MOLECULAR AND CELLULAR BIOLOGY, June 1985, p. 1498–1151
0270-7306/85/061498-14$02.00/0

Vol. 5, No. 6

Linker Scanning Mutagenesis of the 5′-Flanking Region of the Mouse β-Major-Globin Gene: Sequence Requirements for Transcription in Erythroid and Nonerythroid Cells

PATRICK CHARNAY,† PAMELA MELLON,‡ AND TOM MANIATIS*

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

Received 28 December 1984/Accepted 19 February 1985

We analyzed the sequences required for transcription of the mouse β-major-globin gene by introducing deletion and linker scanning mutations into the 5′-flanking region and then studying the effects of these mutations on β-globin gene transcription in a HeLa cell transient expression assay or after stable introduction into mouse erythroleukemia cells. Consistent with earlier studies, we found that three distinct regions upstream from the RNA capping site are required for efficient β-globin gene transcription in HeLa cells: the ATA box located 30 base pairs upstream from the mRNA capping site (−30), the CCAAT box located at −75, and the distal sequence element CCACACCC located at −90. In the ATA and CAAT box regions, the sequences necessary for efficient transcription extend beyond the limits of the canonical sequences. Mutations in the sequences located between the three transcriptional control elements do not significantly affect transcription in HeLa cells. Although the promoter defined in HeLa cell transfection experiments is also required for efficient transcription in mouse erythroleukemia cells, none of the mutations tested affects the regulation of β-globin gene transcription during mouse erythroleukemia cell differentiation. Thus, DNA sequences downstream from the mRNA cap site appear to be sufficient for the regulation of β-globin gene expression during the differentiation of mouse erythroleukemia cells in culture.

A prerequisite for understanding the detailed mechanisms of β-globin gene regulation during erythroid cell differentiation is the identification of the *cis*-acting DNA sequences required for accurate and efficient initiation of transcription. Comparison of the DNA sequences immediately upstream from the mRNA capping site of a number of mammalian β-globin genes revealed three highly conserved regions (17, 28, 32). The first region is the ATA box, which is located approximately 30 base pairs (bp) upstream from the mRNA cap site. This sequence, which was initially noted in *Drosophila* histone genes (M. Goldberg, Ph.D. thesis, Stanford University, Stanford, Calif., 1979), is found at the same relative position in most eucaryotic genes transcribed by RNA polymerase II (7, 11). The second region is the CCAAT box, which is located approximately 80 bp upstream from the mRNA cap site (5, 17). This sequence is also found in a similar position near many other eucaryotic genes (5). A third region, which is conserved among adult mammalian globin genes that are β-like but is not a common feature of other eucaryotic promoters, is the sequence CCA/TCACCCT located approximately 90 or 105 bp upstream from the mRNA cap site (14, 32).

The functional significance of these highly conserved sequence elements has been assayed both in vitro (27) and in vivo (14, 15, 25, 26) in the case of the rabbit β-globin gene. As with many other eucaryotic promoters (7), only the ATA box is required for in vitro transcription (27). Deletion or mutation of the CCAAT or CACCC box sequences has no effect on transcription in whole-cell extracts (27). However, these upstream sequences are required for efficient transcription of the rabbit β-globin gene transiently transfected into cells in culture (14, 25, 26). Deletion or point mutations in any of the three conserved sequences result in a decrease in the level of transcription of β-globin genes, so the conserved sequences appear to play an important role in transcription in vivo. In this paper we extend this analysis by examining the effect of linker scanning (LS) (42) mutations on β-globin gene transcription in HeLa cells as a means of determining whether sequences which lie outside of the conserved transcription element are necessary for normal levels of transcription.

Studies have also been carried out to identify *cis*-acting DNA sequences required for globin gene regulation during erythroid cell differentiation (8, 10, 60, 61). In these studies, cloned globin genes are stably introduced into mouse erythroleukemia (MEL) cells which are derived from adult hematopoietic precursor cells transformed with the Friend virus complex (19; see reference 36 for review). Treatment of these cells with dimethyl sulfoxide (DMSO) or a variety of other chemical inducers leads to a series of morphological and biochemical modifications that mimic events in normal erythroid cell differentiation (19, 36), including a dramatic increase in the rate of transcription of adult globin genes (8, 29, 49). When cloned mouse or human β-globin genes are introduced into MEL cells by DNA-mediated gene transfer, they are coregulated with the endogenous mouse globin genes (8, 60). These results indicate that the DNA fragments used in the transformation experiments contain the sequences required in *cis* for the transcriptional activation of the genes during erythroid cell differentiation. Recently, analysis of the transcription of hybrid globin genes in MEL cells revealed that at least two distinct regulatory sequences

* Corresponding author.
† Present address: European Molecular Biology Laboratory, 6900 Heidelberg, Federal Republic of Germany.
‡ Present address: Regulatory Biology Laboratory, The Salk Institute, La Jolla, CA 92037.

are sufficient for globin gene regulation during MEL cell differentiation, one located 5' to the gene (61) and the other located downstream from the mRNA capping site of the gene (10, 61). An analysis of 5' deletions in the rabbit β-globin gene revealed that regulation in MEL cells does not require more than 58 bp of flanking sequence (61). In the experiments described in this paper we examined the effect of 5' deletions and LS mutations on the regulation of the mouse β-major-globin gene in MEL cells. We found that none of the 5' deletion or LS mutations tested qualitatively affects the regulation of the cloned mouse β-major-globin gene during MEL cell differentiation. These results are consistent with the conclusion that sequences located downstream from the β-globin promoter are sufficient for β-globin gene regulation in MEL cells (10, 61).

MATERIALS AND METHODS

Construction of β-globin gene plasmids. Plasmid πMHβSV-330 (R. Myers, unpublished data) was constructed by subcloning an *Hin*dIII-*Pst*I DNA fragment carrying the hybrid mouse/human β-globin gene into miniplasmid πSVHPlac (33). Plasmid pMHβ20 was constructed in the following way. A derivative of plasmid pMHβ (8) carrying the hybrid mouse/human β-globin gene was obtained by deletion of pBR322 sequences and mouse β-major-globin gene 5'-flanking sequences between a *Pvu*II site and a *Sac*I site and insertion of an *Xho*I linker at the junction. Subsequently, the sequences between a *Cla*I site in pBR322 DNA and a *Bgl*II site in the 3'-flanking region of the human β-globin gene were deleted. Finally, the 5' region of the mouse β-major-globin gene was replaced by an homologous region carrying a *Bgl*II linker insertion in an *Hin*cII site at position +28. Plasmid pMHβ30 carrying the pseudo-wild-type gene was constructed by filling in the *Bgl*II site of plasmid pMHβ20 with the Klenow fragment of *Escherichia coli* DNA polymerase I, in the presence of the four deoxyxanthosine 5'-triphosphates. Plasmid pMHβ40 was derived from plasmid pMHβ20 in the following way. Plasmid pMHβ20 DNA was digested with *Xho*I and *Bgl*II, and the termini were filled in with the DNA polymerase Klenow fragment, in the presence of the four deoxyxanthosine 5'-triphosphates, and then ligated. In this process both sites are preserved. Plasmid pMHβSV-600 was obtained by ligation of *Pst*I-*Eco*RI fragments from pMHβ30 and pMHβ1 (the latter plasmid contains the hybrid mouse/human β-globin gene, the simian virus 40 [SV40] origin of replication and 72-bp repeats, and the M13 origin of replication: Myers, unpublished data). Plasmid pMHβSV-330 was obtained by ligation of *Hin*dIII-*Eco*RI DNA fragments from pMHβ30 and pMHβ1. Plasmid πSV^{thal} contains the −87 mutant human β-globin gene and the SV40 control region on a πVX vector (56). Plasmid pAβ30 was constructed by recombination through an *Xba*I restriction site of plasmid pMHβ30 and plasmid pTAP (P. Mellon, unpublished data), carrying the pBR322 tetracycline resistance gene and the hamster aprt gene. This results in partial direct duplication of pBR322 sequences.

Construction of deletion and LS mutants of the pseudo-wild-type gene. The LS mutants were obtained by matching 5' and 3' deletions covering the promoter region. The series of 3' deletions was constructed as follows. Plasmid pMHβ20 (Fig. 1) was linearized by digestion with restriction enzyme *Bgl*II and then treated with nuclease BAL 31 (New England BioLabs, Inc.). Samples were taken every 2 min, and the reaction was stopped by the addition of ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*'-tetraacetic acid to 20 mM. A portion of each sample was digested with *Hin*dIII, and the DNA was labeled at its 3' ends and then analyzed by polyacrylamide gel electrophoresis to evaluate the extent of the BAL 31 digestion. Deletions extending on the average to positions +10 to −150 were found in samples treated with BAL 31 for 2 to 14 min, respectively. These samples were pooled, and the DNA was ethanol precipitated and suspended in 13 μl of buffer containing 66 mM Tris (pH 7.6), 1 mM ATP, 1 mM spermidine, 10 mM $MgCl_2$, 15 mM dithiothreitol, 200 μg of bovine serum albumin per ml, 1 μg of phosphorylated *Bgl*II linker (CAGATCTG; New England BioLabs), and 30 ng of ^{32}P-phosphorylated *Bgl*II linker. T4 DNA ligase (200 U; New England BioLabs) was added, and the sample was incubated for 12 h at 15°C. The DNA was then diluted into 400 μl of restriction buffer and after the addition of 200 U of *Bgl*II restriction enzyme was incubated at 37°C for 2 h. Another 200 U of *Bgl*II restriction enzyme was then added, and incubation at 37°C was continued for 15 h. Finally, 40 U of *Hin*dIII restriction enzyme was added, and the sample was incubated for 2 h at 37°C. The DNA was then ethanol precipitated and fractionated by electrophoresis on a 5% polyacrylamide gel. After autoradiography, the region of the gel containing *Hin*dIII-*Bgl*II promoter DNA fragments with deletions extending to positions +10 to −130 was divided into 23 slices, and DNA was eluted from each slice and suspended into 20 μl of TE buffer (10 mM Tris [pH 8], 1 mM EDTA).

