JOURNAL OF VIROLOGY, May 1988, p. 1617–1624
0022-538X/88/051617-08$02.00/0

Vol. 62, No. 5

Localization and Footprinting of an Enhancer within the Avian Sarcoma Virus *gag* Gene

KRISTEN CARLBERG,† THOMAS A. RYDEN, AND KAREN BEEMON*

Biology Department, the Johns Hopkins University, 34th and Charles Streets, Baltimore, Maryland 21218

Received 3 September 1987/Accepted 28 January 1988

A *cis*-acting regulatory element within the *gag* gene of avian retroviruses has been localized by deletion analysis, and sites of protein interaction have been studied by DNase I footprinting. Unidirectional deletions were made from both the 5′ and 3′ ends of a 656-base-pair fragment of the *gag* gene of Fujinami sarcoma virus. These deletion mutants were tested for enhancer activity in a chloramphenicol acetyltransferase transient expression assay. A sharp 5′ boundary for enhancer activity was observed between 776 and 786 nucleotides downstream from the transcription initiation site. In contrast, deletion from the 3′ side resulted in a gradual loss of enhancer activity, reaching a near basal level of activity by nucleotide 868. Internal deletion of 76 nucleotides just downstream of the 5′ boundary abolished enhancement. Mutagenesis of a consensus enhancer core sequence (GTGGTTTG) showed that this sequence was not necessary for enhancer activity in our transient assays. DNase I footprinting with both a highly purified enhancer-binding protein from rat liver (EBP20) and a partially purified chicken liver nuclear extract showed specific protection of nucleotides 813 to 872 within the localized enhancer region. Footprinting of unidirectional deletion mutants that had lost activity indicated that this binding was not sufficient to confer enhancement.

Transcription of the retroviral genome is regulated by enhancer, promoter, and polyadenylation sequences in the long terminal repeats (LTRs) (40). Retroviral LTR enhancers have been identified by their ability to increase the transcription rate of linked heterologous genes independent of their orientation or position (7, 16, 20, 25) and have been localized within the U3 region of Rous sarcoma virus (RSV) LTRs by mutational analysis (2, 17). There is also evidence that viral sequences outside the LTR may help to regulate retroviral gene expression (17, 18, 20, 26, 38).

Arrigo et al. (1) have previously shown that sequences within the highly conserved *gag* genes of several avian retroviruses are capable of enhancing expression of a linked chloramphenicol acetyltransferase (CAT) gene (*cat*) in a transient assay. Insertion of a 656-base-pair (bp) fragment (containing nucleotides 533 to 1189 from the transcription initiation site) of the Fujinami sarcoma virus (FSV) *gag* gene into the enhancerless plasmid pSV1*cat* activates *cat* gene expression more than 20-fold in chicken embryo fibroblasts (CEFs). Similar results were obtained with the corresponding *gag* fragment from RSV and Rous-associated virus-0. The increase in CAT activity is due to increased steady-state levels of *cat* mRNA (1). A smaller FSV *gag* fragment, extending from the *Bam*HI site at nucleotide 533 to the *Nae*I site at nucleotide 1017, retained approximately 60% of the maximum enhancer activity. However, deletion of a central 330-bp *Ava*I fragment (nucleotides 631 to 961) abolished enhancer activity (1).

Although we have not yet established *gag* enhancer activity in the intact viral genome, there is some evidence from viral mutants that the *gag* region contributes to regulation of viral expression (18, 38). Experiments with the virus are complicated by the fact that the enhancing sequences lie within coding sequences of a viral gene. Extensive viral deletions may also affect the stability or splicing of viral RNA. Before minimal enhancer mutations could be constructed in the viral genome, it was necessary to localize the enhancer sequences more precisely. In this study, we defined the boundaries of the *gag* enhancer activity and obtained data on protein-DNA interactions in this region.

Unidirectional deletion mutants were constructed from both the 5′ and 3′ ends of the 656-bp FSV *gag* fragment and tested in a transient expression assay. The 5′ deletion series showed a discrete border with a precipitous drop in activity between nucleotides 776 and 786, whereas the 3′ set showed a gradual decline. Internal deletions and site-specific mutations were made to investigate the role of specific sequence elements within the enhancer region. DNase I footprinting and dimethyl sulfate (DMS) alkylation of the *gag* gene fragment localized sites of protein interactions. A highly purified enhancer-binding protein from rat liver, previously identified as EBP20 (13), and a partially purified chicken liver nuclear extract (CLNE) protected a region near the 5′ boundary of the defined enhancer sequence.

MATERIALS AND METHODS

Plasmid construction. p1*gag*S was constructed by inserting a *Bam*HI-to-*Pvu*I *gag* fragment from a subclone of FSV, pBR-F26 (provided by M. Shibuya and H. Hanufusa [35]), extending from nucleotide 533 to 1189, into a *Bam*HI-*Aat*II-cut pSV1*cat* vector (obtained from B. Howard [8]). After an initial ligation between *Bam*HI-cohesive ends, blunt ends were generated, *Sal*I linkers were added, and the molecules were religated. The resulting insert was downstream and in the positive (sense) orientation relative to the simian virus 40 (SV40) early promoter and the *cat* gene (Fig. 1A).

The p1*gag*S plasmid was used to construct a progressive series of unidirectional 5′ and 3′ deletion mutants, essentially as described previously (23, 29). For the 5′ series, 50 μg of p1*gag*S DNA was linearized with *Bam*HI and incubated with 4.5 U of exonuclease III (New England BioLabs, Inc.) at 23°C in a 150-μl reaction. Samples (7.5 μl) were removed at 30-s intervals and pooled on ice. S1 nuclease (75 U; P-L Biochemicals, Inc.) was added, and the mixture was

* Corresponding author.
† Present address: Fred Hutchinson Cancer Research Center, Seattle, WA 98104.

