcription. The finding that replacement of the A/T-rich region by G/C residues strongly reduces transcription is in keeping with this.

This is the first mutational analysis of an ASFV promoter. Future work will focus on elucidating the function of the TATA element in ASFV late transcription as well as on identifying the possible transcriptional factors which could interact with it.

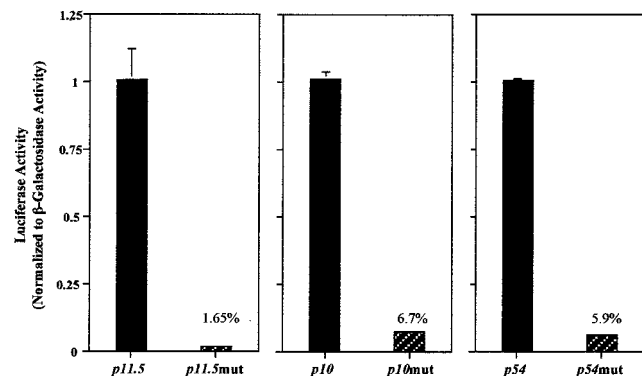


FIG. 5. Requirement of the TATA sequence in the transcription initiation region for activity of other viral late promoters. Plasmids containing wild-type and mutant versions of late promoters directing the expression of the luciferase gene were cotransfected with p72GAL10T as described in the legend of Fig. 2, and luciferase activity was normalized to that of β-galactosidase. All values were normalized to that of wild-type activity in each promoter. The percentages of mutant activities with respect to the wild-type levels are indicated.

We thank J. F. Rodríguez for help in setting up the analysis of promoters in ASFV, M. L. Salas and J. Salas for helpful discussions and critical reading of the manuscript, and M. Mencía and A. Alejo for comments about the manuscript.

This work was supported by Dirección General de Investigación Científica y Técnica grant PB96-0902-C02-01, European Community grant FAIR5-CT97-3441, Comunidad Autónoma de Madrid grant 07B/0032/1997, Ministerio de Educación y Cultura grant AGF98-1352-CE, and an institutional grant from Fundación Ramón Areces.