To clone the deleted promoter fragments, plasmid pMHβ20 was treated with *Bgl*II and *Hin*dIII, and the large restriction DNA fragment was purified by preparative centrifugation at 31,000 rpm on a 5 to 20% sucrose gradient in TE buffer containing 1 M NaCl, in a Beckman SW41 rotor at 20°C for 15.5 h. The vector DNA fragment was then ligated to the promoter DNA fragments eluted from each gel slice. The mixture was then used to transform competent *E. coli* HB101. A total of 10 to 100 ampicillin-resistant colonies were obtained per transformation. For each eluted DNA fraction, minipreparations of plasmid DNA were carried out for 12 to 24 clones (35). About 80% of the clones contained DNA with a deletion in the promoter region, as analyzed by agarose gel electrophoresis. The sizes of the deletions were then compared by polyacrylamide gel electrophoresis, and the different deletions were further characterized by DNA sequencing by the Maxam and Gilbert procedure (39).

5' deletions of the promoter region were constructed by a similar procedure, starting from plasmid pMHβ30 (Fig. 1). Deletions were initiated from a unique *Hin*dIII restriction site, by digestion with *Hin*dIII, followed by treatment with BAL 31 nuclease. *Bgl*II linkers were ligated to the truncated DNA molecules, and the *Bgl*II-*Sac*I fragments carrying the deletions were subsequently cloned in place of the small *Bgl*II-*Sac*I fragment in plasmid pMHβ40 (Fig. 1).

The LS mutants were assembled by joining pairs of 3' and 5' deletions through the *Bgl*II restriction site. In a few cases, the *Bgl*II termini were filled before ligation with the Klenow fragment of *E. coli* DNA polymerase I, in the presence of the four deoxynucleotide triphosphates. The linker substitutions of the mutants obtained by this last procedure were subsequently sequenced by the Maxam and Gilbert procedure (39).

Cell culture and DNA transfection and transformation. HeLa cells were grown in Dulbecco modified Eagle medium with 10% fetal bovine serum. The aprt$^-$ MEL cell line (13) was maintained in Dulbecco modified Eagle medium supplemented with 15% horse serum and 50 μg of diaminopurine

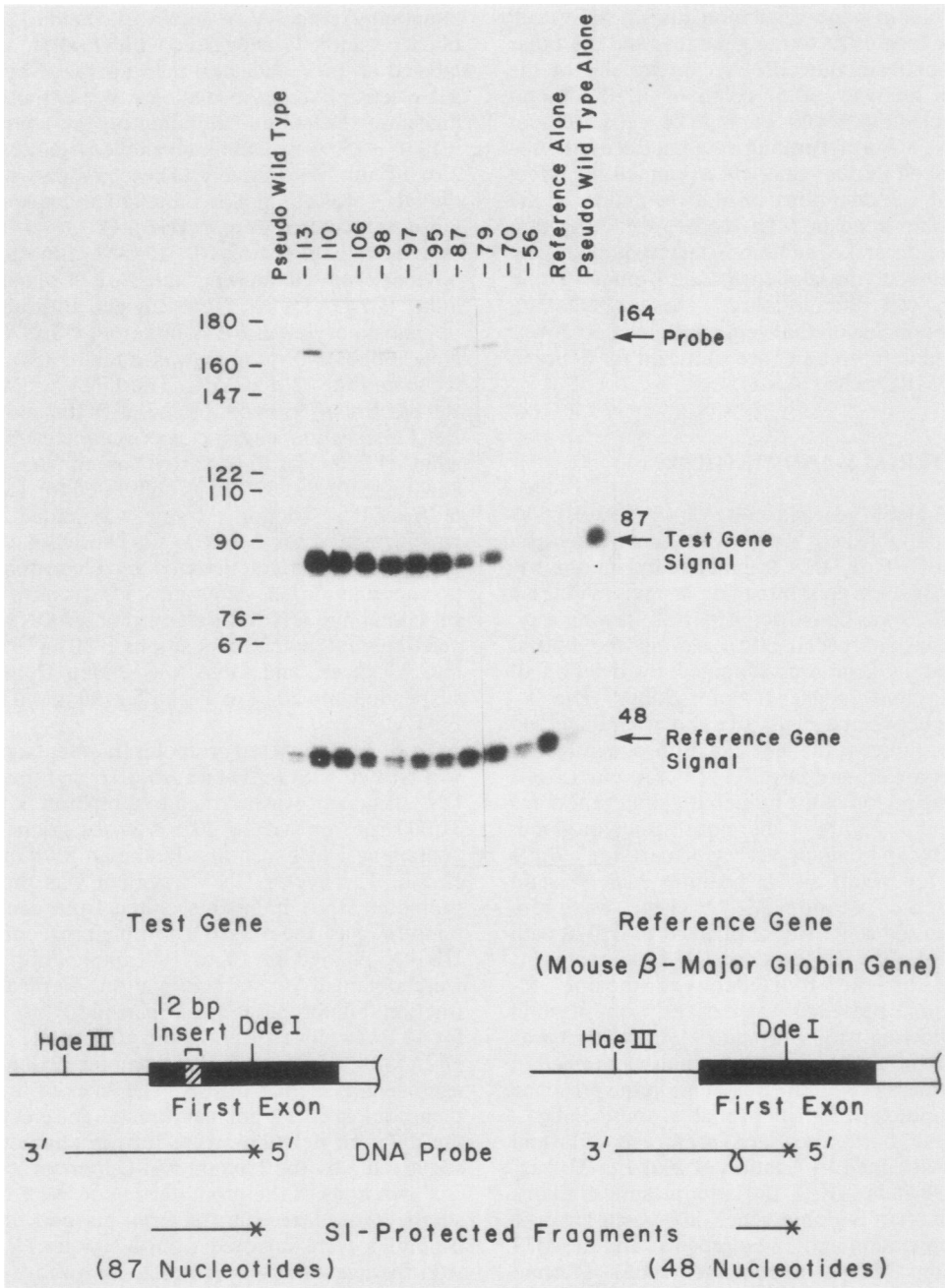

FIG. 1. Effect of 5′ deletions on β-globin gene transcription in HeLa cells. Shown is an autoradiogram of an S1 nuclease analysis of 20 μg of total RNA from HeLa cells transfected with equal amounts of the test plasmid and the reference plasmid. The RNA was hybridized to a 5′-end-labeled single-stranded DNA probe derived from the pseudo-wild-type gene (see diagram at the bottom for a description of the probe). The S1 nuclease-resistant DNA fragments were analyzed by electrophoresis on an 8% polyacrylamide sequencing gel. Hybridization of the probe with pseudo-wild-type mRNA protects a series of fragments about 87 nt in length, whereas hybridization with reference mRNA protects a series of fragments about 48 nt in length. The first 11 lanes correspond to RNA extracted from cells transfected with the reference plasmid and either the pseudo-wild-type gene or 5′-external-deletion mutant genes. The last two lanes correspond to RNA extracted from cells transfected with only the reference gene or the pseudo-wild-type gene. The low level of the reference gene signal in the lane containing the reference gene alone is atypical. Usually, the reference gene signal is the same in side-by-side transfections of the reference gene alone or in combination with the test gene. The same low-level reference gene RNA sample was used in all of the lanes with the reference gene alone in Fig. 1 and 5.

per ml. The aprt$^+$ transformants were maintained in Dulbecco modified Eagle medium containing 10% fetal bovine serum, 4 μg of azaserine per ml, and 15 μg of adenine per ml. HeLa cells were transfected at 30 to 50% confluence by exposure to a DNA calcium phosphate coprecipitate as described previously (3, 58). In each case, 10 μg each of the test plasmid and of the reference plasmid were transfected per 9-cm plate. After exposure to the precipitate for 12 to 16 h at 37°C, cells were washed twice with warm phosphate-buffered saline, and then fresh medium was added. Stable

transformation of aprt⁻ MEL cells was carried out by either calcium phosphate precipitation or protoplast fusion procedures as described previously (8, 10, 54, 55).

Induction of MEL cell differentiation and RNA preparation. MEL cell differentiation was induced by growth for 4 days in Dulbecco modified Eagle medium containing 12% fetal bovine serum and 2% DMSO. Cytoplasmic RNA was extracted from uninduced and DMSO-induced cells after homogenization as described previously (8, 9). Total cellular RNA from HeLa cells was prepared as described previously (57, 58).

S1 nuclease analysis. We used restriction fragments from the pseudo-wild-type gene and the human β-globin gene spanning the RNA capping site to detect the 5' ends of transcripts from these genes. The fragments were labeled at their 5' ends with polynucleotide kinase in the presence of [γ-^{32}P]ATP (7,000 Ci/mmol; New England Nuclear Corp.), and single strands were isolated by electrophoresis on a urea-polyacrylamide sequencing gel. Hybridization, S1 nuclease treatment, and analysis of protected fragments were carried out as described previously (6, 10, 57, 59). Exposures to X-ray film were without an intensifying screen. The quantitation of different RNA species was obtained by exposing the autoradiograms for different lengths of time and by scanning the bands with a densitometer.

RNA mapping with SP6 RNA probes. The plasmid used as the template for synthesis of the complementary strand SP6 RNA probe was constructed by inserting a *Bgl*II-*Dde*I restriction fragment covering the RNA capping site from a deletion mutant of the pseudo-wild-type β-globin gene into SP6-PL1, a plasmid in which a polylinker sequence is inserted 42 bp downstream from the site of initiation of transcription by SP6 RNA polymerase (44). The plasmid was linearized by treatment with *Eco*RI before transcription. Transcription with SP6 RNA polymerase, labeled RNA purification, hybridization, and RNase treatments were as described previously (10, 44, 62).

RESULTS

Effect of 5' deletions in the mouse β-major-globin gene promoter on transcription in HeLa cells. To study the effects of promoter mutations on β-globin gene transcription in HeLa and MEL cells, we made use of a hybrid gene in which the 5' end of the mouse β-major-globin gene is joined to the 3' end of the human β-globin gene through a common *Bam*HI site in the second exon (8). This gene was chosen because it is appropriately regulated when introduced into MEL cells, and its transcripts can be distinguished from those of the resident mouse β-major-globin gene by a 3' S1 nuclease assay (8). To further facilitate the analysis of hybrid gene transcripts we used a derivative of the hybrid mouse/human β-globin gene that carries a 12-bp insert at position +28 in the 5'-untranslated region. This insert allows discrimination between accurately initiated transcripts from the transfected gene, or of its mutant derivatives, and the endogenous mouse β-major-globin gene by 5' S1 nuclease mapping. We will refer to this marked gene as the pseudo-wild-type gene (see Fig. 2C for details of the construction). The plasmid vectors used for the HeLa cell experiments also contain an SV40 transcription enhancer sequence (Fig. 2E and F), since as in the case of the rabbit (3) and human (30, 58) β-globin genes, the hybrid mouse/human β-globin gene must be linked to an enhancer sequence to observe transcription in HeLa cells (R. Myers, personal communication).