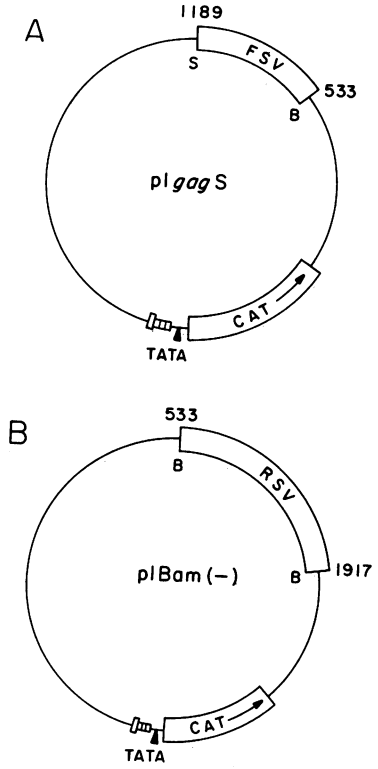

FIG. 1. (A) Structure of plgagS, which was used to make the 5′ and 3′ deletion mutants, as described in Materials and Methods. (B) Structure of plBam(−), which was used to make the internal *gag* gene deletions. Nucleotide endpoints shown for the inserts are from the sequenced genomes of FSV (34) (A) and RSV (33) (B). S, *Sal*I; B, *Bam*HI.

incubated for 15 min at 23°C and then for 15 min on ice. Samples were extracted with phenol-chloroform (1:1), precipitated with ethanol, and ligated to *Bam*HI linkers. After digestion with *Bam*HI and *Sal*I, the *gag* fragments were roughly sized on a 2% agarose minigel, isolated, and ligated to a *Bam*HI-*Sal*I-cut plgagS vector.

The 3′ deletion mutants were constructed similarly, except that plgagS was first linearized with *Sal*I and *Sal*I linkers were added to the deleted ends. Deletion endpoints of the inserts were determined to ±10 bp on a 20-cm 2.5% agarose gel. Endpoints near the activity borders of the enhancer were sequenced by the method of Sanger et al. (30), as modified by Wallace et al. (41), with synthetic oligonucleotide primers complementary to vector sequences adjoining the inserts.

Plasmid plBam(−) (a gift from S. Arrigo), containing a 1,384-bp *Bam*HI *gag* fragment (nucleotides 533 to 1917) from Prague C RSV in the *Bam*HI site of pSV1*cat* in the negative (antisense) orientation (Fig. 1B), was used to construct internal *gag* gene deletions. plBam(−) was linearized with *Nar*I (base 798), deleted with exonuclease III, and digested with S1 nuclease as described above. After *Cla*I linkers were ligated to the deleted ends, inserts were excised with *Cla*I and *Bgl*II, roughly sized by gel electrophoresis, and ligated to a *Nar*I-*Bgl*II-cut plBam(−) vector. Deletion endpoints were determined by dideoxy sequencing (30, 41).

Site-specific mutagenesis. Site-directed mutations (46) were made with M13FSV (provided by G. Weinmaster [43]), which contains a full-length insert of the FSV genome in the antisense orientation at the *Sac*I site of M13. The mutagenic oligonucleotide primer used to generate mutant 1 was a 24-mer (TATGTGGGGAGCCCTTTGTATCCT) in which the three underlined nucleotides were altered from the wild-type sequence TGG. Mutant 2 was derived by extension of another 24-mer (CCTCCTCCTTAT...TTGTATC CTTCC); the wild-type sequence GTGGGGAGTGGT has been deleted from the center of this oligonucleotide. The mutants were sequenced by the dideoxy method with single-stranded M13 DNA (30). Wild-type and mutant M13FSV double-stranded replicative-form DNAs were digested with *Bam*HI and *Aat*II, and the 710-bp *gag* fragments were ligated to *Bam*HI-*Aat*II-cut pSV1*cat*. The resulting plgagA plasmids contained the wild-type or mutant *gag* fragments in the downstream, positive transcriptional orientation.

Cell culture and DNA transfection. CEFs were cultured as previously described (1). Transfection assays used 5 μg of plasmid DNA per 6-cm plate of CEFs and 200 μg of DEAE-dextran per ml as described by Lopata et al. (19). Cells were incubated with DNA for 4 to 6 h and then shocked for 2 min with 10% dimethyl sulfoxide.

CAT assay. Approximately 48 h after DNA addition, CAT activities were assayed as described by Gorman et al. (8), with modifications by Arrigo et al. (1). Cells were treated for 10 to 30 min with pancreatic DNase, scraped from the plate, and lysed by freeze-thawing. After clarification, half (50 μl) of the lysate from each plate was used for the CAT assay and the remaining lysate and pellet were used for DNA quantitation.

DNA quantitation. Pellets were suspended in the remaining lysate and digested for 2 h at 37°C in 200 μl of 0.5-mg/ml pronase (Boehringer Mannheim Biochemicals) and 0.5% sodium dodecyl sulfate. The samples were extracted twice with phenol-chloroform (1:1) and once with chloroform and then were precipitated with ethanol. DNA pellets were digested with *Sal*I and *Eco*RI, electrophoresed on a 1% agarose gel, blotted to nitrocellulose (37), and hybridized to a *Hin*dIII-digested pSV1*cat* DNA probe. Plasmid DNA was quantitated by scintillation counting of the radioactive bands. Plasmid DNA quantitation in some experiments was from Hirt supernatants as described by Arrigo et al. (1) rather than from total cellular DNA as described above.

Nuclear extracts. Rat liver nuclear extract (RLNE) and CLNE were isolated as described by Woll et al. (44), with modifications by Graves et al. (9). Frozen rat livers were obtained from Pel-Freez, Rogers, Ark. Livers from freshly sacrificed chickens (White Rock) were the generous gift of Dover Poultry, Baltimore, Md. Tissue (330 g) yielded approximately 3 ml of nuclear extract at a protein concentration of 3 mg/ml. For heat-treated CLNE, samples were incubated at 68°C for 10 min, chilled on ice, and clarified by centrifugation at 16,000 × *g* for 10 min in a microfuge. The 130-fold-purified enhancer binding protein (EBP20) was the generous gift of P. Johnson, B. Graves, and S. McKnight and has been described as the FPLC Mono S fraction (13).