## REFERENCES

- Alejo, A. 1999. Ph.D. thesis. Universidad Autónoma de Madrid, Madrid, Spain.
- Alejo, A., R. J. Yáñez, J. M. Rodríguez, E. Viñuela, and M. L. Salas. 1997. African swine fever virus trans-prenyltransferase. *J. Biol. Chem.* **272**:9417–9423.
- Almazán, F., J. M. Rodríguez, G. Andrés, R. Pérez, E. Viñuela, and J. F. Rodríguez. 1992. Transcriptional analysis of multigene family 110 of African swine fever virus. *J. Virol.* **66**:6655–6667.
- Almazán, F., J. M. Rodríguez, A. Angulo, E. Viñuela, and J. F. Rodríguez. 1993. Transcriptional mapping of a late gene coding for the p12 attachment protein of African swine fever virus. *J. Virol.* **67**:553–556.
- Almazán, F., J. R. Murguía, J. M. Rodríguez, I. de la Vega, and E. Viñuela. 1995. A set of African swine fever virus tandem repeats shares similarities with SAR-like sequences. *J. Gen. Virol.* **76**:729–740.
- Andrés, G., R. García-Escudero, C. Simón-Mateo, and E. Viñuela. 1998. African swine fever virus is enveloped by a two-membraned collapsed cisterna derived from the endoplasmic reticulum. *J. Virol.* **72**:8988–9001.
- Andrés, G., C. Simón-Mateo, and E. Viñuela. 1997. Assembly of African swine fever virus: role of polyprotein pp220. *J. Virol.* **71**:2331–2341.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.). 1994. *Current protocols in molecular biology*. Greene Publishing Associates/Wiley-Interscience, New York, N.Y.
- Broyles, S. S., X. Liu, M. Zhu, and M. Kremer. 1999. Transcription factor YY1 is a vaccinia virus late promoter activator. *J. Biol. Chem.* **274**:35662–35667.
- Carrascosa, J. L., P. Gonzalez, A. L. Carrascosa, B. García-Barreno, L. Enjuanes, and E. Viñuela. 1986. Localization of structural proteins in African swine fever virus particles by immunoelectron microscopy. *J. Virol.* **58**:377–384.
- Chacón, M. R., F. Almazán, M. L. Nogal, E. Viñuela, and J. F. Rodríguez. 1995. The African swine fever virus IAP homolog is a late structural polypeptide. *Virology* **214**:670–674.
- Davison, A. J., and B. Moss. 1989. Structure of vaccinia virus late promoters. *J. Mol. Biol.* **210**:771–784.
- Enjuanes, L., A. L. Carrascosa, M. A. Moreno, and E. Viñuela. 1976. Titration of African swine fever virus. *J. Gen. Virol.* **32**:471–477.
- Galindo, I., F. Almazán, M. J. Bustos, E. Viñuela, and A. L. Carrascosa. 2000. African swine fever virus EP153R open reading frame encodes a glycoprotein involved in the hemadsorption of infected cells. *Virology* **266**:340–351.
- Galindo, I., E. Viñuela, and A. L. Carrascosa. 2000. Characterization of the African swine fever virus protein p49: a new late structural polypeptide. *J. Gen. Virol.* **81**:59–65.
- García, R., F. Almazán, J. M. Rodríguez, M. Alonso, E. Viñuela, and J. F. Rodríguez. 1995. Vectors for the genetic manipulation of African swine fever virus. *J. Biotechnol.* **40**:121–131.
- García-Beato, R., M. L. Salas, E. Viñuela, and J. Salas. 1992. Role of the host cell nucleus in the replication of African swine fever virus DNA. *Virology* **188**:637–649.
- García-Escudero, R., G. Andrés, F. Almazán, and E. Viñuela. 1998. Inducible gene expression from African swine fever virus recombinants: analysis of the major capsid protein p72. *J. Virol.* **72**:3185–3195.
- Kuznar, J., M. L. Salas, and E. Viñuela. 1980. DNA-dependent RNA polymerase in African swine fever virus. *Virology* **101**:169–175.
- Lu, A., and L. K. Miller. 1997. Regulation of baculovirus late and very late gene expression, p. 193–216. *In* L. K. Miller (ed.), *The baculoviruses*. Plenum, New York, N.Y.
- Lue, N. F., A. R. Buchman, and R. D. Kornberg. 1989. Activation of yeast RNA polymerase transcription by a thymidine-rich upstream element in vitro. *Proc. Natl. Acad. Sci. USA* **86**:486–490.
- Moss, B., B.-Y. Ahn, B. Amegadzie, P. D. Gershon, and J. G. Keck. 1991. Cytoplasmic transcription system encoded by vaccinia virus. *J. Biol. Chem.* **266**:1355–1358.
- Moss, B. 1996. Poxviridae: the viruses and their replication, p. 2637–2671. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), *Fields virology*, 3rd ed. Lippincott-Raven, Philadelphia, Pa.
- Oliveros, M., R. García-Escudero, A. Alejo, E. Viñuela, M. L. Salas, and J. Salas. 1999. African swine fever virus dUTPase is a highly specific enzyme required for efficient replication in swine macrophages. *J. Virol.* **73**:8934–8943.
- Rodríguez, F., V. Ley, P. Gómez-Puertas, R. García, J. F. Rodríguez, and J. M. Escribano. 1996. The structural protein p54 is essential for African swine fever virus viability. *Virus Res.* **40**:161–167.
- Rodríguez, J. M. 1993. Ph.D. thesis. Universidad Autónoma de Madrid, Madrid, Spain.
- Rodríguez, J. M., F. Almazán, E. Viñuela, and J. F. Rodríguez. 1992. Genetic manipulation of African swine fever virus: construction of recombinant viruses expressing the  $\beta$ -galactosidase gene. *Virology* **188**:67–76.
- Rodríguez, J. M., R. J. Yáñez, F. Almazán, E. Viñuela, and J. F. Rodríguez. 1993. African swine fever virus encodes a CD2 homolog responsible for the adhesion of erythrocytes to infected cells. *J. Virol.* **67**:5312–5320.
- Rodríguez, J. M., M. L. Salas, and E. Viñuela. 1996. Intermediate class of mRNAs in African swine fever virus. *J. Virol.* **70**:8584–8589.
- Salas, M. L., J. Kuznar, and E. Viñuela. 1981. Polyadenylation, methylation, and capping of the RNA synthesized in vitro by African swine fever virus. *Virology* **113**:484–491.
- Salas, M. L., J. Rey-Campos, J. M. Almendral, A. Talavera, and E. Viñuela. 1986. Transcription and translation maps of African swine fever virus. *Virology* **152**:228–240.
- Simón-Mateo, C., J. M. P. Freije, G. Andrés, C. López-Otín, and E. Viñuela. 1995. Mapping and sequence of the gene encoding protein p17, a major African swine fever virus structural protein. *Virology* **206**:1140–1144.
- Weis, L., and D. Reinberg. 1992. Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *FASEB J.* **6**:3300–3309.
- Yáñez, R. J., J. M. Rodríguez, M. L. Nogal, L. Yuste, C. Enríquez, J. F. Rodríguez, and E. Viñuela. 1995. Analysis of the complete sequence of African swine fever virus. *Virology* **208**:249–278.