MOLECULAR AND CELLULAR BIOLOGY, Dec. 1989, p. 5340–5349
0270-7306/89/125340-10$02.00/0

Vol. 9, No. 12

A *cis*-Acting Element in the Promoter Region of the Murine c-*myc* Gene Is Necessary for Transcriptional Block

HARVEY MILLER, CLAUDE ASSELIN,[2] DANIEL DUFORT,[1] JIAN-QING YANG,[2] KALPANA GUPTA,[1] KENNETH B. MARCU,[2] AND ALAIN NEPVEU[1]*

Ludwig Institute for Cancer Research, Montreal Branch, 687 Pine Avenue West, Montreal, Canada H3A 1A1,[1] and Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-5215[2]

Received 9 June 1989/Accepted 18 August 1989

A block to elongation of transcription has been shown to occur within the first exon of the human and murine c-*myc* genes. The extent of this block was found to vary with the physiological state of cells, indicating that modulation of the transcriptional block can serve to control the expression of this gene. To determine which sequences are required in *cis* for the transcriptional block, we generated a series of constructs containing various portions of murine c-*myc* 5′-flanking and exon 1 sequences. We established populations of HeLa and CV-1 cells stably transfected with these constructs. The transcription start sites were determined by S1 nuclease mapping analysis, and the extent of transcriptional block was measured by nuclear run-on transcription assays. Our results demonstrate that at least two *cis*-acting elements are necessary for the transcriptional block. A 3′ element was found to be located in the region where transcription stopped and showed features reminiscent of some termination sites found in procaryotes. A 5′ element was positioned between the P_1 and P_2 transcription start sites. We recently demonstrated that two nuclear factor-binding sites reside between P_1 and P_2 (C. Asselin, A. Nepveu, and K. B. Marcu, Oncogene 4:549–558, 1989). Removal of the more 3′ binding site abolished the transcriptional block.

Recent studies suggest that the expression of several eucaryotic genes can be regulated at the level of transcription elongation. Nuclear run-on transcription assays revealed that the density of RNA polymerase was often higher in the 5′ portion than in the 3′ portion of a gene, indicating that elongation of transcription in vivo can be blocked within a gene. Intragenic transcriptional block has successively been demonstrated by run-on assays for the hen and murine β-globin genes, simian virus 40 (SV40) late transcription unit, murine immunoglobulin μ heavy-chain gene, human and murine c-*myc* genes, murine c-*myb* and hamster c-*fos* genes, human immunodeficiency virus type 1 (HIV-1) long terminal repeat, human L-*myc* gene, and minute virus of mice (4–6, 9, 13–15, 19–21, 24, 28, 31, 36, 42, 45, 47, 48, 51). In addition, intragenic transcriptional block has been revealed by in vitro transcription studies in the adenovirus major late transcription unit, the *Drosophila hsp70* gene, and the human histone 3.3 and c-*myc* genes (11, 16, 26, 29, 41). Since relatively few genes have been assayed in this manner so far, it is reasonable to assume that many more genes will eventually be added to this list.

The extent of transcriptional block within a given gene is not necessarily constant but can vary with the physiological state of the cells, indicating that modulation of gene expression can be exerted at the level of transcription elongation. For example, enhanced transcriptional block was observed within the human c-*myc* gene upon differentiation of HL60 cells (5, 10, 33). Similarly, an enhanced transcriptional block was shown to contribute to the decrease in expression of the c-*myc* and c-*myb* genes when mouse erythroleukemia cells were chemically induced to differentiate (37, 49).

The transcriptional block has been mapped close to the exon 1-intron 1 junction within the human c-*myc* gene. In the *Xenopus* oocyte injection system, RNA polymerase II was shown to terminate transcription at two stretches of T's located in the vicinity and on either side of the exon 1-intron 1 junction (6). These results confirmed the mapping obtained by nuclear run-on analysis (5, 6, 9). In vitro transcription studies using partially purified RNA polymerase II also identified sites of transcription termination and pausing in the same region (26). For the murine c-*myc* gene, nuclear run-on transcription studies suggest that the site of blockage is located somewhat upstream of the exon 1-intron 1 junction (35, 51). Indeed, in *Xenopus* oocytes, transcription of the murine gene was found to stop at a stretch of five T's located 112 base pairs (bp) upstream of intron 1 (6).

Sequences located in the 3′ half of exon 1 have been implicated in transcriptional arrest in both the human and murine c-*myc* genes. A fragment of 95 bp from the 3′ half of the human c-*myc* exon 1 was shown to confer transcriptional blockage when inserted downstream of the herpes simplex virus *tk* promoter or the Rous sarcoma virus long terminal repeat (6). However, no block was observed when the same fragment was linked to the adenovirus major late promoter, suggesting that sequences present at the 5′ end of the RNA or in the promoter itself could also play at role (6). A recent report revealed that a 180-bp fragment from the 3′ half of the murine c-*myc* first exon caused blockage of transcription when inserted within the human α-globin gene (51).

As a first step to study the modulation of transcription elongation in mammalian cells, we identified the sequences of the murine c-*myc* gene necessary for the transcriptional block. Our results demonstrate that two portions of the murine c-*myc* exon 1 are necessary for the block: the promoter region between the P_1 and P_2 start sites and the region where transcription stops. We propose that these two regions respectively contribute a 5′ and a 3′ element, both of which are required for the transcriptional block.

MATERIALS AND METHODS

Cell culture and transfection. HeLa, CV-1, and COS cells were grown in Dulbecco modified Eagle medium supple-

* Corresponding author.

mented with 10% calf serum. Plasmid DNA was introduced by the calcium phosphate precipitation technique (50). Typically, 0.2 μg of pSV2Neo and 10 μg of c-*myc* plasmid were added to 10^6 cells in a 100-mm-diameter plate. Transfections were performed in duplicate, and 48 h later cells were split 1:6 in selective medium containing 1 mg of G418 per ml. Two weeks later, the resistant colonies were pooled. We usually obtained between 500 and 1,000 colonies in HeLa cells and 100 to 200 in CV-1 or COS cells. Constructs 1 to 5 and 9 to 11 were tested in both HeLa and CV-1 cells, and others were tested in HeLa (7, 8, 13–15), CV-1 (6), and COS (12) cells.

Nuclear run-on assays. Cells were scraped in phosphate-buffered saline with 1 mM EDTA, centrifuged for 5 min at 800 × *g*, and lysed in 2 ml of buffer A (46) adjusted to 0.1% Nonidet P-40. Lysed cells were centrifuged for 5 min at 1,500 rpm. The supernatant was saved for cytoplasmic RNA purification. The nuclei were washed once in nuclei storage buffer (46) and suspended in the same buffer containing 100 U of RNasin (Promega Biotec) per ml. The in vitro elongation reactions were carried out at 26°C for 30 min in a cocktail containing 100 mM Tris hydrochloride (pH 7.9), 50 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 0.4 mM EDTA, 300 mM $(NH_4)_2SO_4$, 0.1 mM phenylmethylsulfonyl fluoride, 1.2 mM dithiothreitol, 1 mM each ATP, CTP, and GTP, 150 μCi of [α-^{32}P]UTP (800 Ci/mmol), 10 mm creatine phosphate, 20 U of RNasin (Promega Biotec) per ml, 29% glycerol, and about 2 × 10^7 nuclei. Purification of the labeled nascent transcripts, hybridization, and washing were done as previously described (36, 37).