DNase I and DMS footprinting. DNase I footprinting was performed by the method of Jones et al. (15) as modified by Graves et al. (10). DMS footprints (36) were prepared as described by Johnson et al. (13). Reactions used 1 to 2 ng of DNA, end labeled to a specific activity of $>10^6$ cpm/μg with T4 polynucleotide kinase and 0 to 20 μl of nuclear extract. DNA fragments were sequenced by the method of Maxam and Gilbert (21).

RESULTS

Enhancer activity within avian sarcoma virus *gag* genes was identified previously in a fragment containing nucleo-

tides 533 to 1189 from the start of transcription (1). To determine the boundaries of the *gag* enhancer more precisely, we constructed a 5′ and a 3′ series of unidirectional deletion mutants. The parental plasmid p1*gag*S contained a 656-bp FSV *gag* gene fragment in the *Bam*HI site of pSV1*cat* (Fig. 1A). (pSV1*cat* contains the promoter from SV40, but the enhancer has been deleted [8].) The *gag* insert was downstream and in the sense orientation relative to the promoter and the *cat* gene. As with similar constructs tested previously (1), this 656-bp *gag* fragment elevated CAT activity about 20-fold in CEFs (data not shown).

5′ Deletion mutants. To find the 5′ boundary of the enhancer element, the *gag* gene was progressively deleted from the *Bam*HI site at the 5′ end of the insert (FSV nucleotide 533). Deletion mutants were transfected into CEFs and assayed 48 h later for CAT activity. Parental p1*gag*S and the enhancerless pSV1*cat* were tested in parallel with the deletion mutants, and CAT activity for each sample was normalized to intracellular plasmid DNA levels. The CAT activity of the mutants was not significantly reduced until the *gag* gene was deleted beyond nucleotide 776 (Fig. 2A). Deletion to nucleotide 786 reduced CAT activity to about 20% of that of the undeleted plasmid. Further deletion to nucleotide 802 resulted in activity barely above that of pSV1*cat*.

3′ Deletion mutants. A similar series of deletion mutants was constructed from the 3′ end of the *gag* insert at nucleotide 1189, transfected into CEFs, and assayed for CAT activity (Fig. 2C). In contrast to the sharp boundary observed in the 5′ series, the 3′ deletion series exhibited a more gradual decline in enhancer activity. The first significant drop in activity occurred between deletion endpoints at nucleotides 1136 and 1079, reducing CAT activity to approximately 70% of the activity of full-length constructs. Further deletion to nucleotide 952 reduced activity to about 50%. CAT activity declined more rapidly as the *gag* insert was deleted from nucleotide 952 to nucleotide 868. Deletion beyond nucleotide 868 resulted in no detectable enhancer activity. Thus, the *gag* sequences necessary for full enhancement of CAT activity were localized between nucleotides 776 and 1136 in the FSV genome.

Internal *gag* gene deletions. Unidirectional deletion analysis suggested that sequences in the 5′ portion of the enhancer region might contribute more to enhancer activity than

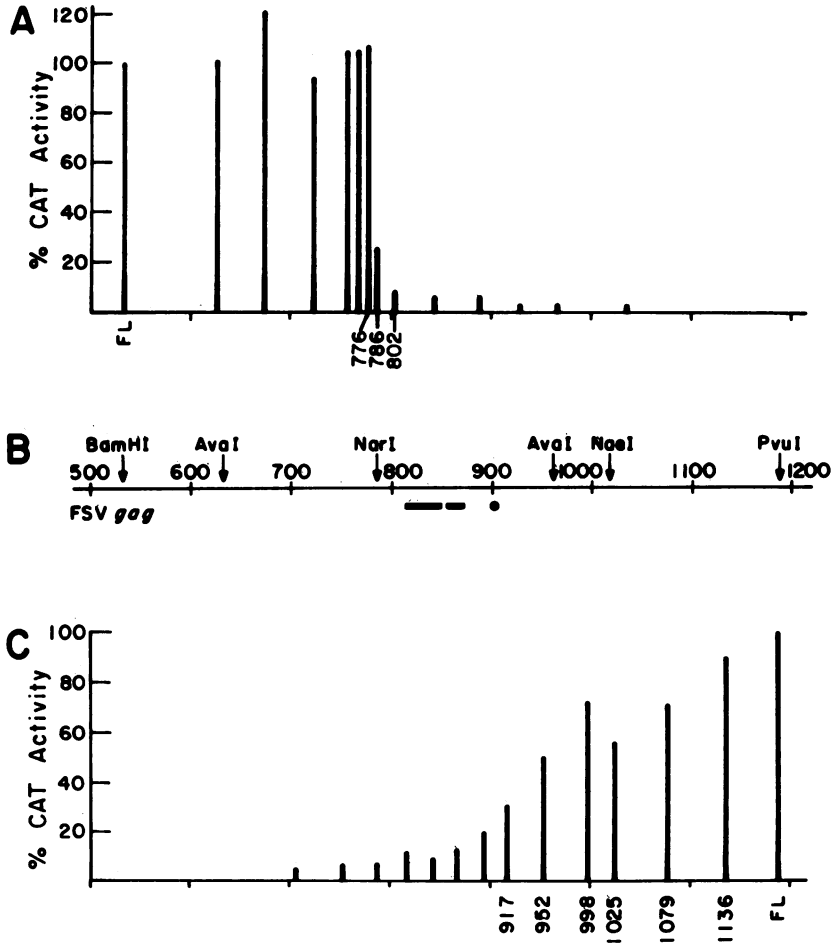

FIG. 2. CAT activity of 5′ and 3′ deletion mutants, shown as the percentage of the activity of p1*gag*S, which has a full-length (FL) insert. The value given for each mutant is the average of two experiments, each of which was normalized to the amount of intracellular plasmid DNA. The enhancerless plasmid pSV1*cat* had 3% of the activity of p1*gag*S. (A) 5′ Deletion mutants. The 5′ endpoints of mutants near the border of the observed enhancer activity are indicated. (B) Map of the *gag* gene of FSV. Numbers refer to the FSV nucleotide sequence (34). Symbols: ●, position of the SV40 consensus enhancer core sequence, GTGGTTTG; ▬, region of the sense strand protected by EBP20 or heat-treated CLNE from DNase I digestion. (C) 3′ Deletion mutants. The 3′ endpoints of sequenced mutants are indicated.