Transient expression assays were carried out by cotransfecting HeLa cells with two plasmids, one carrying the pseudo-wild-type gene or a mutant derivative (test gene) and the other (pMHβSV-330, Fig. 2A) carrying a reference gene. The reference gene was the hybrid mouse/human β-globin gene lacking the 12-bp insertion at position +28. The hybridization probe used to analyze RNA from transfected HeLa cells was a 5'-^{32}P-labeled restriction fragment spanning the RNA cap site of the pseudo-wild-type gene (Fig. 1). Transcripts initiating at the mRNA cap site of the test genes should protect a 5'-end-labeled DNA fragment 87 nucleotides (nt) in length from S1 nuclease digestion (31). As expected, a series of fragments of approximately 87 nt are detected when the hybridization assay is carried out with RNA from HeLa cells transfected with the pseudo-wild-type gene (Fig. 1, pseudo-wild-type gene alone). Some full-length probe protection is also observed, indicating the presence of transcripts that are initiated upstream from the promoter (Fig. 1). Transcripts initiating from the cap site of the reference gene should protect a fragment of 48 nt from S1 nuclease digestion when hybridized to the 5'-end-labeled pseudo-wild-type DNA probe (Fig. 1). As expected, RNA from HeLa cells transfected with both the test and reference genes protects 87- and 48-nt fragments from S1 nuclease digestion (Fig. 1, pseudo-wild-type gene). The relative levels of transcription of the test and reference genes can be quantitated by determining the relative intensities of the 87- and 48-nt bands. By using the convention of Dierks et al. (14), the relative transcription level (RTL) is defined as the ratio of mutant gene transcripts to reference transcripts divided by the ratio of pseudo-wild-type transcripts to reference transcripts. Although variability in the reference gene signal was observed (Fig. 1; see Fig. 5), due presumably to differences in transfection efficiency and RNA recovery, the RTL determined in two independent experiments never differed by more than 10%. Moreover, we have not observed competition between the cotransfected reference gene and wild-type test genes. Thus, we believe that the RTL is a valid measure of promoter strength (also see reference 14).

To determine the 5' boundary of the DNA sequences required for efficient transcription from the mouse β-major-globin promoter, we constructed a series of 5' deletions with end points located between −600 and −56 bp upstream from the mRNA cap site. Each deletion mutant was then tested by cotransfection with the reference gene into HeLa cells. The mutant genes were in the context of the pseudo-wild-type gene in plasmid pMHβSV-600 (Fig. 2E). None of the deletions affect the site at which transcription is initiated, since accurately transcribed RNA is observed with every deletion mutant tested (Fig. 1). However, certain deletions cause a significant decrease in the level of accumulated transcripts (Fig. 1 and 3). A four- to fivefold decrease in the RTL is observed when the sequences between −95 and −87 are removed, and an additional fivefold decrease is observed when sequences between −79 and −70 are deleted (Fig. 3). Removal of sequences between −70 and −55 does not lead to a further decrease in the level of transcription. Since the −113 to −95 deletion mutants produce similar levels of transcription and since this level is about threefold higher than that of the pseudo-wild-type gene with 600 bp of 5'-flanking sequences (Fig. 3), we normalized the RTL by using the level of transcription of the −113 deletion mutant rather than that of the pseudo-wild-type gene (Fig. 3). The increase in the level of transcription observed when the sequences between −600 and −113 are removed may be the result of deleting a sequence that partially prevents the action of the SV40 enhancer sequence.

In conclusion, we find that only 95 nt upstream from the

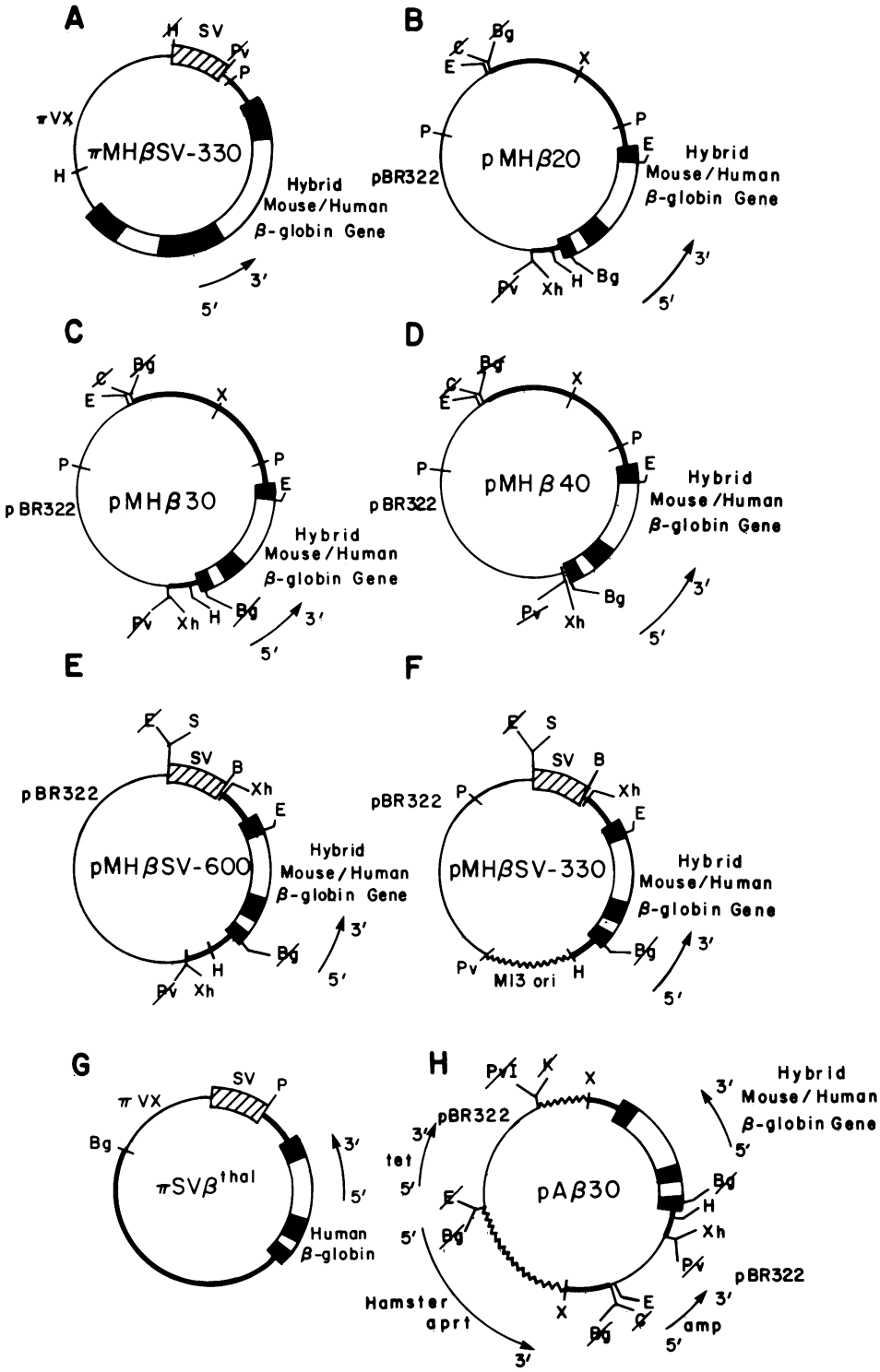

FIG. 2. Structure of β-globin gene plasmids used in the HeLa and MEL cell transfection studies. The construction of plasmids A through H is described in the text. (A) πMHβSV-330: SV, SV40 *Hin*dIII-*Pvu*II fragment containing the origin of replication and the transcription enhancer sequence; πVX, small *supF*-containing plasmid constructed by B. Seed (56); arrow labeled 5′ to 3′, direction of β-globin transcription. (B) pMHβ20: pBR322, sequences from plasmid pBR322. (C) pMHβ30. (D) pMHβ40. (E) pMHβSV-600. (F) pMHβSV-330: wavy line, phage M13 origin of replication. (G) $\pi SV\beta^{thal}$. (H) pAβ30: wavy line, hamster aprt gene. The restriction enzyme cleavage sites are labeled as follows: P, *Pst*I; H, *Hin*dIII; Pv, *Pvu*I; B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; C, *Cla*I; X, *Xba*I; Xh, *Xho*I; S, *Sal*I.

-110 -100 -90 -80 -70 -60 -50		RTL Exp.1	RTL Exp.2
ACCGAAGCCTGATTCCGTAGAGCCACACCCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	Pseudo-wild type	0.4	0.3
GAAGCCTGATTCCGTAGAGCCACACCCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	Deletion -113	1.	1.
GCCTGATTCCGTAGAGCCACACCCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	-110	0.8	0.9
GATTCCGTAGAGCCACACCCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	-106	1.1	1.2
AGAGCCACACCCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	-98	1.1	1.4
GAGCCACACCCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	-97	0.8	1.1
GCCACACCCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	-95	0.8	1.2
CCTGGTAAGGGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	-87	0.2	0.3
GGCCAATCTGCTCACACAGGATAGAGAGGGCAGGA	-79	0.3	0.3
GCTCACACAGGATAGAGAGGGCAGGA	-70	0.04	0.07
GAGAGGGCAGGA	-56	0.03	0.08

FIG. 3. nt sequence and relative transcription levels of 5'-deletion mutants of the mouse β-major-globin gene promoter. The sequence of a portion of the wild-type mouse β-major-globin gene promoter region is displayed on the top line. The sequence of the deletion mutants is represented up to the point of divergence from the wild-type sequence. The sequences immediately upstream from the deletion endpoints are identical, except in cases in which the *Bgl*II DNA linker that was added during the construction of the deletion mutant reconstitutes a portion of the deleted sequence. The RTLs were calculated from two independent experiments.

mRNA capping site are required for efficient transcription from the mouse β-major-globin promoter in HeLa cells. At least two distinct elements are required for transcription. Their 5' limits are located between −95 and −87 and between −79 and −70. The region from −95 to −87 contains the conserved CCACACCC sequence, whereas the region from −79 to −70 includes the CCAAT box.

Effect of LS mutations in the mouse β-major-globin gene promoter on transcription in HeLa cells. A more detailed analysis of the sequences required for transcription from the mouse β-major-globin gene promoter in HeLa cells was accomplished by constructing 21 LS mutations located between −110 and −2 bp upstream from the mRNA cap site. The LS mutants were analyzed in HeLa cells in the context of the pseudo-wild-type gene in plasmid pMHβSV-330, which contains approximately 330 bp of 5'-flanking sequences (Fig. 2F). Each LS mutation is designated by two numbers, which refer to the nt immediately 5' and 3' to the inserted linker sequence (Fig. 4). The base substitutions in the promoter are indicated by the white letters on the black background. The different LS mutations span the entire region between positions −110 and −2. Of the mutations, 11 do not alter the spacing of the adjacent sequences (Fig. 4).