RNA isolation. The cytoplasmic fraction obtained after cell lysis (see above) was mixed with 4 volumes of guanidine thiocyanate, and RNA was prepared by the guanidine thiocyanate-hot phenol method essentially as described by Feramisco et al. (12).

S1 nuclease protection analysis. S1 nuclease protection analysis was performed essentially as described previously (52). A 50-μg sample of cytoplasmic RNA was annealed to 2 × 10^5 cpm of end-labeled probe or 5 × 10^5 cpm of uniformly labeled probe at 56°C for 8 h in 80% formamide–0.4 M NaCl–0.4 M piperazine-*N*,*N*-bis(2-ethanesulfonic acid) (PIPES) (pH 6.4)–1 mM EDTA. RNA-DNA hybrids were digested with 500 U of S1 nuclease (Bethesda Research Laboratories, Inc.) per ml at 25°C for 45 min, followed by a 15-min incubation at 37°C.

In vitro transcription. Whole-cell extracts were prepared from HeLa cells by the method of Manley et al. (30). The DNA template was previously linearized with *Eco*RI. The in vitro reactions were carried out at 30°C for 1 h in a total volume of 25 μl that included 1 μg of DNA template, 15 μl of HeLa whole-cell extract (8 mg/ml), 0.14 mM EDTA, 1 mM creatine phosphate, and 0.5 mM each ATP, CTP, GTP, and UTP. At the end of the reaction, 1 μl of RNase-free DNase (Bethesda Research Laboratories) was added, and the tubes were incubated at 30°C for 10 min. Transcripts were purified as described by Manley et al. (30) except that they were suspended in 30 μl of H_2O after the first ethanol precipitation. Half of the RNA was used to map the 5′ ends with a *Sma*I-*Not*I c-*myc* DNA probe end labeled at the *Not*I site. The other half served to map the 3′ ends, using as a probe the *Not*I-*Nco*I fragment from the template plasmid end-labeled at the *Not*I site with the Klenow fragment of *Escherichia coli* polymerase 1 (32). Hybridization was done at 54°C, and S1 nuclease digestion was performed as described above.

Plasmid construction. All constructs except construct 12 originate from the construct pCAT(ΔEP), which is a derivative of pSV2CAT without the SV40 enhancer and promoter (17, 52). Constructs 1 and 2 have been described by Yang et al. (52); constructs 6, 7, 9, 10, and 11 have been described by Asselin et al. (1). Construct 3, deleted of bases 335 to 516, was made by cutting construct 1 with *Not*I and *Xho*I and resealing the ends with *Xho*I linkers. Construct 4 was made by inserting the *Bam*HI-*Xho*I fragment of probe A3 (Fig. 1) into the *Xho*I site of construct 3. Construct 5 was made as follows: construct 2 was cleaved with *Xho*I, treated with BAL 31 nuclease, and recircularized with *Xho*I linkers. The *Xho*I-*Eco*RI fragment from this plasmid was then replaced with the corresponding fragment from construct 2, effectively restoring the integrity of the sequences downstream of *Xho*I. This resulted in a deletion between nucleotide +438 (relative to P1) and the *Xho*I site. Probes A1, A2, A3, and A4 (Fig. 1) were generated as follows: the replicative form of an M13 DNA clone containing the *Bam*HI-*Bgl*II fragment from c-*myc* exon 1 (probe A; Fig. 2c) was cleaved with *Bam*HI, treated with BAL 31 nuclease, and recircularized with *Bam*HI linkers. The deletions were mapped by DNA sequencing. Construct 8 was made by inserting the *Bam*HI-*Bgl*II fragment from c-*myc* exon 1 into the *Bgl*II site of pSVCAT(ΔEP). For construct 13, a 2-kilobase-pair (kb) *Hin*dIII from plasmid pH2CAT was inserted into the corresponding site of construct 8 (27). The *Hin*dIII site distal to the c-*myc* sequences was then removed by partial digestion, filling in of the ends with Klenow polymerase, and blunt-end ligation. Constructs 14 and 15 were obtained by inserting the *Hae*III-*Hin*dIII fragment (from −6 to +138 relative to P_1) into the unique *Hin*dIII site of this plasmid.

RESULTS

Mapping of the transcriptional block. A block to elongation of transcription has previously been identified close to the junction of exon 1 and intron 1 (36, 51). S1 nuclease protection assays using preparations of total cellular or nuclear RNA have failed to reveal short transcripts that would be generated by cleavage at the site(s) of the block. We assume that these transcripts, if generated, are rapidly degraded in vivo. We therefore used two different approaches to more precisely map the site(s) of transcriptional block.

First, we performed nuclear run-on transcription assays, using nuclei isolated from the $54cl_2$ cell line and short single-stranded DNA probes derived from exon 1 (Fig. 1a). The c-*myc* gene is amplified about 20-fold in the $54cl_2$ cell line and shows strong transcriptional blockage (35, 36). To determine the rate of transcription within exon 1, we used M13 single-stranded DNA clones that contained various lengths of exon 1 sequences. We reasoned that probes which did not contain sequences upstream of the site(s) of transcriptional block would give weaker signals. For comparison, we included a probe containing 3.4 kb of sequences downstream of exon 1. In agreement with our previous results, no signals were detected in the presence of 1 μg of α-amanitin per ml (data not shown). This control verified that the signals detected in our run-on assays were due to transcription by RNA polymerase II. The signals obtained with probes A1, A2, and A3 were of greater intensity than that of probe C, but the signal obtained with probe A4 was much weaker. Since probe C is much larger than any probe A, the results suggest that probes A1, A2, and A3 contain sequences upstream of the site(s) of blockage, whereas probe A4 contains few or none of these sequences. This allowed us to map the site of blockage somewhere between