TABLE 1. Internal *gag* gene deletion mutants

Plasmid	Relative CAT activity (%)[a]
p1Bam(+), wild type	260
p1Bam(−), wild type	100
p1Bam Δ 76	5
p1Bam Δ 189	5
pSV1*cat*	7

[a] CAT activity is shown as the percentage of the activity of the negative orientation plasmid p1Bam(−), which was used for deletion mutagenesis.

sequences in the 3′ portion. To examine the enhancer activity of the 3′ sequences alone, we made two internal deletions within the 5′ portion of the enhancer. These were constructed by using plasmid p1Bam(−), which contains a 1,384-bp *Bam*HI *gag* fragment from Prague C RSV in the downstream negative orientation (Fig. 1B). The enhancer activities of FSV and RSV *gag* genes have been shown to be indistinguishable (1). As with previous constructs assayed in mouse L cells (1), the enhancer activity was found to depend on the orientation of the *gag* insert. When tested for CAT activity in CEFs, the positive orientation clone, p1Bam(+), had a 2.6-fold-higher activity than the negative orientation clone, p1Bam(−) (Table 1). The orientation dependence observed in CEFs was less pronounced than that previously seen in mouse L cells (1).

The 5′ boundary of the *gag* enhancer in FSV was very near the *Nar*I site (FSV nucleotide 786). The corresponding *Nar*I site in the RSV genome is at nucleotide 798 (33). p1Bam(−) was progressively deleted from the *Nar*I site in the 5′ to 3′ direction as described in Materials and Methods. Two mutants that had 76 bp (RSV nucleotides 798 to 874) and 189 bp (RSV nucleotides 798 to 987) deleted were selected. When tested in CEFs, both mutants showed a complete loss of enhancer activity (Table 1). Therefore, we have found sequences essential for enhancer function within the 76 bp just downstream of the *Nar*I site in *gag*.

Mutagenesis of consensus enhancer core sequence. Results with the internal deletion mutants showed that sequences in the 3′ portion of the enhancer region alone were not sufficient for enhancer activity. The gradual loss of activity in the 3′ deletion series suggested, however, that additional sequence elements might contribute to the enhancer activity. The enhancer region defined by unidirectional deletion mutagenesis contains several of the sequence motifs commonly found in transcriptional regulatory elements, including an SV40 consensus enhancer core sequence GTGGTTTG (nucleotides 900 to 907) (42).

To test the significance of the enhancer core sequence within the *gag* enhancer region, this sequence was altered by site-directed mutagenesis. The DNA template for mutagenesis was an M13 clone containing the full-length FSV genome (43). Two different oligonucleotides complementary to the antisense FSV insert were synthesized and used for mutagenesis. The sequences of the two resulting mutants are shown in Table 2. The three nucleotides altered in FSV *gag* mutant 1 were selected because mutation of the corresponding nucleotides in the SV40 enhancer caused a dramatic reduction in both viral growth and enhancer function (11, 45).

Mutant 2 was designed to have a 12-nucleotide deletion, including most of the consensus enhancer core sequence and a related GTGG sequence upstream. However, sequence analysis showed it had spontaneously deleted an additional six nucleotides consisting of two of four tandemly repeated CCT triplets present just upstream of the consensus enhancer core sequence (Table 2).

The 710-bp *Bam*HI-*Aat*II *gag* fragment from each mutant was inserted into the pSV1*cat* vector, and the resulting plasmids, p1*gag*A mut 1 and p1*gag*A mut 2, were transfected into CEFs and assayed for CAT activity. Neither mutation reduced CAT activity in our transient assays (Table 2). Thus, the consensus core sequence was not necessary for *gag* enhancer activity.

DNA-protein interactions. Enhancer-mediated activation of transcription appears to require binding of *trans*-acting cellular proteins to the enhancer DNA (4, 24, 31, 32). We have therefore used DNase I footprinting (5) and DMS alkylation (36) to look for sequence-specific binding of nuclear proteins to the *gag* enhancer. When crude RLNE was used for the initial DNase I footprints, several large protected regions, extending beyond the functional boundaries defined by deletion mutagenesis (data not shown), were observed. We therefore turned to the use of purified proteins.

Johnson et al. (13) have recently purified from RLNE an enhancer binding protein (EBP20) which binds to the consensus enhancer core sequences (GTGG$^{TTT}_{AAA}$G) of SV40, polyomavirus, and the Moloney murine sarcoma virus LTR. We were interested in whether EBP20 would also bind to the consensus enhancer core in the avian sarcoma virus *gag* gene. EBP20, purified 130-fold from RLNE, was used for DNase I footprinting of the *gag* enhancer; 180 ng of this extract protected two adjacent domains on the coding strand which encompassed about 60 nucleotides within the 5′ half of the *gag* enhancer region. This binding affinity was similar to that of this extract for the consensus enhancer core sequence of polyomavirus (Peter Johnson, personal communication). The tandem footprints extended from nucleotides 813 to 848 and 855 to 872 (Fig. 3A). The protected region overlapped the enhancer domain, whose importance was established by an internal 76-bp deletion (Fig. 4). Surprisingly, the consensus enhancer core sequence at nucleotides 900 to 907 was not protected by this extract.