The results of a typical transient expression analysis of the different LS mutations are presented in Fig. 5, and the quantitation of the relative transcriptional levels of the mutant genes is presented in Fig. 4. Consistent with earlier studies of the rabbit β-globin gene promoter (14, 15, 25, 26), we can identify three separate regions in the mouse β-major-globin gene promoter that are necessary for efficient transcription in HeLa cells. In the first region, LS mutation −93/−84 which alters the ACACC sequence within the conserved CCACACCC sequence leads to a five- to sevenfold decrease in the level of transcription. This result is consistent with the fivefold decrease in transcription observed with the −87 deletion (Fig. 3). In the second region, which includes the CCAAT box, LS mutations −83/−70, −77/70, −74/62, −70/62, and −70/56 lead to a decrease in the level of transcription by a factor of 2.5 to 7. Although the first three mutations alter the CCAAT box, the last two lie immediately downstream, indicating that the functionally important sequences extend beyond the highly conserved canonical CCAAT box sequence. In the third region, LS mutations −42/−33, −38/−24, and −37/24, which lie within or immediately upstream from the ATA box, decrease transcription by a factor of 2.5 to 7. Interestingly, LS mutation −42/−33, which lies entirely outside of the canonical ATA box sequence, leads to a 2.5-fold decrease in the level of transcription, suggesting that sequences immediately upstream from the ATA box are required for efficient transcription. These three transcriptional control elements are surrounded by sequences that are sensitive to LS mutagenesis and thus are not essential for efficient transcription. These regions are located between −110 and −96, −86 and −81, −66 and −61, and −20 and −3 (Fig. 5).

In addition to the quantitative effects of the LS mutations mentioned above, three LS mutations appear to affect the site of initiation of transcription. LS mutation −11/−2 leads to the appearance of one or more additional bands corresponding to accurately initiated transcripts (Fig. 5). These bands might reflect a real difference in the site of transcription initiation. Alternatively, they might correspond to RNA molecules initiated more upstream and map the position of divergence between the probe and mutant sequence. LS mutations −37/−24 and −36/−24, which alter the ATA box sequence, generate an additional cluster of bands corresponding to transcripts initiating near position −11 (Fig. 5).

Effect of 5'-deletion and LS mutants on the regulated expression of the hybrid mouse/human β-globin gene in MEL cells. Cloned human or hybrid mouse/human β-globin genes are appropriately regulated when introduced into MEL cells by DNA cotransformation procedures (8, 10, 60). Treatment with DMSO of MEL cells carrying these genes leads to a 5- to 50-fold increase in the level of mRNA from the exogenous β-globin gene (8, 10, 60). To determine whether sequences located in the 5'-flanking region of the hybrid mouse/human

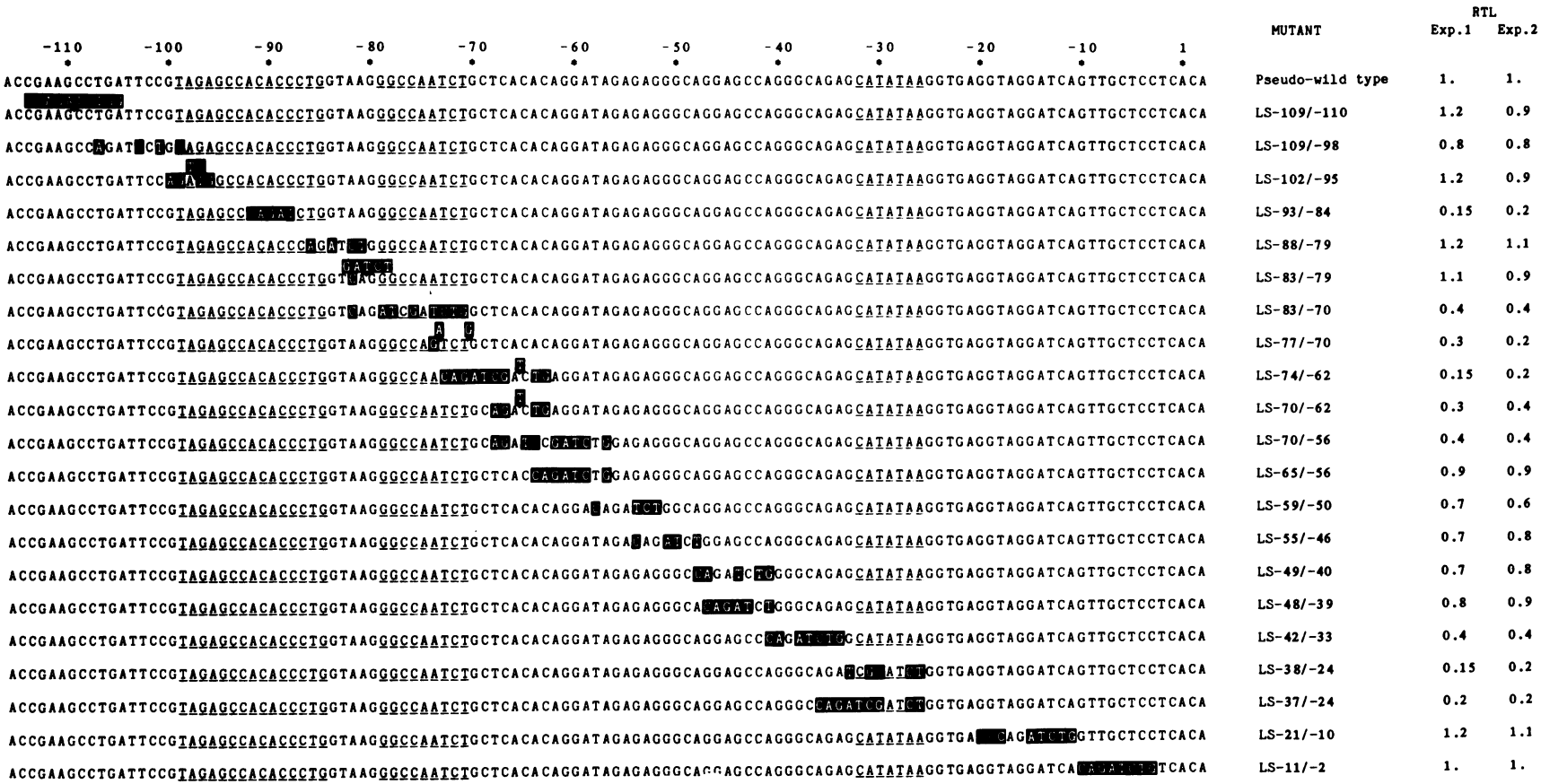

FIG. 4. nt sequence and relative transcription levels of LS mutants in the mouse β-major-globin gene promoter. The nt sequence of the wild-type mouse β-major-globin gene promoter region is displayed on the top line (31). Each LS mutant is designated by the positions of the 5′ and 3′ nt that are immediately adjacent to the linker sequence. The nt that are changed by substitution of promoter sequences by the synthetic DNA linker are indicated by white letters on a black background. In the case of LS mutations that are not exact substitutions, the alignment has been adjusted to provide the best sequence conservation. Additions are represented by nt above the sequence and deletions by bars. LS mutations −83/−70, −74/−62, −70/−56, −38/−24, and −37/−24 were obtained by filling in the *Bgl*II termini before ligation. The 15-nt-long upstream sequence, which is repeated in the rabbit β-globin gene and includes the CACCC box, is underlined Similarly, the highly conserved sequences that include the CCAAT box and the ATA box sequences are also underlined. The RTLs were calculated from two separate experiments.