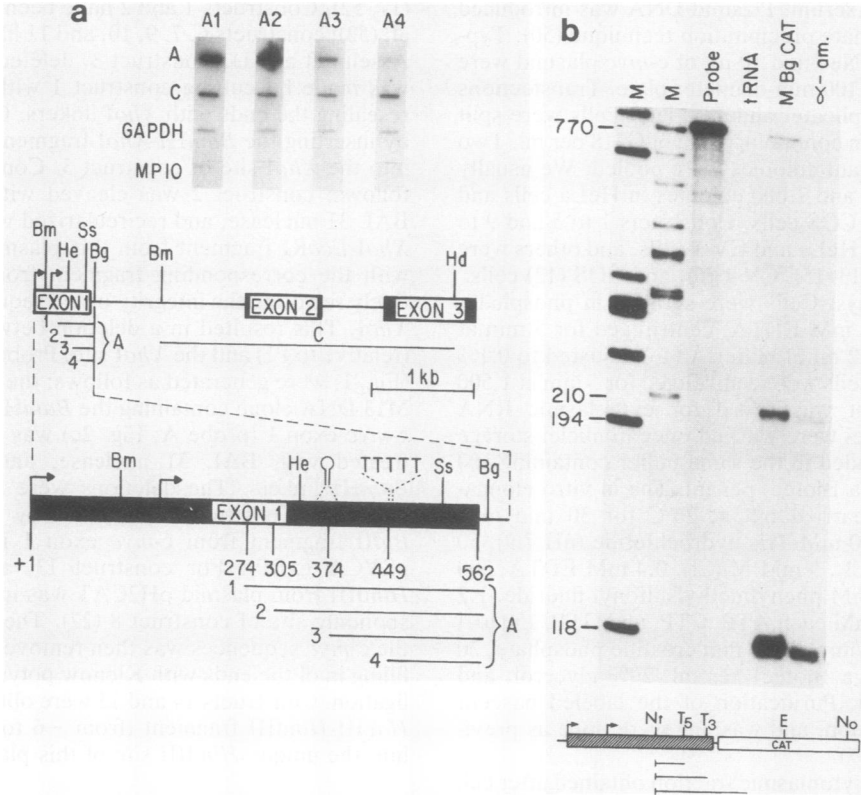

FIG. 1. Mapping of the transcriptional block. (a) Nuclear run-on transcription analysis of the c-*myc* gene in 54cl$_2$. Nascent transcripts were labeled in isolated nuclei and hybridized to single-stranded M13 DNA probes. The c-*myc* sequences present in the A and C probes are shown in the map. The left boundary of each A probe is indicated in nucleotides relative to the P_1 start site (+1). Uridine contents of the regions covered by the c-*myc* probes are as follows: 62 in A1, 56 in A2, 45 in A3, 36 in A4, and 584 in C. The GAPDH probe is the *Pst*I fragment of the rat glyceraldehyde-3-phosphate dehydrogenase cDNA clone pRGAPDH13 (39). The MP10 probe is the single-stranded DNA of the M13 vector mp10. Arrows represent the P_1 and P_2 start sites. Also indicated are the position of one pair of inverted repeats and the stretch of five T's. Restriction sites: Bm, *Bam*HI; *He*, *Hae*III; *Ss*, *Sst*I; *Bg*, *Bgl*II. (b) S1 nuclease mapping of transcription termination sites in vitro. Construct pMBgCAT (Fig. 2c) was cleaved with *Eco*RI and incubated with HeLa whole-cell extracts for 60 min at 30°C with or without α-amanitin (2 μg/ml). RNA was purified, and before digestion by S1 nuclease, half of it was hybridized to a *Not*I-*Nco*I DNA fragment end labeled at the *Not*I site. The fragments protected by terminated transcripts are indicated below the map. M, ϕX174 DNA digested with *Hae*III (first lane) or *Hin*cII (second lane); α-am, in vitro transcription performed in the presence of α-amanitin (2 μg/ml). Symbols for map: ▨, c-*myc* exon 1; ▭, CAT gene. The line between the two boxes represents 9 bp of c-*myc* intron 1 sequences. Abbreviations: Nt, *Not*I; E, *Eco*RI; No, *Nco*I; T_5 and T_3, stretches of five and three T's.

the 5′ ends of probes A3 and A4, about 100 to 200 bp upstream of the exon 1-intron 1 junction.

To map more precisely the site(s) of the transcriptional block, we then performed in vitro transcription with whole-cell extracts (30). As a template, we used plasmid pMBgCAT linearized at the *Eco*RI site within the chloramphenicol acetyltransferase (CAT) gene. This plasmid contains c-*myc* exon 1 and 1,141 bp of 5′-flanking sequences (Fig. 3). The 3′ ends of the in vitro-transcribed c-*myc* transcripts were mapped by S1 nuclease protection assay, using an end-labeled probe (Fig. 1b). We observe one major band of 115 to 120 bases, which was resolved into several discrete bands upon shorter exposure (Fig. 1b and data not shown). The sizes of these bands were consistent with termination of transcription occurring in vitro within or near a stretch of five T's on the sense DNA strand (bases 452 to 456 relative to P_1). This was found to be the unique site of termination in the *Xenopus* oocyte system (6). We also found additional minor bands corresponding to transcripts terminated further downstream. A band of 200 bases identified a second site of termination within or near another T-rich sequence (bases 539 to 546 relative to P_1). Comparison of the sequences of the murine and human c-*myc* genes showed that this latter site corresponded to the T1 termination-pausing site identified in the human c-*myc* gene (6, 26). A band of 330 bases identified a third site of termination near a T-rich sequence in the CAT gene. These short c-*myc* transcripts appeared to be generated by premature termination in vitro, since they could not be chased into longer products upon prolonged incubation (data not shown). We cannot exclude the possibility that termination of transcription occurred downstream and was followed by rapid processing of the primary transcripts to generate the observed 3′ ends. A certain proportion of the transcripts detected in the in vitro transcription assay were in fact generated by RNA polymerase III. Indeed, performing the same assay in the presence of α-amanitin reduced the signal obtained by fivefold. This result is in accordance with those other in vitro studies performed on the human c-*myc* gene (8, 26). It should be stressed, however, that we have never detected RNA polymerase III transcription of the c-*myc* gene in nuclear run-on transcription assays. These results suggest that the specificity of c-*myc* promoters for different RNA polymerases is not as stringent in vitro.