To determine whether a factor analogous to the rat EBP20 exists in the natural host of the retrovirus, we prepared CLNE. As was the case with crude RLNE, binding of crude

TABLE 2. Sequences and activities of enhancer core mutants

gag enhancer	Core sequence[a] from nucleotide 881 to 910	Relative CAT activity (%)[b]
Wild type FSV	CCTCCTCCTTATGTGGGGAGTGGTTTGTAT	100
Mutant 1	CCTCCTCCTTATGTGGGGAGCCCTTTGTAT	103
Mutant 2	CCT.....TAT.....TTGTAT	106

[a] The enhancer core consensus sequence is in boldface, the three nucleotide changes in mutant 1 are underlined and the deleted regions in mutant 2 are indicated by dots.

[b] CAT activity of the mutants is shown as a percentage of the activity of the wild-type plasmid. Values are the averages of three experiments, each normalized to levels of intracellular plasmid DNA.

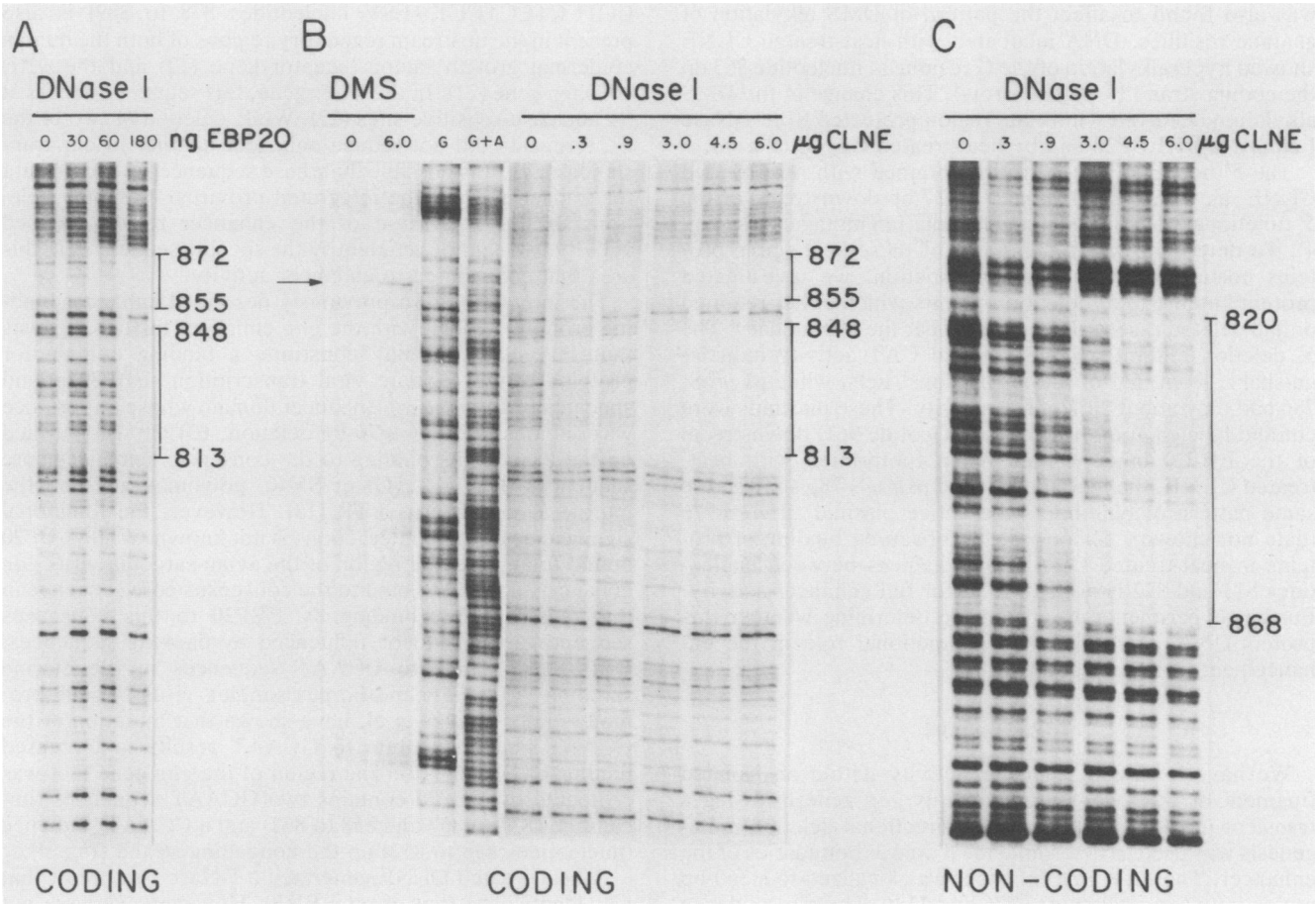

FIG. 3. DNase I footprints of the *gag* enhancer region. (A) Footprint of the coding strand with purified EBP20 from rat liver nuclei. A 5′ deletion mutant (5′ endpoint at nucleotide 636) was digested with *Bam*HI and end labeled with polynucleotide kinase. After digestion with *Sal*I, a 575-bp fragment was gel purified. Footprint reactions used 0, 20, 60, or 180 ng of the EBP20-containing Mono S fraction (13). (B) DNase I footprinting and DMS methylation studies of the coding strand of the *gag* enhancer with heat-treated CLNE. The same 5′ deletion mutant used for panel A was used as a DNA probe. The DMS methylation reactions used 0 or 6.0 μg of extract. The arrow denotes the hypersensitive G residue at nucleotide 863. The DNase I footprinting reactions contained 0, 0.3, 0.9, 3.0, 4.5, or 6.0 μg of extract. G and A+G sequence ladders of the end-labeled DNA fragment were electrophoresed in parallel. (C) DNase I footprinting, with heat-treated CLNE, of the noncoding strand of the *gag* enhancer. A 5′ deletion mutant (endpoint 776) was digested and labeled at the *Ava*I site (nucleotide 961) and then incubated with 0, 0.3, 0.9, 3.0, 4.5, or 6.0 μg of heat-treated CLNE. After DNase I digestion, products were run on 8% sequencing gels. At the right of the footprints are diagrams of the protected regions, with nucleotide numbers for endpoints.