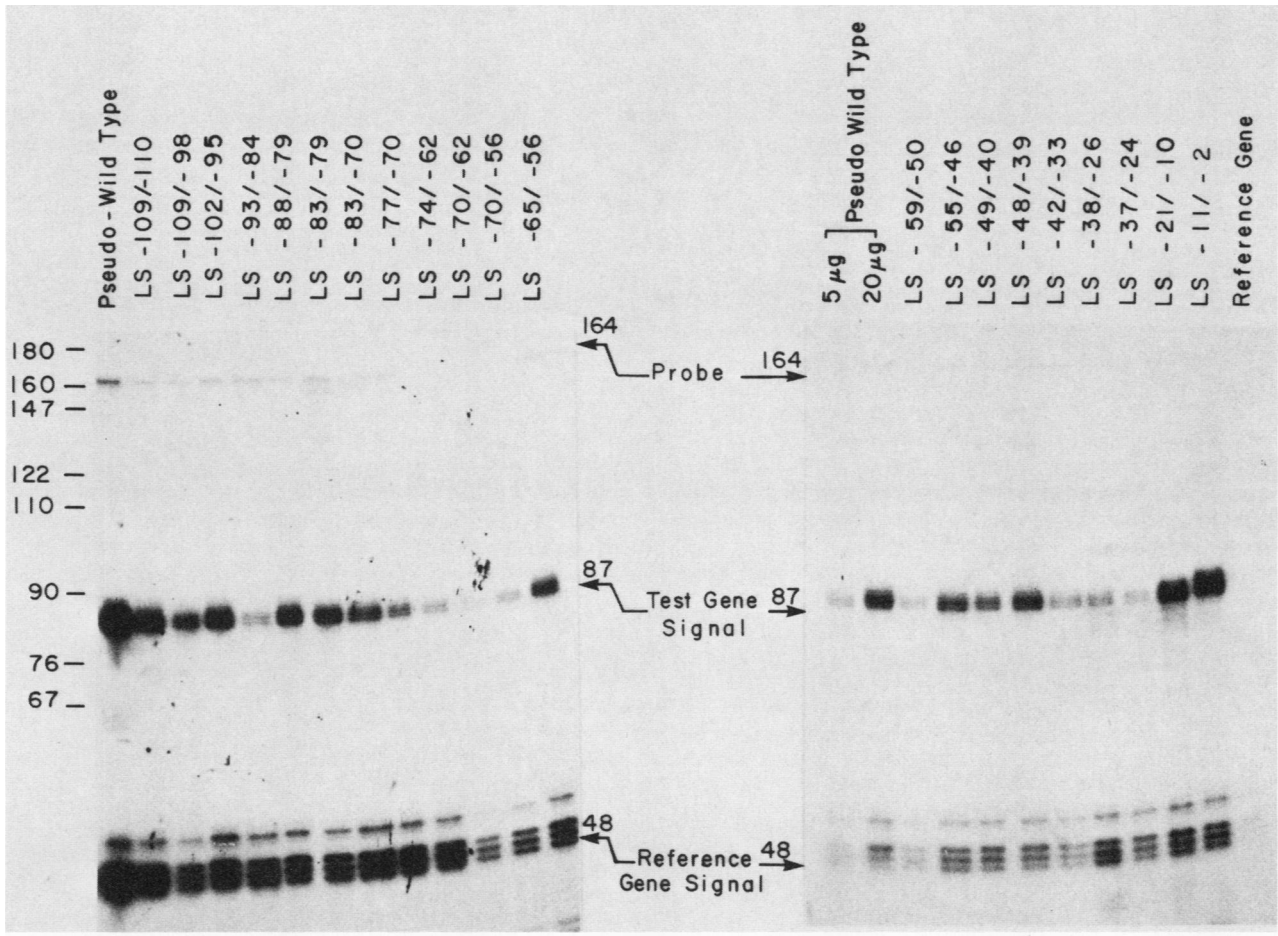

FIG. 5. Effect of LS mutations on β-globin gene transcription in HeLa cells. Shown is an autoradiogram of an S1 nuclease analysis of 20 μg of total RNA from HeLa cells transfected with equal amounts of the test plasmid and the reference plasmid. The RNA was hybridized with the 5'-end-^{32}P-labeled single-stranded DNA probe covering the RNA cap site described in Fig. 1. S1 nuclease-resistant DNA fragments were analyzed by electrophoresis on an 8% polyacrylamide sequencing gel. The first 24 lanes correspond to RNA extracted from cells transfected with the reference gene and with either the pseudo-wild-type gene or the LS mutant genes. The last lane corresponds to RNA extracted from cells transfected only with the reference gene.

β-globin gene are required for this regulation, we analyzed the transcription of deletion and LS mutants in MEL cells. To accomplish this, the pseudo-wild-type mouse/human β-globin gene and its mutant derivatives were introduced into an adenine phosphoribosyltransferase-deficient (aprt$^-$) MEL cell line (8, 13) by cotransformation with the hamster aprt gene (34), by a protoplast fusion procedure (10, 54, 55). The prototype plasmid used in these experiments (pAβ30) carries both the β-globin gene and the aprt gene (Fig. 2H). RNA from pools of 50 to 100 independent colonies of transformed MEL cells was examined before and after DMSO-induced differentiation. The levels of both exogenous and endogenous β-globin mRNAs in the transformed MEL cells were determined by a 5' S1 nuclease assay or by an analogous procedure which involves the use of uniformly ^{32}P-labeled complementary SP6 RNA probes (44, 62).

The results of an analysis of a series of 5'-deletion mutants with endpoints located between −106 and −47 are shown in Fig. 6. Although some deletions decrease the level of RNA from the exogenous β-globin gene in both induced and uninduced MEL cells, all mutants appear to be appropriately regulated. DMSO-induced differentiation of the transformed MEL cells results in a 3- to 10-fold increase in the level of RNA from all of the deleted mutants, including the −47 deletion. The absence of a sequence necessary for β-globin gene regulation in the 5'-flanking sequences suggests that a *cis*-acting regulatory sequence sufficient for regulated expression of the gene is located downstream from position −47. We note that the induction ratios observed for the same deletion mutant in independently derived pools or after independent DMSO treatments of the same pool are variable. Because of this variability we cannot provide an accurate estimate of the induction ratios. Thus, the possibility that other sequences upstream from position −47 also modulate quantitative expression of the gene in a tissue-specific manner cannot be ruled out.

Our results also reveal that deletion of sequences located between −98 and −87, which includes the conserved CCACACCC sequence, leads to a significant decrease in the levels of hybrid mouse/human β-globin RNA before and after DMSO-induced differentiation (Fig. 6). Thus, these sequences appear to be involved in the modulation of the efficiency of transcription of the mouse β-major-globin gene in MEL cells as well as in HeLa cells. This conclusion is supported by the results of a comparative analysis of clones of MEL cells transformed with either the human β-globin

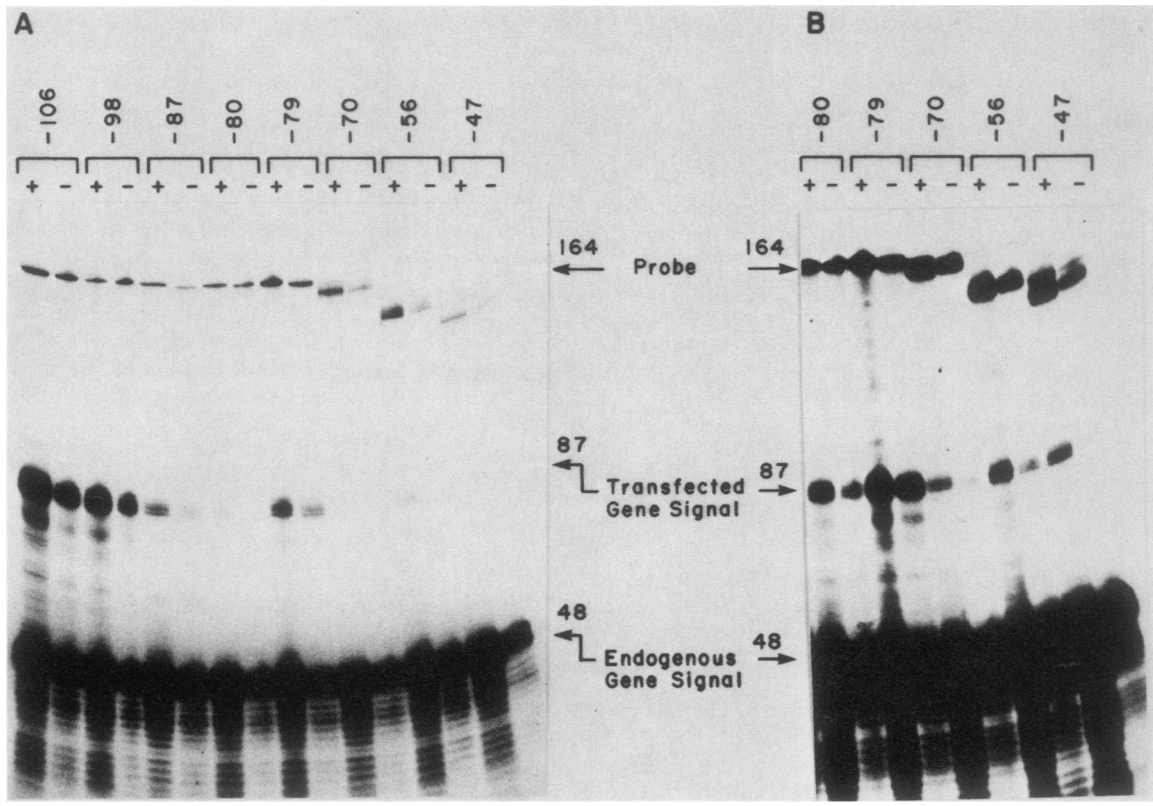

FIG. 6. Effect of 5'-deletion mutations on β-globin gene transcription in MEL cells. Pools of about 100 independent MEL cell colonies transformed with the 5'-external-deletion mutant genes were obtained by the protoplast fusion procedure. The pools were grown for 4 days in the presence (+) or absence (−) of 2% DMSO. Cytoplasmic RNA (20 μg) was hybridized to the 5'-end-^{32}P-labeled single-stranded DNA probe covering the RNA capping site described in Fig. 1. S1 nuclease-resistant DNA fragments were analyzed by electrophoresis on an 8% polyacrylamide sequencing gel. (A) The autoradiogram was exposed for 16 h. (B) Portion of an autoradiogram of the same gel exposed for 1 week.

gene or a thalassemia mutant of the same gene carried on plasmid πSVβthal (Fig. 2G). This mutation is a C-to-G transversion at position −87, within the CACCC sequence (47). The level of transcription of this gene is 10-fold less than that of the normal human β-globin gene in the HeLa cell transient expression assay (57). Similarly, the level of expression of the thalassemia gene is markedly reduced in the three clones analyzed, compared with that of the wild-type gene, both before and after DMSO treatment of the cells (Fig. 7). However, the ratio between the levels of expression before and after cell differentiation is not affected by the mutation. Thus, the −87 mutation is likely to reduce the level of expression of the gene by altering a nonspecific transcriptional control element rather than a tissue-specific regulatory sequence.

We also examined the transcription of several LS mutants of the hybrid mouse/human β-globin gene transferred into MEL cells by protoplast fusion. The level of exogenous globin mRNA was quantitated with uniformly ^{32}P-labeled complementary SP6 RNA probes (Fig. 8). Hybridization of the probe with RNA transcribed from the pseudo-wild-type gene protects a cluster of fragments about 87 nt in length, whereas hybridization of the same probe with endogenous mouse β-major-globin mRNA protects fragments about 48 nt in length (Fig. 8). In the case of the LS mutants analyzed in Fig. 8, DMSO induction leads to a 4- to 15-fold increase in the level of hybrid mouse/human β-globin mRNA. Similarly, for three other LS mutants examined (−88/−79, −65/−56, and −55/−46), we observed a 4- to 20-fold accumulation of exogenous globin mRNA upon differentiation of the cells, in at least one of the two pools analyzed (data not shown). Thus, none of the LS mutations tested significantly affects the regulation of the hybrid mouse/human β-globin gene during MEL cell differentiation.