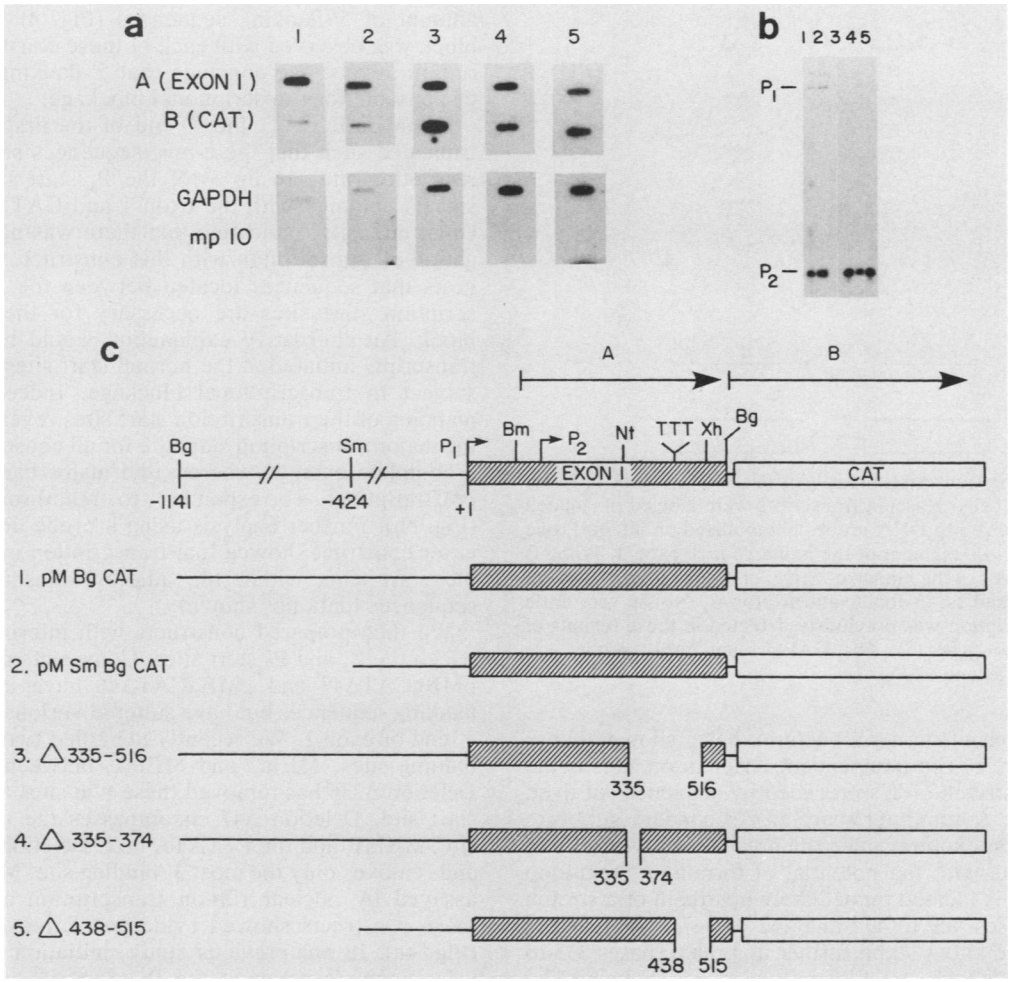

FIG. 2. Transcription analysis of constructs with deletions in the 3′ half of c-*myc* exon 1. (a) Run-on transcription assays. The constructs were stably transfected into HeLa cells. Nuclei were isolated and used in run-on assays. Labeled nascent transcripts were hybridized to DNA probes immobilized on filters. Lane numbers correspond to construct numbers. Probe A (exon 1) is a single-stranded M13 DNA containing the *Bam*HI-*Bgl*II fragment of c-*myc* exon 1 (uridine content, 95). Probe B (CAT) is the *Bgl*II-*Nco*I fragment of the CAT gene (uridine content, 156). (b) S1 nuclease mapping of the transcription start sites. Before digestion by S1 nuclease, cytoplasmic RNA isolated from the same cells was hybridized to a *Sam*I-*Not*I fragment end labeled at the *Not*I site. Lane numbers correspond to construct numbers. Positions of the fragments protected by P_1- and P_2-initiated transcripts are shown. (c) Partial maps of the constructs. The line to the left of exon 1 represents c-*myc* 5′-flanking sequences. The boundaries of the deletions are indicated in nucleotides relative to P_1 (+1). Abbreviations: Bg, *Bgl*II; Sm, *Sma*I; Bm, *Bam*HI; Nt, *Not*I; Xh, *Xho*I; TTT, stretch of T's.

In conclusion, using two different approaches, we have mapped the sites of transcriptional block to the same region of exon 1.

Determination of the sequences necessary for the transcriptional block: differential transcriptional regulation of the exogenous and endogenous c-*myc* genes. To map the essential *cis*-acting sequences for transcriptional blockage, we prepared a series of hybrid constructs in which various c-*myc* sequences were attached to either the CAT or rabbit β-globin gene. The c-*myc* sequences present in each construct are shown in Fig. 2c, 4b, and 6b. The block to elongation of transcription was assayed by nuclear run-on transcription assays, using nuclei isolated from populations of stably transfected cells. No signal was obtained for exon 1 transcription when the experiment was done with nuclei isolated from nontransfected cells (Fig. 4a, lane 0). Thus, with the hybridization and washing conditions we used, there was no cross-hybridization between the murine and the human or monkey c-*myc* exon 1 sequences.

The construct pMBgCAT contains the entire c-*myc* exon 1, 11 bp of intron 1 sequences, and 1,141 bp of 5′-flanking sequences (Fig. 2c). When assayed in HeLa and in CV-1 cells, this construct gave a stronger signal with the exon 1 probe (A) than with the CAT probe (B) (Fig. 2a and 4a). This result indicated that a block to elongation of transcription occurred when this construct was used as a template for transcription. Two conclusions can be derived from this result. First, this construct contains all of the *cis*-acting elements necessary for the transcriptional block to occur. Second, the transcriptional machinery in HeLa and CV-1 cells can recognize the *cis*-acting elements present in the murine c-*myc* sequences and modulate elongation of transcription in a fashion similar to that observed in murine cell lines. However, we found no evidence of transcriptional block within the endogenous c-*myc* gene of either HeLa or CV-1 cells (Fig. 3).

The 3′ half of the murine c-*myc* exon 1 contains an element essential for blockage. We prepared constructs with various deletions in the 3′ half of the c-*myc* first exon and assessed their degree of transcriptional blockage (Fig. 2). Studies of

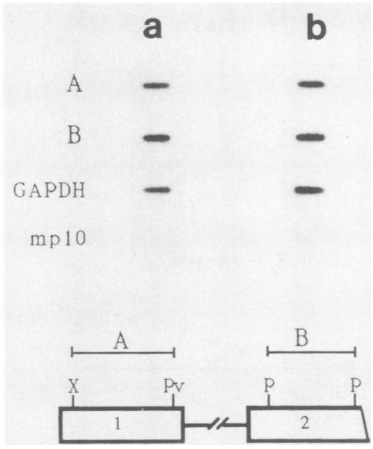

FIG. 3. Run-on transcription analysis of the c-*myc* gene in HeLa (a) and CV-1 (b) cells. Nascent transcripts were labeled in isolated nuclei and hybridized to DNA probes immobilized on filters. Probe A is the *Xho*I-*Pvu*II fragment of the human c-*myc* exon 1. Probe B is the *Pst*I fragment of the human c-*myc* exon 2. The uridine content of the sense strand is 86 for A and 67 for B. No or very little antisense transcription was previously detected in these regions of the human c-*myc* gene (5). The GAPDH and mp10 probes are described in the legend to Fig. 1.

transcription termination in bacteria have shown that a constant feature of rho-independent termination sites is the presence of a stretch of T's preceded by sequences of dyad symmetry (40). A computer search for secondary structure in the c-*myc* first exon revealed the presence of two pairs of inverted repeats with the potential of forming a stem-loop structure. One is located immediately upstream of a stretch of five T's (bases 429 to 434 and 442 to 447), whereas the other is situated about 90 bp further upstream (bases 333 to 339 and 353 to 359). The position of the latter is shown in Fig. 1a. Upon deletion of the more 5′ inverted repeats, we observed complete readthrough transcription (Fig. 2a, lane 4). Similarly, no transcriptional block was detected when the second inverted repeat and the adjacent run of five T's were deleted (Fig. 2a, lane 5). We conclude that this entire region contains sequences required for the transcriptional block.

We used these constructs to determine whether the block has a similar effect on the elongation of primary transcripts initiated at the P_1 and P_2 start sites. The steady-state level of P_2-initiated transcripts is generally higher than that of P_1-initiated transcripts (25, 35). This may reflect the relative strength of the two promoters; alternatively, the lower level of P_1 transcripts may result from a stronger block to elongation of primary transcripts initiated at this site. If this were the case, one would expect that abrogation of the transcriptional block would effectively increase the P_1/P_2 ratio. To investigate this possibility, we performed S1 nuclease protection assays on cytoplasmic RNA isolated from cells transfected with these constructs (Fig. 2b). We found that the quantity of P_1-initiated transcripts was not increased in the absence of a transcriptional block and that the relative abundance of P_1- and P_2-initiated transcripts was not altered. Thus, the P_1/P_2 ratio must normally reflect the relative strengths of the two promoters.