CLNE was observed over a large region of the *gag* gene. Since EBP20 is unusually heat stable and can be purified 10- to 20-fold by heat treatment (13), we subjected the crude CLNE to a similar heat treatment. The heated CLNE protected the same region of the coding strand of the *gag* enhancer from DNase I digestion as did the highly purified EBP20 preparation (Fig. 3B). On the noncoding strand, the heat-treated CLNE protected a region with boundaries at nucleotides 820 and 868 (Fig. 3C). Therefore, we have observed proteins derived from both RLNE and CLNE which are heat stable and which bind to the same sequence within a discrete region of the *gag* gene that is essential for enhancer activity.

Binding of the heat-treated CLNE to the enhancer DNA

FIG. 4. Summary of binding sites within the central region of the *gag* enhancer. The binding of rat EBP20 and heat-treated CLNE to the coding strand of FSV DNA is shown by the line above the sequence, while binding of heat-treated CLNE to the noncoding strand is denoted by the line below the sequence. Arrows designate GCAAT and CCAAT boxes. The hypermethylated G residue at nucleotide 863 is marked by an asterisk. The sharp drop in enhancer activity in the 5′ deletion series occurred at endpoints between nucleotide 776 and 786; 3′ deletion beyond nucleotide 868 also resulted in no enhancement. The location of a 76-bp internal deletion which eliminated enhancer activity is shown by brackets.

was also found to affect the pattern of DMS alkylation of guanine residues. DNA incubated with heat-treated CLNE showed hyperalkylation of the G residue at nucleotide 863 on the coding strand (Fig. 3B, arrow). This change in the DMS alkylation occurred within the region protected from DNase I cleavage by EBP20 and by heat-treated CLNE (Fig. 4).

The 5′ border of the footprint obtained with heat-treated CLNE, as with EBP20, was 27 to 37 bp downstream of the 5′ functional boundary defined by deletion mutagenesis (Fig. 4). To determine whether binding of these heat-stable proteins correlated with enhancer function, we investigated protein binding to 5′ deletion mutants which either retained or had deleted essential sequences near the 5′ boundary. The 5′ deletion mutant pl*gag*S-776 had CAT activity indistinguishable from that of full-length constructs, while pl*gag*S-786 had only about 20% of full activity. These plasmids were cut and labeled at an *Ava*I site (nucleotide 961) downstream of the EBP20 binding site. When footprinted with heat-treated CLNE, the inactive plasmid pl*gag*S-786 showed the same pattern of binding as the active plasmid pl*gag*S-776 (data not shown). Therefore, the observed binding of proteins in heat-treated CLNE to sequences between nucleotides 813 and 872 is not sufficient for full enhancer activity. Further experiments are needed to determine whether this protein-DNA interaction has a functional role in the enhancement of transcription.

DISCUSSION

We have localized enhancer activity within an isolated fragment of the avian sarcoma virus *gag* gene by using a transient CAT assay system. Unidirectional deletion mutagenesis was used to determine the 5′ and 3′ boundaries of the enhancer. The FSV *gag* enhancer was localized to a 360-bp region between nucleotides 776 and 1136. There is evidence that several domains within this region are required for full enhancer activity. The importance of sequences upstream of the *Nar*I site at nucleotide 786 was demonstrated by the 5′ deletion series. A CAT plasmid bearing a deletion from the 5′ end of the *gag* insert to nucleotide 776 was fully active, whereas one deleted to nucleotide 786 showed an 80% reduction in enhancer activity.

Sequences just downstream of the *Nar*I site were also shown to be critical for *gag* enhancer activity. A purified enhancer binding protein, EBP20, was shown to bind to this region of the *gag* gene. An internal deletion of 76 bp just 3′ of the *Nar*I site eliminated all enhancement. It was shown previously that insertion of linkers at the *Nar*I site does not reduce enhancer activity (1). This suggests that at least two domains—one upstream and one downstream of the *Nar*I site—are required for full enhancement.

A 3′ deletion extending to approximately nucleotide 870 also abolished enhancer activity, thus defining a third critical domain, which is downstream of the EBP20-binding site. The gradual loss of activity seen with the 3′ deletion series may be evidence of multiple enhancer elements acting in concert with upstream sequences. Similar gradual reductions have been seen in the murine sarcoma virus LTR enhancer (9) and the immunoglobulin κ enhancer (27). There are a number of previously described enhancer and promoter motifs downstream of the critical 5′ sequences, including Sp1-binding sequences (3) and nuclear-factor-1-binding sequences (14). Mutagenesis of the SV40 consensus enhancer core sequence (41) (nucleotides 900 to 907) demonstrated that it was not necessary for *gag* enhancer function as measured in transient CAT assays. The repeated sequence CCTCCTCCTCCT (FSV nucleotides 878 to 889) is also present in the upstream regulatory regions of both the human epidermal growth factor receptor gene (12) and the α2(I) collagen gene (22). In the latter gene, this sequence occurs at S1 nuclease-sensitive sites (22). While deletion of two of the CCT repeats did not reduce enhancer activity in our transient assays, it is possible that these sequences are important for transcription of the integrated provirus. Although deletions of the 3′ portion of the enhancer region reduced activity, we cannot yet identify the specific sequences in this area that contributed to enhancer activity.