DISCUSSION

DNA sequences required for accurate and efficient transcription in HeLa cells. In this paper we report a detailed study of the sequence requirements for efficient transcription from the mouse β-major-globin gene promoter in the HeLa and MEL cells. Consistent with studies of the rabbit β-globin gene (14, 15, 25, 26), we find that the mouse β-major-globin gene promoter region contains at least three separate DNA sequence elements that are necessary for efficient transcription. These elements are located between positions −40 and −21, −80 and −65, and −95 and −87 and are designated the ATA, the CCAAT, and the CACCC box elements, respectively. Mutations located between these different transcription control elements do not significantly affect the transcription of the gene. Although the HeLa cell transient expression assay used here requires an SV40 enhancer sequence for detectable levels of β-globin transcription, the assay does appear to provide an accurate indication of the sequences required for globin gene transcription in erythroid cells. First, a single base change in the ATA or CACCC sequence

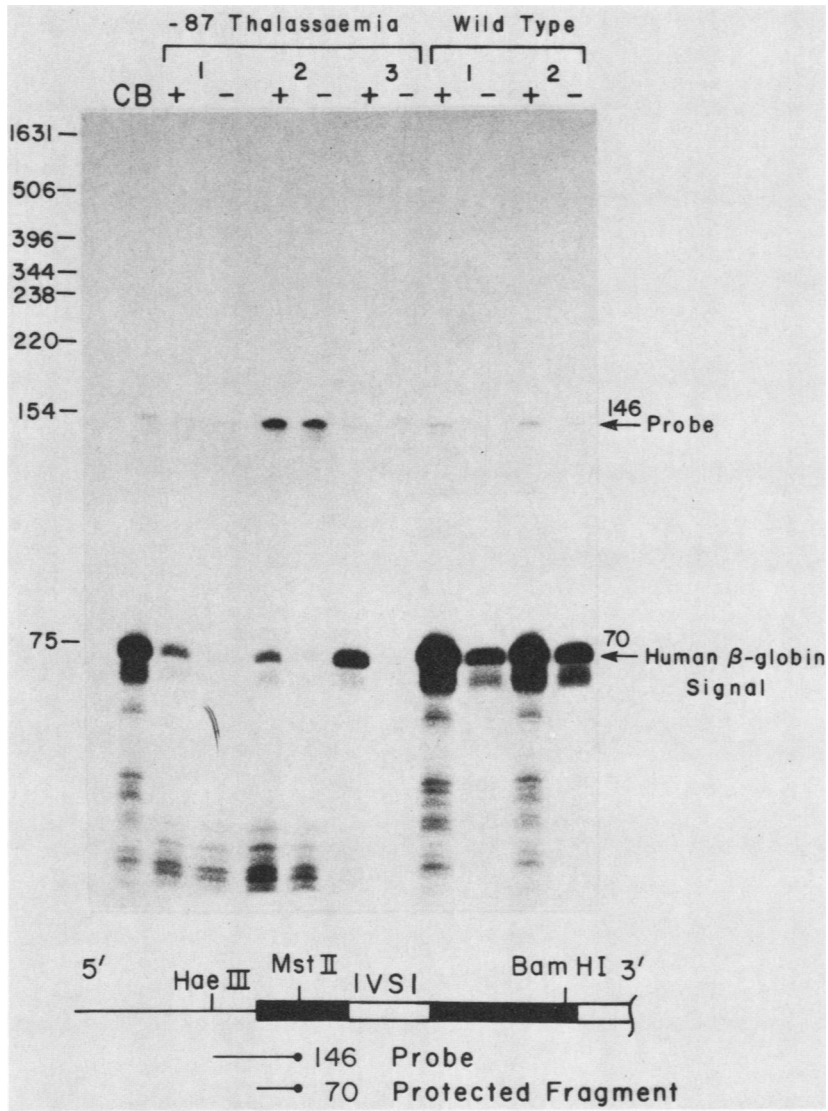

FIG. 7. Comparative analysis of wild-type and −87 mutant human β-globin gene transcription in transformed MEL cell lines. Cloned MEL cell transformants containing either the −87 mutant or the wild-type human β-globin gene were grown for 4 days in the presence (+) or absence (−) of 2% DMSO. Cytoplasmic RNA (20 μg) was hybridized to the 5'-end-^{32}P-labeled single-stranded DNA probe covering the RNA cap site shown at the bottom. S1 nuclease-resistant DNA fragments were analyzed by electrophoresis on an 8% polyacrylamide sequencing gel. Human β-globin mRNA protects a 70-nt-long fragment. The first lane corresponds to RNA extracted from human cord blood. Three independent lines transformed with the thalassemic gene and two independent lines transformed with the wild-type gene were analyzed.

associated with β-thalassemia decreases the level of transcription in both the HeLa cell assay and in the erythroid cells of the patient (48, 57). Second, in this paper we show that mutations that decrease the level of transcription from the mouse β-major-globin promoter in the HeLa cell assay also decrease the level of transcription in MEL cells before and after induction. Since the globin genes introduced into MEL cells are not linked to an enhancer and since the SV40 enhancer does not affect β-globin gene transcription in MEL cells (unpublished data), it is unlikely that the enhancers used in the HeLa cell experiments alter the sequence requirements for transcription from the β-globin promoter.

ATA box element. Most eucaryotic promoters contain an A-T-rich sequence 30 bp upstream from the transcription initiation site (7; Goldberg, Ph.D. thesis). This sequence is required for efficient transcription of most eucaryotic genes tested in vitro and in vivo (see reference 7 for review), and in vitro transcription factors that interact with this sequence have been described previously (12, 52). In the case of some viral promoters, however, there is no discernible ATA box or it can be deleted without affecting the level of transcription (1, 4, 18, 20, 37, 45, 51). The ATA box is essential for mammalian β-globin gene transcription. Base substitutions in the rabbit or human ATA box region decrease the level of transcription (Fig. 9). Within the rabbit β-globin ATA box sequence an A-to-G substitution at −27, or changing the AA sequence to a GT sequence at −28 and −29, significantly decreases the level of transcription (Fig. 9). An A-to-G substitution at position −28 of the human gene decreases the level of transcription by a factor of 5 (Fig. 9). Similarly, LS mutations affecting the ATA box in the mouse β-major-globin gene promoter cause a five- to sevenfold decrease in the level of transcription (Fig. 9).

LS mutation −42/−33, which modifies sequences imme-

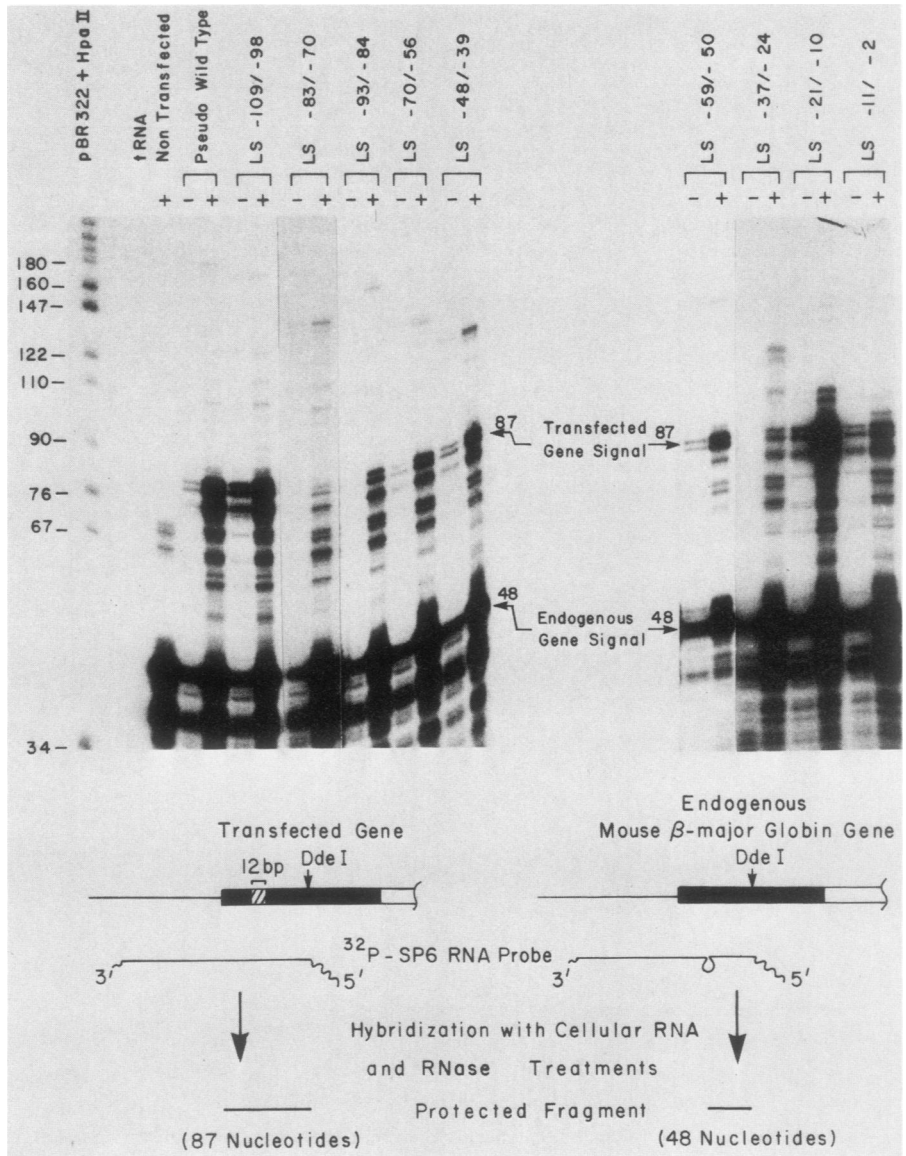

FIG. 8. Transcription of LS mutants in pools of transformed MEL cell colonies. Pools of about 100 independent MEL cell colonies transformed by protoplast fusion with either the pseudo-wild-type gene or LS mutant genes were grown for 4 days in the presence (+) or absence (−) of 2% DMSO. Cytoplasmic RNA (20 μg) was hybridized to an excess of a uniformly ^{32}P-labeled complementary strand SP6 RNA probe covering the RNA cap site and described at the bottom. Fragments protected from a subsequent RNase digestion were analyzed by electrophoresis on an 8% polyacrylamide sequencing gel. Hybridization of the probe to pseudo-wild-type mRNA protects a series of fragments about 87 nt in length, whereas hybridization to the endogenous mouse β-major-globin gene transcripts protects a series of fragments about 48 nt in length. First lane, plasmid pBR322 DNA digested with *Hpa*II; second lane, hybridization of the probe to tRNA; third lane, RNA extracted from untransformed, DMSO-treated MEL cells.

diately upstream from the canonical ATA box sequence, reduces the level of mRNA transcribed from the gene by a factor of 2.5 (Fig. 9). This observation suggests that the functional ATA box element extends beyond the limits of the evolutionarily conserved sequence. However, we cannot exclude the possibility that the decreased level of transcription is an indirect effect of introducing the linker close to the ATA sequence.

The ATA box is also considered to be an essential component in determining the start site of transcription (4, 14, 20, 22, 25, 37, 51). Consistent with this conclusion, we observe RNA molecules transcribed from the mutant gene LS −37/−24 that appear to initiate at approximately position −11, a site which is not used for initiation in the wild-type gene. Nevertheless, a substantial proportion of the transcripts from this mutant gene have correct 5′ ends. This result is consistent with the possibility that the mRNA capping site sequence may contribute to the determination of the transcription initiation site (14).