Are sequences 5′ of either the c-*myc* P_1 or P_2 start sites required for blockage? We verified the importance of DNA sequences 5′ of the P_1 start site for blockage. Constructs pMSmBgCAT, pMBgCATΔ 96, and pMHeBgCAT contain all of exon 1 plus 11 bp of intron 1 sequences and variable amount of 5′-flanking sequences (Fig. 4). Transcriptional block was observed with each of these constructs in nuclear run-on assays. We conclude that 5′-flanking sequences are dispensable for transcriptional blockage.

In pMBmBgCAT, the 5′ end of the first exon has been truncated such that the c-*myc* sequences start at a *Bam*HI restriction site 116 bp 3′ of the P_1 start site. The run-on signals obtained with the exon 1 and CAT probes were of equal intensity, indicating that there was no block to elongation of transcription with this construct. This result suggests that sequences located between the P_1 and P_2 transcription start sites are necessary for the transcriptional block. An alternative explanation would be that only the transcripts initiated at the normal start sites, P_1 and P_2, are subject to transcriptional blockage. Indeed, S1 nuclease mapping of the transcription start sites revealed that P_2 was the major transcription start site for all constructs containing a complete exon 1, whereas the major band observed for pMBmBgCAT corresponded to readthrough transcripts (Fig. 5a). Further analysis using a probe derived from this latter construct showed that transcription initiated at multiple start sites within the adjacent plasmid and cellular sequences (data not shown).

We then prepared constructs with internal deletions between the P_1 and P_2 start sites. Constructs pMBgCATΔ236, pMBgCATΔ47 and pMBgCATΔ46 have intact c-*myc* 5′-flanking sequences but have suffered various deletions at the 5′ end of exon 1. We recently identified two nuclear factor-binding sites, MEIa2 and MEIa1, between P_1 and P_2 (1). Deletion Δ236 has removed these two sites as well as the P_2 start site. Deletion Δ47 encompasses the more 3′ binding site, MEIa1, and the P_2 TATA box. Deletion Δ46 is of 16 bp and removes only the most 3′ binding site, MEIa1 (1). When assayed by nuclear run-on transcription assays, none of these constructs showed evidence of transcriptional block (Fig. 4a). In our previous study, initiation of transcription was shown to occur at the P_1 start site with these three constructs and, in the case of pMBgCATΔ46, to also occur at the P_2 start site (1). Indeed, an S1 nuclease protection assay with an end-labeled probe demonstrated that this deletion of 16 bp did not abolish initiation at the P_2 start site, although it may have reduced the level of utilization, as shown by the relative intensities of the bands corresponding to the P_1- and P_2-initiated transcripts (Fig. 5b; 1).

Two conclusions can be derived from these results. First, initiation of transcription at the P_1 and P_2 start sites does not ensure a transcriptional block. Second, a deletion of 16 bp at the 5′ end of the c-*myc* first exon is sufficient to abrogate the block to elongation of transcription. This deletion removes one of two nuclear factor-binding sites present between P_1 and P_2.

The 3′ half of the c-*myc* exon 1 does not confer transcriptional block when placed downstream of two heterologous promoters. A fragment of the murine c-*myc* gene containing all of exon 1 except the first 138 bp was inserted downstream of two different promoters: the SV40 early and the *H-2*^{k} major histocompatibility class gene (MHC *H-2*^{k}) promoters. When tested by nuclear run-on transcription assay, these constructs showed no evidence of a transcriptional block (Fig. 6a). The transcription start sites of the SV40-*myc* construct have been determined previously (52) and are indicated in Fig. 6b. For the MHC *H-2*^{k}-*myc* construct, mapping of the transcription start sites is shown in Fig. 5b. Using a uniformly labeled probe derived from this construct, we observed one major band corresponding to transcripts initiated at the start site of the MHC *H-2*^{k} gene. These results