The interaction of a previously described enhancer binding protein, EBP20, with the *gag* enhancer further suggests that this sequence may constitute a binding domain for proteins which regulate viral transcription. EBP20 bound specifically within a *gag* enhancer domain whose importance was demonstrated by a 76-bp deletion. EBP20 was purified on the basis of its binding to the consensus enhancer core sequences ($GTGG^{TTT}_{AAA}G$) of SV40, polyomavirus, and the murine sarcoma virus LTR (13). However, the functional significance of these interactions is not known. While EBP20 bound to an essential region of the avian sarcoma virus *gag* enhancer, it failed to bind to the consensus core sequence in this region. Thus, binding by EBP20 to this consensus sequence appears to be influenced by flanking sequences. EBP20 also binds to CCAAT sequences in the murine sarcoma virus LTR and herpes simplex virus *tk* gene promoters (10). Graves et al. have shown that mutation of the CCAAT pentanucleotide to GCAAT results in increased binding of EBP20 (10). The region of the *gag* gene that was protected by EBP20 contains two GCAAT sequences (nucleotides 833 to 837 and 858 to 862) and a CCAAT sequence (nucleotides 848 to 852) on the noncoding strand (Fig. 4).

A heat-treated CLNE generated a DNase I footprint that was identical to that of rat EBP20. However, we have not yet determined to what extent these two heat-stable proteins are related. Although one or more heat-stable proteins from CLNE bound to an essential region of the *gag* enhancer, this binding is apparently not sufficient for full enhancer activity. We found no difference in binding affinity to two constructs that had either full or diminished enhancement due to 5′ deletions. Since multiple essential domains have been observed in the *gag* enhancer region, it is likely that transcriptional activation requires binding of additional proteins. We have observed binding of EBP20 to the RSV and FSV LTR enhancers as well as to the *gag* regions of both viruses (T. A. Ryden and K. Beemon, manuscript in preparation). This suggests that the *gag* and LTR enhancers interact with the same transcriptional factors.

Because these experiments were performed with the *gag* fragment outside its natural context, we cannot conclude that the *gag* region acts as an enhancer in the intact virus. However, there are some suggestions from deletion mutants that the *gag* region helps to regulate viral expression. Levantis et al. have shown that transformation of rat cells by a *src* gene fragment is increased 20- to 30-fold by adjacent *gag* sequences in the absence of an LTR (18). Stoltzfus et al. showed that retention of the *gag* enhancer region in deleted Prague A RSV clones is associated with increased transformation efficiency in CEF and NIH 3T3 cells (38). In both of these studies, the viral genome was extensively deleted, so the *gag*-induced changes may not reflect normal viral controls. Further, these increases in transformation efficiency have not yet been shown to be due to increased transcription. In most bursal lymphomas, induced by avian leukosis virus (6, 28) or reticuloendotheliosis virus (39), the *gag*

enhancer region has been deleted from the integrated viral genome. Surprisingly, these deleted proviruses, which have integrated upstream of c-*myc*, give rise to stable transcripts initiated at the 3′ LTRs but not at intact 5′ LTRs (6, 39). This suggests that the *gag* enhancer may function to focus the cell transcriptional apparatus on the 5′ LTR in preference to the 3′ LTR. Both the proximity of the *gag* enhancer to the 5′ LTR and its orientation dependence support this possibility. Experiments with appropriate viral mutants will be necessary to determine what role the *gag* enhancer plays in the life cycle of the virus.

ACKNOWLEDGMENTS

We thank S. Arrigo for helpful discussions and T. Csepany and T. C. Lee for help with experiments. We also thank P. Johnson, B. Graves, and S. McKnight for generously providing EBP20 and for helpful discussions; B. Howard for *cat* plasmids; M. Shibuya and H. Hanafusa for FSV subclones; and G. Weinmaster and T. Pawson for the M13FSV clone.

This work was supported by Public Health Service grant CA33199 from the National Cancer Institute and by Biomedical Research Support grant S07 RR07041. K. B. is a recipient of a Faculty Research Award from the American Cancer Society.