CCAAT box element. The second mutation-sensitive region revealed by our studies includes the CCAAT box. The corresponding region was shown to be required for transcription of the rabbit β-globin gene (14, 25, 26) and the human α-globin gene (43) in transient expression assays. However, the CCAAT box is not required for transcription of the sea urchin histone H2A and the herpes simplex virus tk genes

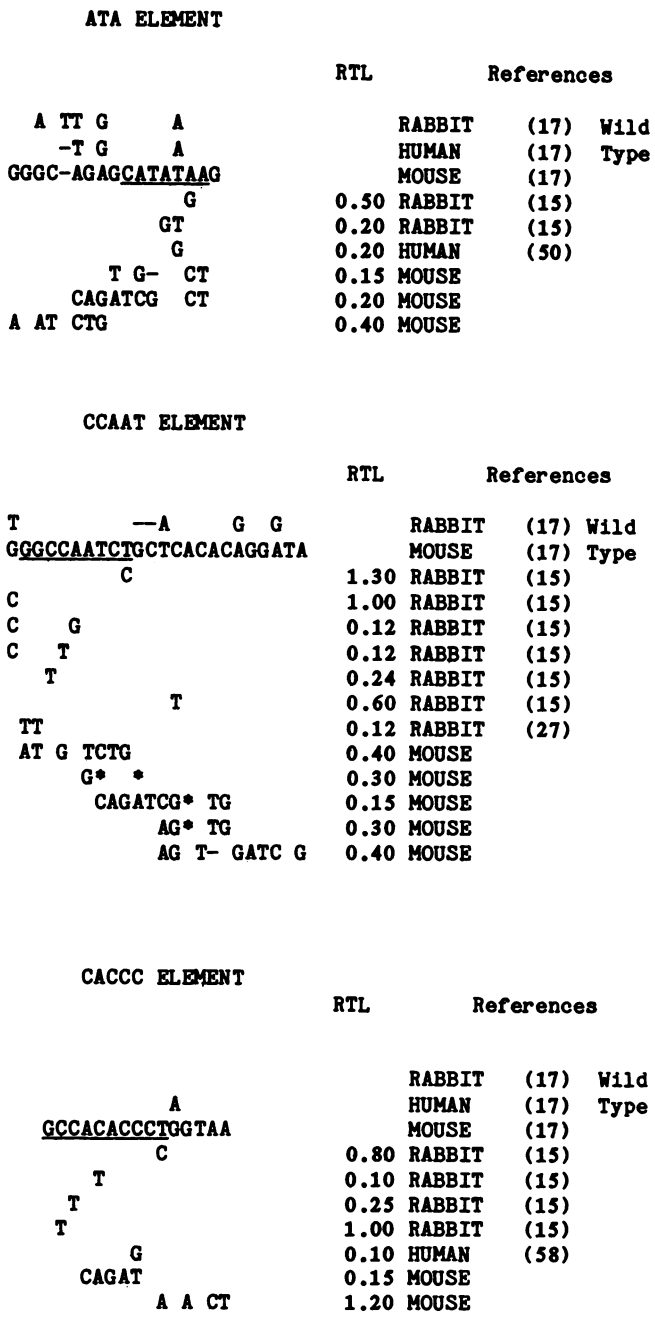

FIG. 9. Summary of base substitution mutations in mammalian β-globin gene promoters. Shown is a compilation of published base substitution mutations that affect β-globin gene transcription in the HeLa cell transient expression assay. Mutations in each of the three conserved elements are presented separately. The wild-type sequences for the rabbit, human, and mouse promoters are shown at the top of each list. The highly conserved sequences are underlined. A dash indicates that the base is missing, and an asterisk indicates the insertion of a base.

(22–24, 40–42). Direct evidence for a role of the CCAAT sequence in β-globin gene transcription was provided by the analysis of single- or double-base mutations in the canonical CCAAT box sequence of the rabbit β-globin gene, which decrease the level of transcription by as much as 10-fold (Fig. 9). In the case of the mouse β-major-globin gene promoter, LS mutations affecting the CCAAT sequence have a three- to sevenfold effect on the level of transcription (Fig. 9). In addition, we found that LS mutations affecting the sequence TCAC located immediately downstream from the CCAAT box (between positions −68 and −65) decrease the level of transcription by a factor of 3 (Fig. 9). It is possible that the TCAC sequence is part of the same transcriptional control element as is the CCAAT sequence. This possibility is suggested by the observation that the decrease in the level of transcription caused by the −70 deletion which removes the CCAAT sequence is the same as that caused by the −55 deletion which removes both the CCAAT and TCAC sequences. Similarly, LS mutation −74/−62, which modifies both sequences, reduces the level of transcription only slightly more than do the LS mutations that alter the CCAAT sequence alone. Although the TCAC sequence is relatively well conserved among mammalian β-globin genes (17, 28, 32), it is not found in the same position with respect to the CCAAT sequence in many other eucaryotic genes which we examined. Thus, we cannot exclude the possibility that the phenotype of LS mutations −70/−62 and −70/−56 is due to an indirect perturbation of the neighboring CCAAT box transcriptional control element. However, it is worth noting that the TCAC sequence is also found in the conserved CCT/ACACCCT sequence present in the adult mammalian β-globin genes (14). Furthermore, Dierks and collaborators reported that a C-to-T transition in the TCAC sequence of the rabbit β-globin gene results in a nearly twofold decrease in the level of transcription (14; Fig. 9).

CACCC box element. A third promoter element was identified in the mouse β-major-globin gene promoter, upstream from the CCAAT sequence and within the sequence GCCACACCC located between positions −95 and −87. LS mutation −93/−84, which affects only the sequence ACACC, decreases the level of transcription by a factor of 5 to 7 (Fig. 4). In contrast, LS mutations −102/−95 and −88/−79 do not affect the efficiency of transcription (Fig. 4). The rabbit and human β-globin gene promoters contain imperfect tandem repeats of a 15-nt sequence which includes the CACCC box (14). Transcription studies of the rabbit β-globin gene indicate that both copies of this sequence are required for maximal levels of transcription. Analysis of the substitution mutations available in that sequence indicates that the CA/TCACCC sequence is the essential component of these repeated transcriptional control elements (Fig. 9). Thus, although the same sequence element appears to be required for efficient transcription of the rabbit (14, 26), human (57), and mouse β-globin genes, only one copy is present in the mouse promoter, whereas two are found in the other two promoters. As previously noted, the CACCC box is highly conserved among mammalian β-globin genes (17, 28, 32). It is possible that the CACCC box is a recognition site for a promoter-specific transcription factor like the HeLa cell factor SP1 which binds to the upstream sequences of the SV40 early promoter (16).

In addition to identifying the DNA sequences necessary for transcription, the LS mutagenesis procedure also identifies sequences within the promoter that are insensitive to mutations. For example, our data indicate the existence of four separate regions in which linker substitutions do not lead to a significant decrease in the level of transcription. These regions are located between positions −110 and −96, −86 and −81, −64 and −41, and −20 and −3 (Fig. 4). Interestingly, these regions are reasonably well conserved among adult β-globin genes (17, 28). Thus, although these regions are not required for promoter function in non-

erythroid cells, they may be involved in globin gene regulation. An alternative explanation is that some of the transcriptional control sequences are redundant. For example, the sequence AGGGCAG is duplicated in the region between the CAATT and ATA boxes (Fig. 4, −53 to −47 and −41 to −37). Thus, this sequence may be necessary for wild-type levels of transcription, but a linker scan in only one of the two copies may not be sufficient to disrupt promoter function.

DNA sequences required for regulated expression of the hybrid mouse/human β-globin gene in MEL cells. Analysis of the 5′-deletion and LS mutants in MEL cells did not reveal any sequences that are required for appropriate regulation of the hybrid mouse/human β-globin gene in these cells. Wright et al. (61) reported that deletions leaving only 58 bp of 5′-flanking sequences do not affect the regulation of the rabbit β-globin gene in MEL cells. These results in conjunction with those from studies of the transcription of hybrid genes in MEL cells suggest that regulatory sequences sufficient for globin gene regulation are located downstream from the mRNA cap site (10, 61). When cloned human α- and β-globin genes are introduced into MEL cells by DNA transformation, the human β-globin gene is appropriately regulated, whereas the human α-globin gene is expressed at the same level in induced and uninduced MEL cells (10). Analysis of the transcription of hybrid human α/β-globin genes revealed that the sequences responsible for these differences in expression are located on the 3′ side of the mRNA capping site of the genes (10). Furthermore, Wright et al. (61) recently demonstrated that a hybrid gene consisting of 5′-flanking sequences from a mouse H2 class I histocompatibility gene fused to the human β-globin structural gene is inducible at the level of transcription in MEL cells. These investigators also demonstrated that a hybrid gene consisting of the human β-globin gene 5′-flanking sequence fused to a mouse H2 class I histocompatibility structural gene is activated during DMSO-induced MEL cell differentiation. Thus, at least two regulatory elements may be involved in the erythroid cell-specific expression of β-globin genes, one located within the 5′-flanking sequences and one located downstream from the mRNA cap site. Regulatory sequences also have been located within immunoglobulin genes (2, 21, 46, 53) and within the adenovirus E1a gene (50). The fact that we were unable to detect regulatory mutations within the promoter region indicates that the sequences within the gene alone are sufficient for globin gene activation during MEL cell differentiation. By examining the transcription of hybrid genes in which mutant β-globin promoters are fused to a nonglobin structural gene, it may be possible to identify sequences within the 5′-flanking region of the mouse β-globin promoter that are necessary for the regulated expression in MEL cells.

ACKNOWLEDGMENTS

We are grateful to Richard Axel for many useful discussions and for comments on the manuscript. We thank R. Myers for providing plasmids and for helpful discussions and Catherine Riney for excellent technical assistance.

P.C. was supported by fellowships from the Fogarty International Center of the National Institute of Health and the Philippe Foundation, and P.M. was supported by a Special Fellowship from the Leukemia Society of America. This work was funded by a grant from the National Science Foundation to T.M.