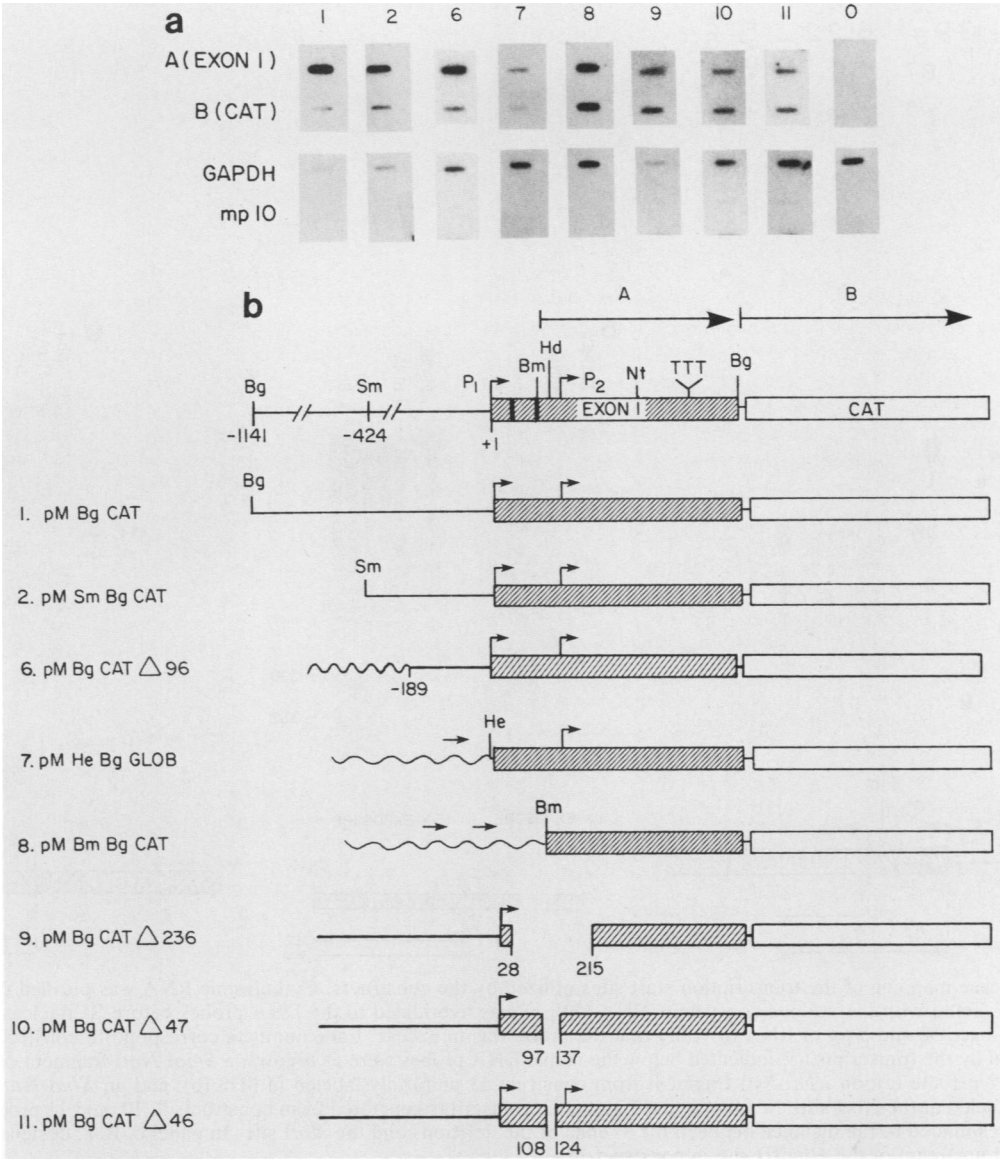

FIG. 4. (a) Run-on transcription analysis of constructs with deletions in the promoter region. The constructs were stably transfected into HeLa cells, CV-1 cells, or both. Nuclei were isolated and used in run-on assays. Labeled nascent transcripts were hybridized to DNA probes immobilized on filters. Lane numbers correspond to construct numbers. Lane 0, Untransfected HeLa cells. The probes are the same as in Fig. 2 except for construct 7, for which probe B is a 437-bp *Hin*dIII-*Bgl*II fragment of the rabbit β-globin cDNA (34). (b) Partial maps of the constructs. The line to the left of the exon 1 box represents c-*myc* 5′-flanking sequences; wavy lines represent vector or cellular sequences. Nucleotide numbers are relative to P_1 (+1). The two thick vertical bars within the exon 1 box represent the ME1a2 and ME1a1 protein-binding sites. Arrows above the maps indicate the major transcription start sites as determined by S1 nuclease analysis (Fig. 5 and data not shown). Arrows in constructs 7 and 8 represent cryptic start sites in the vector and cellular sequences. Abbreviations: Bg, *Bgl*II; Sm, *Sma*I; Bm, *Bam*HI; Hd, *Hin*dIII; Nt, *Not*I; TTT, stretch of T's.

demonstrate that a fragment containing most of the murine c-*myc* exon 1, including the sites of blockage, is insufficient to confer transcriptional blockage when inserted downstream of these heterologous promoters.

The 5′ end of the murine c-*myc* exon 1 contains an element required for transcriptional block. To further confirm that the 5′ end of the murine c-*myc* exon 1 contains a *cis*-acting element necessary for transcriptional blockage, we inserted a 144-bp fragment of this region into the MHC *H-2*^{k}-*myc* construct. This fragment extends from the *Hae*III restriction site (at −6 relative to P_1) to the *Hin*dIII site at +138. Insertion of this fragment in its sense orientation effectively restored the integrity of the c-*myc* exon 1 (MHC HeBgCAT; Fig. 6b). We also prepared a construct in which the same fragment was inserted in the opposite orientation (MHC-HdHeHdBgCAT; Fig. 6b). We found transcriptional block with MHC-HeBgCAT (Fig. 6a, lane 14) but not with MHC-HdHeHdBgCAT (Fig. 6a, lane 15). The sites of transcription initiation for MHC-HeBgCAT corresponded to the MHC *H-2*^{k} major site and the P_2 start site (Fig. 5c). Whether blockage affected transcription initiated at either site is not known. For MHC-HdHeHdBgCAT, three bands were visible: two corresponding to transcripts initiating close to P_2 and a higher-intensity band representing readthrough tran-

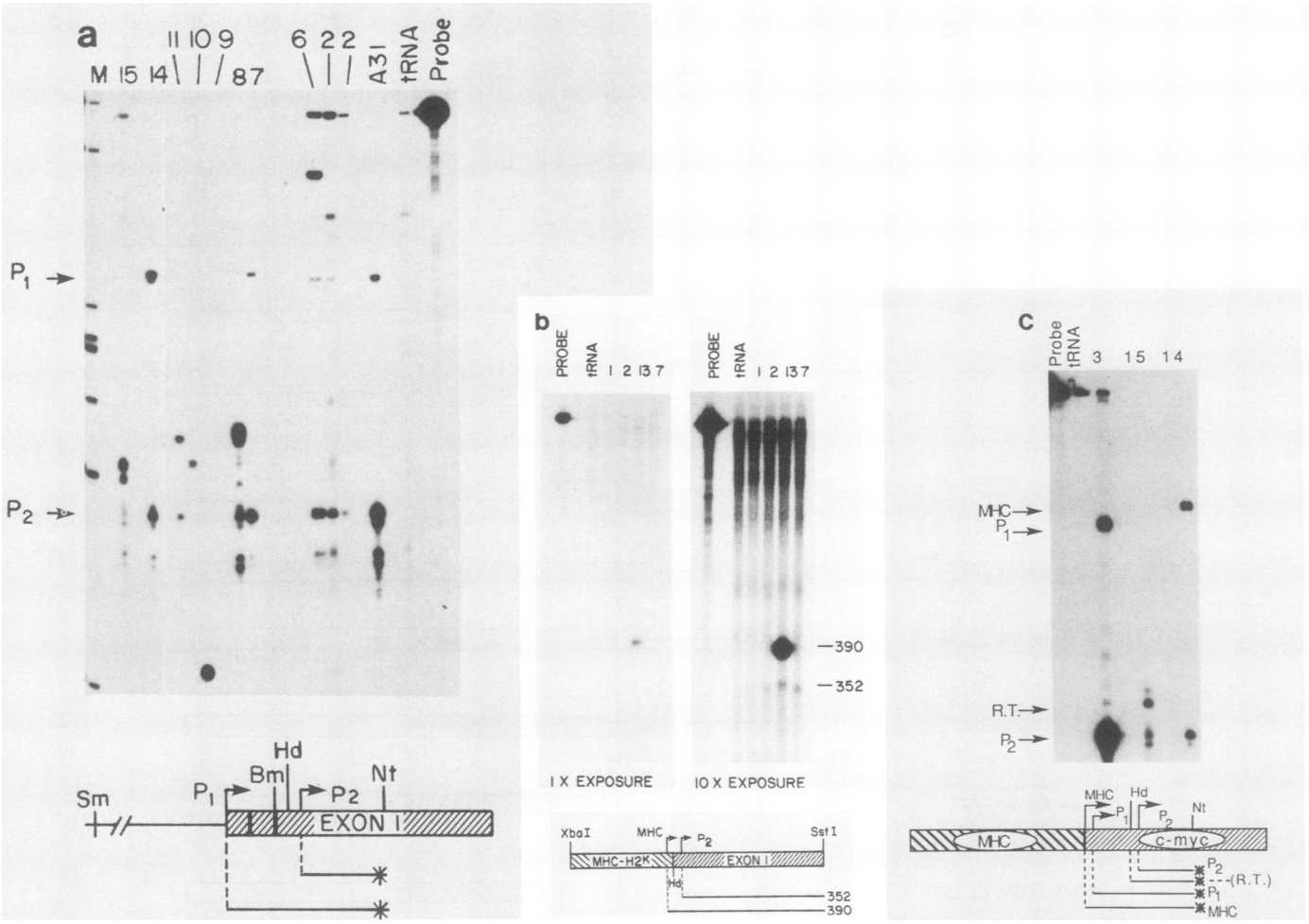

FIG. 5. S1 nuclease mapping of the transcription start sites utilized by the constructs. Cytoplasmic RNA was purified from populations of cells stably transfected with the various constructs. RNA (50 μg) was hybridized to the DNA probes before S1 nuclease digestion. For construct 2 (a), we used 50 and 5 μg of RNA to verify that the probe was in excess. Lane numbers correspond to construct numbers. The fragments protected by the transcripts are indicated below the maps. DNA probes were as follows: a *Sma*I-*Not*I fragment of the c-*myc* gene end labeled at the *Not*I site (a); an *Xba*I-*Sst*I fragment from construct 13 uniformly labeled in M13 (b); and an *Xba*I-*Not*I fragment from construct 14 end labeled at the *Not*I site. In panel a, the P_1-initiated transcripts generated from constructs 9, 10, and 11 protected fragments whose lengths corresponded to the distance between the 3′ ends of the deletions and the *Not*I site. In panel c, R.T. designates readthrough transcripts initiated upstream of the *Hin*dIII site in construct 15.

scripts that were initiated upstream of the c-*myc* sequences and probably correspond to the MHC *H-2*^{k} start site (Fig. 5c).