LITERATURE CITED

1. **Arrigo, S., M. Yun, and K. Beemon.** 1987. *cis*-Acting regulatory elements within *gag* genes of avian retroviruses. Mol. Cell. Biol. **7**:388–397.
2. **Cullen, B. R., K. Raymond, and G. Ju.** 1985. Transcriptional activity of avian retroviral long terminal repeats directly correlates with enhancer activity. J. Virol. **53**:515–521.
3. **Dynan, W. S., and R. Tjian.** 1983. The promoter-specific transcription factor Sp1 binds to upstream sequences of the SV40 early promoter. Cell **35**:79–87.
4. **Ephrussi, A., G. M. Church, S. Tonegawa, and W. Gilbert.** 1985. B lineage-specific interactions of immunoglobulin enhancer with cellular factors *in vivo*. Science **227**:134–140.
5. **Galas, D., and A. Schmitz.** 1978. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. **5**:3157–3170.
6. **Goodenow, M. M., and W. S. Hayward.** 1987. 5′ Long terminal repeats of *myc*-associated proviruses appear structurally intact but are functionally impaired in tumors induced by avian leukosis viruses. J. Virol. **61**:2489–2498.
7. **Gorman, C. M., G. T. Merlino, M. C. Willingham, I. Pastan, and B. H. Howard.** 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA **79**:6777–6781.
8. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
9. **Graves, B. J., R. N. Eisenman, and S. L. McKnight.** 1985. Delineation of transcriptional control signals within the Moloney murine sarcoma virus long terminal repeat. Mol. Cell. Biol. **5**:1948–1958.
10. **Graves, B. J., P. F. Johnson, and S. L. McKnight.** 1986. Homologous recognition of a promoter domain common to the MSV LTR and HSV *tk* gene. Cell **44**:565–576.
11. **Herr, W., and J. Clarke.** 1986. The SV40 enhancer is composed of multiple functional elements that can compensate for one another. Cell **45**:461–470.
12. **Ishii, S., Y. H. Xu, R. H. Stratton, B. A. Roe, G. T. Merlino, and I. Pastan.** 1985. Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. Proc. Natl. Acad. Sci. USA **82**:4920–4924.
13. **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.
14. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48**:79–89.
15. **Jones, K. A., K. B. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell **42**:559–572.
16. **Laimins, L. A., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA **79**:6453–6457.
17. **Laimins, L. A., P. Tsichlis, and G. Khoury.** 1984. Multiple enhancer domains in the 3′ terminus of the Prague strain of Rous sarcoma virus. Nucleic Acids Res. **12**:6427–6442.
18. **Levantis, P., D. A. F. Gillespie, K. Hart, M. J. Bissell, and J. A. Wyke.** 1986. Control of expression of an integrated Rous sarcoma provirus in rat cells: role of 5′ genomic duplications reveals unexpected patterns of gene transcription and its regulation. J. Virol. **57**:907–916.
19. **Lopata, M. A., D. W. Cleveland, and B. Sollner-Webb.** 1984. High level transient expression of a chloramphenicol acetyltransferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethylsulfoxide or glycerol shock treatment. Nucleic Acids Res. **12**:5707–5717.
20. **Luciw, P. A., J. M. Bishop, H. E. Varmus, and M. R. Capecchi.** 1983. Location and function of retroviral and SV40 sequences that enhance biochemical transformation after micro-injection of DNA. Cell **33**:705–716.
21. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
22. **McKeon, C., A. Schmidt, and B. de Crombrugghe.** 1984. A sequence conserved in both the chicken and mouse α2(I) collagen promoter contains sites sensitive to S1 nuclease. J. Biol. Chem. **259**:6636–6640.
23. **McKnight, S. L., and E. R. Gavis.** 1980. Expression of the herpes thymidine kinase gene in *Xenopus laevis* oocytes: an assay for the study of deletion mutants constructed *in vitro*. Nucleic Acids Res. **8**:5931–5948.
24. **McKnight, S. L., and R. Tjian.** 1986. Transcriptional selectivity of viral genes in mammalian cells. Cell **46**:795–805.
25. **Oskarsson, M., W. L. McClements, D. G. Blair, J. V. Maizel, and G. F. Vande Woude.** 1980. Properties of a normal mouse cell DNA sequence (sarc) homologous to the *src* sequence of Moloney sarcoma virus. Science **207**:1222–1224.
26. **Payvar, F., D. DeFranco, G. L. Firestone, B. Edgar, O. Wrange, S. Okret, J. A. Gustafsson, and K. R. Yamamoto.** 1983. Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. Cell **35**:381–392.
27. **Queen, C., and J. Stafford.** 1984. Fine mapping of an immunoglobulin activator. Mol. Cell. Biol. **4**:1042–1049.
28. **Robinson, H. L., and G. C. Gagnon.** 1986. Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. J. Virol. **57**:28–36.
29. **Sakonju, S., D. F. Bogenhagen, and D. D. Brown.** 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription. I. The 5′ border of the region. Cell **19**:13–25.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
31. **Sassone-Corsi, P., and E. Borrelli.** 1986. Transcriptional regulation by trans-acting factors. Trends Genet. **2**:215–219.
32. **Schöler, H. R., and P. Gruss.** 1984. Specific interactions between enhancer-containing molecules and cellular components. Cell **36**:403–411.
33. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
34. **Shibuya, M., and H. Hanafusa.** 1982. Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its transforming gene with transforming genes of other sarcoma viruses. Cell **30**:787–795.
35. **Shibuya, M., L.-H. Wang, and H. Hanafusa.** 1982. Molecular cloning of the Fujinami sarcoma virus genome and its compar-

ison with sequences of other related transforming viruses. J. Virol. **42**:1007–1016.
36. **Siebenlist, V., R. B. Simpson, and W. Gilbert.** 1980. E. coli RNA polymerase interacts homologously with two different promoters. Cell **20**:269–281.
37. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
38. **Stoltzfus, C. M., L.-J. Chang, T. P. Cripe, and L. P. Turek.** 1987. Efficient transformation by Prague A Rous sarcoma virus plasmid DNA requires the presence of *cis*-acting regions within the *gag* gene. J. Virol. **61**:3401–3409.
39. **Swift, R. A., C. Boerkoel, A. Ridgway, D. J. Fujita, J. B. Dodgson, and H.-J. Kung.** 1987. B-lymphoma induction by reticuloendotheliosis virus: characterization of a mutated chicken syncytial virus provirus involved in c-*myc* activation. J. Virol. **61**:2084–2090.
40. **Varmus, H., and R. Swanstrom.** 1985. Replication of retroviruses, p. 75–134. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses: molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Wallace, R. B., M. J. Johnson, S. V. Suggs, K. Miyoshi, R. Bhatt, and K. Itakura.** 1981. A set of synthetic oligodeoxyribonucleotide primers for DNA sequencing in the plasmid vector pBR322. Gene **16**:21–26.
42. **Weiher, H., M. Konig, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science **219**: 626–631.
43. **Weinmaster, G., M. J. Zoller, M. Smith, E. Hinze, and T. Pawson.** 1984. Mutagenesis of Fujinami sarcoma virus: evidence that tyrosine phosphorylation of $P130^{gag\text{-}fps}$ modulates its biological activity. Cell **37**:559–568.
44. **Woll, W. W., J. J. Duffy, N. A. Giese, and T. J. Lindell.** 1981. Nuclear isolation by a modified method of Hewish and Burgoyne: implications for the study of nuclear enzymology. Life Sci. **29**:2709–2719.
45. **Zenke, M., T. Grundstrom, H. Matthes, M. Wintzerith, C. Schatz, A. Wildeman, and P. Chambon.** 1986. Multiple sequence motifs are involved in SV40 enhancer function. EMBO J. **5**:387–397.
46. **Zoller, M. J., and M. Smith.** 1984. Oligonucleotide-directed mutagenesis: simple method using two oligonucleotide primers and a single-stranded DNA template. DNA **3**:479–488.