LITERATURE CITED

1. **Baker, C. C., J. Herisse, G. Courtois, F. Galibert, and E. Ziff.** 1979. Messenger RNA for the Ad2 DNA binding protein: DNA sequences encoding the first leader and heterogeneity at the mRNA 5′ end. Cell **18**:569–580.
2. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
3. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of β-globin gene is enhanced by remote SV40 DNA sequences. Cell **27**:299–308.
4. **Benoist, C., and P. Chambon.** 1981. In vivo sequence requirements of the SV40 early promoter region. Nature (London) **290**:304–310.
5. **Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon.** 1980. The ovalbumin gene-sequence of putative control regions. Nucleic Acids Res. **8**:127–142.
6. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell **12**:721–732.
7. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for protein. Annu. Rev. Biochem. **50**:349–383.
8. **Chao, M. V., P. Mellon, P. Charnay, T. Maniatis, and P. Axel.** 1983. The regulated expression of β-globin genes introduced into mouse erythroleukemia cells. Cell **32**:483–493.
9. **Charnay, P., and T. Maniatis.** 1983. Transcriptional regulation of globin gene expression in the human erythroid cell line K562. Science **220**:1281–1283.
10. **Charnay, P., R. Treisman, P. Mellon, M. Chao, R. Axel, and T. Maniatis.** 1984. Differences in human α- and β-globin gene expression in mouse erythroleukemia cells: the role of intragenic sequences. Cell **38**:251–263.
11. **Corden, J., B. Wasylyk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon.** 1980. Promoter sequences of the eukaryotic protein-coding genes. Science **209**:1406–1414.
12. **Davidson, B. L., J. M. Egly, E. R. Mulvihill, and P. Chambon.** 1983. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature (London) **301**:680–686.
13. **Deisseroth, A., and D. Hendrick.** 1978. Human α-globin gene expression following chromosomal dependent gene transfer into mouse erythroleukemia cells. Cell **15**:55–63.
14. **Dierks, P., A. von Ooyen, M. D. Cochran, C. Dobkin, J. Reiser, and C. Weissmann.** 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. Cell **32**:695–706.
15. **Dierks, P., A. von Ooyen, N. Mantei, and C. Weissmann.** 1981. DNA sequences preceding the rabbit β-globin gene are required for formation in mouse L cells of β-globin RNA with the correct 5′ terminus. Proc. Natl. Acad. Sci. U.S.A. **78**:1411–1415.
16. **Dynan, W. S., and R. Tjian.** 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell **32**:669–680.
17. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the β-globin gene family. Cell **21**:653–668.
18. **Flavell, A. J., A. Cowie, J. R. Arrand, and R. Kamen.** 1980. Localization of three major capped 5′ ends of polyoma virus late mRNA's within a single tetranucleotide sequence in the viral genome. J. Virol. **33**:902–908.
19. **Friend, C., W. Scher, J. G. Holland, and T. Sato.** 1971. Hemoglobin synthesis in murine induced leukemic cells in vitro: stimulation of erythroid differentiation by dimethylsulfoxide. Proc. Natl. Acad. Sci. U.S.A. **68**:378–382.
20. **Fromm, M., and P. Berg.** 1982. Deletion mapping of DNA required for SV40 early region promoter function in vivo. J. Mol. Appl. Genet. **1**:457–481.
21. **Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa.** 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33**:717–728.
22. **Grosschedl, R., and M. L. Birnstiel.** 1980. Identification of regulatory sequences in the prelude sequences of an H2A

histone gene by the study of specific deletion mutants in vivo. Proc. Natl. Acad. Sci. U.S.A. 77:1432–1436.
23. **Grosschedl, R., and M. L. Birnstiel.** 1980. Spacer DNA sequences upstream of the TATAAATA sequences are essential for promotion of H2A histone gene transcription in vivo. Proc. Natl. Acad. Sci. U.S.A. 77:7102–7106.
24. **Grosschedl, R., and M. L. Birnstiel.** 1982. Delimitation of far upstream sequences required for maximal in vitro transcription of an H2A histone gene. Proc. Natl. Acad. Sci. U.S.A. 79:297–301.
25. **Grosveld, G. C., E. de Boer, C. K. Shewmaker, and R. A. Flavell.** 1982. DNA sequences necessary for transcription of the rabbit β-globin gene in vivo. Nature (London) 295:120–126.
26. **Grosveld, G. C., A. Rosenthal, and R. A. Flavell.** 1982. Sequence requirements for the transcription of the rabbit β-globin gene in vivo: the −80 region. Nucleic Acids Res. 10:4951–4971.
27. **Grosveld, G. C., C. K. Shewmaker, P. Jat, and R. A. Flavell.** 1981. Localization of DNA sequences necessary for transcription of the rabbit β-globin gene in vitro. Cell 25:215–226.
28. **Hardison, R. C.** 1983. The nucleotide sequence of the rabbit embryonic globin gene β4. J. Biol. Chem. 258:8739–8744.
29. **Hofer, E., R. Hofer-Warbinek, and J. E. Darnell.** 1982. Globin RNA transcription: a possible termination site and demonstration of transcription control correlated with altered chromatin structure. Cell 29:887–893.
30. **Humphries, R. K., T. Ley, P. Turner, A. D. Moulton, and A. W. Nienhuis.** 1982. Differences in human α-, β- and δ-globin gene expression in monkey kidney cells. Cell 30:173–183.
31. **Konkel, D. A., S. M. Tilghman, and P. Leder.** 1978. The sequence of the chromosomal mouse β-globin major gene: homologies in capping, splicing and poly A sites. Cell 15:1125–1132.
32. **Lacy, E., and T. Maniatis.** 1980. The nucleotide sequence of a rabbit β-globin pseudogene. Cell 21:545–553.
33. **Little, P., R. Treisman, L. Beirut, B. Seed, and T. Maniatis.** 1983. Plasmid vectors for the rapid isolation and transcriptional analysis of human β-globin gene alleles. Mol. Biol. Med. 1:473–488.
34. **Lowy, I., A. Pellicer, J. F. Jackson, G.-K. Sim, S. Silverstein, and R. Axel.** 1980. Isolation of transforming DNA: cloning the hamster aprt gene. Cell 22:817–823.
35. **Maniatis, T., E. F. Fritsch, and J. Sambrook (ed.).** 1982. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Marks, P. A., and R. A. Rifkind.** 1978. Erythroleukemic differentiation. Annu. Rev. Biochem. 47:419–426.
37. **Mathis, D., and P. Chambon.** 1981. The SV40 early region TATA box is required for accurate in vitro initiation of transcription. Nature (London) 290:310–315.
38. **Mathis, D. J., R. Elkaim, C. Kedinger, P. Sassone-Corsi, and P. Chambon.** 1981. Specific in vitro initiation of transcription on the adenovirus type 2 early and late EII transcription units. Proc. Natl. Acad. Sci. U.S.A. 78:7383–7387.
39. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. 74:560–564.
40. **McKnight, S. L.** 1982. Functional relationships between transcriptional control signals of the thymidine kinase gene of herpes simplex virus. Cell 31:355–365.
41. **McKnight, S. L., E. R. Gavis, R. Kingsbury, and R. Axel.** 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell 25:385–398.
42. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science 217:316–324.
43. **Mellon, P. M., V. Parker, Y. Gluzman, and T. Maniatis.** 1981. Identification of DNA sequences required for transcription of the human α1-globin gene using a new SV40 host-vector system. Cell 27:279–288.
44. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. 12:7035–7056.
45. **Myers, R. M., D. C. Rio, A. K. Robbins, and R. Tjian.** 1981. SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell 25:373–384.
46. **Neuberger, M. S.** 1983. Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. EMBO J. 2:1373–1378.
47. **Orkin, S. H., H. H. Kazazian, Jr., S. E. Antonarakis, S. C. Goff, C. D. Boehm, J. P. Sexton, P. J. Waber, and P. J. V. Giardina.** 1982. Linkage of β-thalassaemia mutations and β-globin gene polymorphisms with DNA polymorphisms in human β-globin gene cluster. Nature (London) 296:627–631.
48. **Orkin, S. H., P. J. Sexton, T. Cheng, S. C. Goff, P. J. C. Giardina, J. I. Lee, and H. H. Kazazian, Jr.** 1983. ATA box transcription mutation in β-thalassemia. Nucleic Acids Res. 11:4727–4734.
49. **Orkin, S. H., and P. S. Swerdlow.** 1977. Globin RNA synthesis in vitro by isolated erythroleukemia cell nuclei: direct evidence for increased transcription during erythroid differentiation. Proc. Natl. Acad. Sci. U.S.A. 74:2475–2479.
50. **Osborne, T. F., D. N. Arvidson, E. S. Tyau, M. Dunsworth-Browne, and A. J. Berk.** 1984. Transcription control region within the protein-coding portion of adenovirus E1a genes. Mol. Cell. Biol. 4:1293–1305.
51. **Osborne, T. F., R. B. Gaynor, and A. J. Berk.** 1982. The TATA homology and the mRNA 5' untranslated sequence are not required for expression of essential adenovirus E1A functions. Cell 29:139–148.
52. **Parker, C. S., and J. Topol.** 1984. A drosophila RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. Cell 36:357–369.
53. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell 33:741–748.
54. **Sandri-Goldin, R. M., A. L. Goldin, M. Levine, and J. C. Glorioso.** 1981. High-frequency transfer of cloned herpes simplex virus type 1 sequences to mammalian cells by protoplast fusion. Mol. Cell. Biol. 1:743–752.
55. **Schaffner, W.** 1980. Direct transfer of cloned genes from bacteria to mammalian cells. Proc. Natl. Acad. Sci. U.S.A. 77:2163–2167.
56. **Seed, B.** 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. Nucleic Acids Res. 11:2427–2443.
57. **Treisman, R., S. H. Orkin, and T. Maniatis.** 1983. Specific transcription and RNA splicing defects in five cloned β-thalassemia genes. Nature (London) 302:591–596.
58. **Treisman, R., N. J. Proudfoot, M. Shander, and T. Maniatis.** 1982. A single-base change at a splice site in a β-thalassemia gene causes abnormal RNA splicing. Cell 29:903–911.
59. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15S β-globin mRNA precursor and 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. 6:1175–1192.
60. **Wright, S., E. deBoer, F. G. Grosveld, and R. A. Flavell.** 1983. Regulated expression of the human β-globin gene family in murine erythroleukemia cells. 305:333–336.
61. **Wright, S., E. deBoer, A. Rosenthal, R. A. Flavell, and F. Grosveld.** 1984. DNA sequences required for regulated expression of β-globin genes in murine erythroleukemia cells. Cell 38:265–273.
62. **Zinn, K., D. DiMaio, and T. Maniatis.** 1983. Identification of two distinct regulatory regions adjacent to the human β-interferon gene. Cell 34:865–879.