These findings demonstrate that the region between P_1 and P_2 contains a *cis*-acting element that is necessary for the transcriptional block and functions in an orientation-dependent fashion.

DISCUSSION

The transcriptional machinery present in HeLa and CV-1 cells can recognize the murine c-*myc* sequences that mediate the block. The *trans*-acting factors involved in this mode of regulation are therefore present in these cells, although transcription of the endogenous c-*myc* gene proceeds without arrest. Thus, the endogenous and exogenous c-*myc* genes are not subject to the same transcriptional regulation. We propose several testable hypotheses to explain this phenomenon. (i) The constructs used in this study may lack some regulatory elements required for readthrough transcription in HeLa and CV-1 cells. These elements may be specific to primates or, if conserved, must be located more than 1,141 bp upstream of P_1 or downstream of the first exon. (ii) The c-*myc* gene in HeLa and CV-1 cells may have suffered mutations in the sequences necessary for the block, as shown in some Burkitt lymphomas (7). Alternatively, the endogenous c-*myc* gene may have previously been imprinted or tagged in such a way that it could no longer be down-regulated by transcriptional block. A somewhat analogous situation has been described previously (2). (iii) The human and monkey c-*myc* promoters may be weaker than that of the murine gene such that transcription does not require to be down regulated by blockage. In this model, the transcriptional block would play the role of fine tuning control for transcription of the c-*myc* gene. We emphasize that this study does not permit a comparison of the promoter strengths of the various constructs used with those of the endogenous genes.

A block to elongation of transcription was observed with all constructs that contain the murine c-*myc* exon 1. We conclude that the sequences necessary in *cis* for the tran-

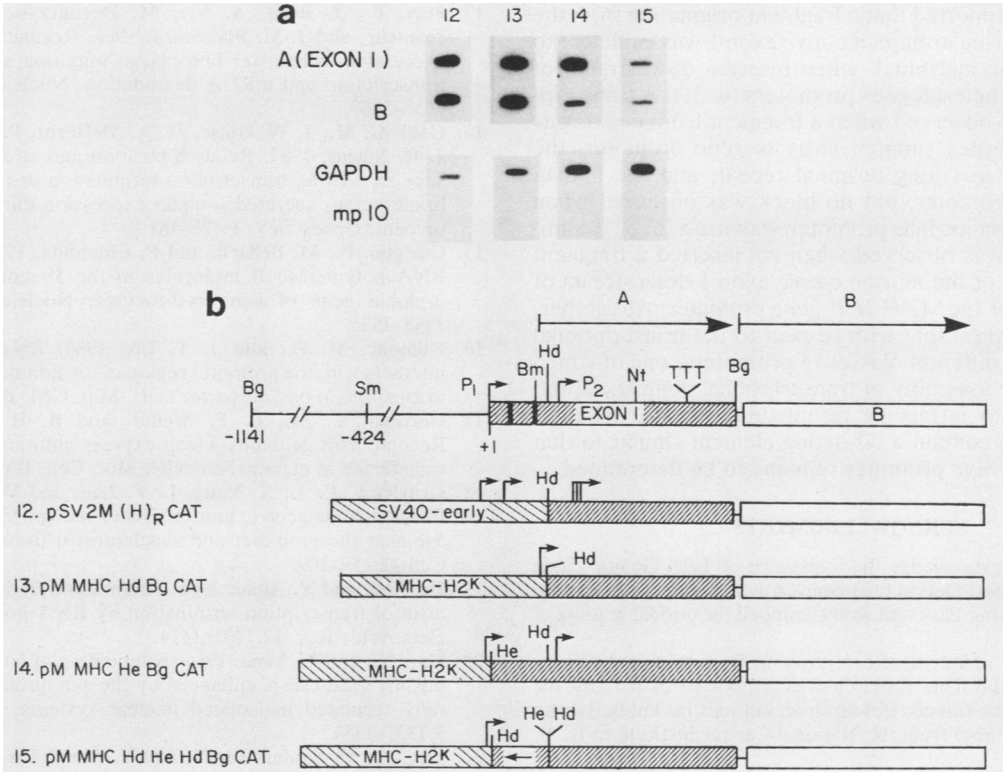

FIG. 6. (a) Run-on transcription analysis of constructs with heterologous promoters. The constructs were stably transfected into HeLa cells. Nuclei were isolated and used in run-on assays. Labeled nascent transcripts were hybridized to DNA probes immobilized on filters. Lane numbers correspond to construct numbers. The probes are the same as in Fig. 2. (b) Partial maps of the constructs. SV40-early, SV40 enhancer and early promoter; MHC-H2^k, 2 kb of 5′-flanking sequences from the MHC *H-2^k* gene. Arrows above the maps indicate the major transcription start sites as determined by S1 nuclease analysis (Fig. 5 and data not shown). The arrow within the c-*myc* box of construct 15 indicates that the *Hae*III-*Hin*dIII fragment has been inserted in the reverse orientation. Abbreviations: Bg, *Bgl*II; Sm, *Sma*I; Bm, *Bam*HI; Hd, *Hin*dIII; Nt, *Not*I; He, *Hae*III; TTT, stretch of T's.

scriptional block are contained entirely within this exon. Deletion analysis suggests that at least two *cis*-acting elements are necessary for the transcriptional block: a 5′ element located between the P_1 and P_2 start sites and a 3′ element close to the main site(s) of transcriptional block. Removal of either the 5′ or the 3′ element effectively abolishes the block.

Various deletions that removed the 3′ element resulted in abrogation of the block. This finding confirms previous results showing that a 180-bp fragment from this region can confer transcriptional blockage when inserted within an intron of the human α-globin gene (51). We demonstrate that the element present in this region is required for the transcriptional block in the context of the c-*myc* gene but, when inserted downstream of two heterologous promoters, is not sufficient in itself to confer transcriptional blockage. The block of elongation of transcription was restored when a fragment from the c-*myc* promoter that contained the 5′ element was inserted into one of these constructs. Correspondingly, the transcriptional block was abolished in all constructs with a deletion of this region between the P_1 and P_2 start sites. The smallest of these deletions removed 16 bp that coincided with a nuclear factor-binding site (1). Transcription initiation within this construct occurs at both P_1 and P_2 start sites (1; this study).

In conclusion, our data show that two regions of the murine c-*myc* exon 1 are involved in modulation of the transcriptional block: the promoter region between the P_1 and P_2 start sites and the region where transcription stops. The nature of the *cis*-acting elements present in these two regions and the exact relationship between these elements remains to be defined. On the basis of our results, we favor a model in which promoter sequences participate in modulation of the block.

The role of promoter sequences in the control of transcription elongation has been most extensively studied in the case of the early and late genes of the λ and λ-related bacteriophages (reviewed in reference 44). Expression of these genes depends on whether transcription stops or continues beyond several sites of termination. Elongation of transcription is in turn controlled by specific interactions ocurring in the promoter region between the polymerase and some termination (Nun) or antitermination (N and Q) factors (3, 18, 43). We suggest that similar interactions between RNA polymerase II and some transcriptional factors determine the fate of transcription elongation within the c-*myc* gene. A transcription termination factor would bind to some promoter sequences and then modify or become part of the transcriptional complex, allowing the polymerase to recognize the 3′ element as it proceeds along the 3′ half of exon 1. Alternatively, looping out of the DNA between the 5′ and 3′ elements would allow interaction between this factor and the site of blockage. Interestingly, the involvement of promoter sequences in the regulation of transcription termination has been reported for other RNA polymerase II-transcribed genes. The 3′-end formation of U1 and U2 small nuclear RNAs, which is intimately linked to transcription termination, also requires promoter elements (22, 23, 38).

Other groups reported that a fragment originating from the 3′ half of the murine or human c-*myc* exon 1 was sufficient to confer transcriptional block when inserted downstream of some but not all heterologous promoters (6, 51). A transcriptional block was observed when a fragment from c-*myc* was linked to the herpes simplex virus *tk* gene promoter, the Rous sarcoma virus long terminal repeat, and the human α-globin gene promoter, but no block was observed when the adenovirus major late promoter was used (6, 51). Similarly, no block was observed when we inserted a fragment containing most of the murine c-*myc* exon 1 downstream of the SV40 early or the MHC *H-2*^{k} gene promoter. Altogether, these results suggest that with respect to the transcriptional block, there are different classes of promoters, one of which can promote the assembly of transcriptional complexes that are responsive to intragenic termination signals. Whether these promoters contain a *cis*-acting element similar to that present in the c-*myc* promoter remains to be determined.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of Terri Genio, Linda Sapienza, and Robert Derval for manuscript and figure preparation. We also thank Morag Park and Rob Campbell for critical reading of the manuscript.

C. Asselin is a recipient of a fellowship from National Cancer Institute of Canada. This research was supported in part by the Ludwig Institute for Cancer Research and in part by Public Health Service grant CA36246 from the National Cancer Institute to K. B. Marcu.

LITERATURE CITED

1. **Asselin, C., A. Nepveu, and K. B. Marcu.** 1989. Molecular requirements for transcriptional initiation of the murine c-*myc* gene. Oncogene **4**:549–558.
2. **Atchison, M. L., and R. P. Perry.** 1987. The role of the κ enhancer and its binding factor NF-κB in the developmental regulation of κ gene transcription. Cell **48**:121–128.
3. **Barik, S., B. Ghosh, W. Whalen, D. Lazinski, and A. Das.** 1987. An antitermination protein engages the elongating transcription apparatus at a promoter-proximal recognition site. Cell **50**: 885–899.
4. **Bender, T. P., C. B. Thompson, and W. M. Kuehl.** 1987. Differential expression of c-*myb* mRNA in murine B lymphomas by a block to transcription elongation. Science **237**:1473–1476.
5. **Bentley, D. L., and M. Groudine.** 1986. A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. Nature (London) **321**:702–706.
6. **Bentley, D. L., and M. Groudine.** 1988. Sequence requirements for premature termination of transcription in the human c-*myc* gene. Cell **53**:245–256.
7. **Cesarman, E., R. Dalla-Favera, D. Bentley, and M. Groudine.** 1987. Mutations in the first exon are associated with altered transcription of c-*myc* in Burkitt lymphoma. Science **238**:1272–1275.
8. **Chung, J., D. J. Sussman, R. Zeller, and P. Leder.** 1987. The c-*myc* gene encodes superimposed RNA polymerase II and III promoters. Cell **51**:1001–1008.
9. **Eick, D., and G. W. Bornkamm.** 1986. Transcriptional arrest within the first exon is a fast control mechanism in c-*myc* gene expression. Nucleic Acids Res. **14**:8331–8346.
10. **Eick, D., R. Berger, A. Polack, and G. W. Bornkamm.** 1987. Transcription of c-*myc* in human mononuclear cells is regulated by an elongation block. Oncogene **2**:61–65.
11. **Evans, R., J. Weber, E. Ziff, and J. E. Darnell.** 1979. Premature termination during adenovirus transcription. Nature (London) **278**:367–370.
12. **Feramisco, J. R., J. E. Smart, K. Burridge, D. M. Helfman, and G. P. Thomas.** 1982. Co-existence of vinculin and a vinculin-like protein of higher molecular weight in smooth muscle. J. Biol. Chem. **257**:11024–11031.
13. **Fort, P., J. Rech, A. Vie, M. Piechaczyk, A. Bonnieu, P. Jeanteur, and J.-M. Blanchard.** 1987. Regulation of c-*fos* gene expression in hamster fibroblasts: initiation and elongation of transcription and mRNA degradation. Nucleic Acids Res. **15**: 5657–5667.
14. **Galli, G. M., J. W. Guise, M. A. McDevitt, P. W. Tucker, and J. R. Nevins.** 1987. Relative position and strengths of poly(A) sites as well as transcription termination are critical to membrane versus secreted μ-chain expression during B-cell development. Genes Dev. **1**:471–481.
15. **Gariglio, P., M. Bellard, and P. Chambon.** 1981. Clustering of RNA polymerase B molecules in the 5′ moiety of the adult β-globin gene of hen erythrocytes. Nucleic Acids Res. **9**: 2588–2595.
16. **Gilmour, M. Z., and J. T. Lis.** 1986. RNA polymerase II interacts with the promoter region of the noninduced *hsp70* gene in *Drosophila melanogaster* cells. Mol. Cell. Biol. **6**:3984–3989.
17. **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.
18. **Grayhack, E. J., X. Yang, L. F. Lau, and W. Roberts.** 1985. Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. Cell **42**:259–269.
19. **Hay, N., and Y. Aloni.** 1984. Attenuation in SV40 as a mechanism of transcription termination by RNA polymerase B. Nucleic Acids Res. **12**:1401–1414.
20. **Hay, N., and Y. Aloni.** 1985. Attenuation of late simian virus 40 mRNA synthesis is enhanced by the agnoprotein and is temporally regulated in isolated nuclear systems. Mol. Cell. Biol. **5**:1327–1334.
21. **Hay, N., H. Skolnik-David, and Y. Aloni.** 1982. Attenuation in the control of SV40 gene expression. Cell **29**:183–193.
22. **Hernandez, N., and R. Lucito.** 1988. Elements required for transcription initiation of the human U2 snRNA gene coincide with elements required for snRNA 3′ end formation. EMBO J. **1**:3125–3134.
23. **Hernandez, N., and A. M. Weiner.** 1986. Formation of the 3′ end of U1 snRNA requires compatible snRNA promoter elements. Cell **47**:249–258.
24. **Kao, S.-Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin.** 1987. Antitermination of transcription within the long terminal repeat of HIV-1 by *tat* gene product. Nature (London) **330**:489–493.
25. **Kelly, K., B. H. Cochran, C. D. Stiles, and P. Leder.** 1983. Cell-specific regulation of the c-*myc* gene by lymphocyte mitogens and platelet-derived growth factor. Cell **35**:603–610.
26. **Kerppola, T. K., and C. M. Kane.** 1988. Intrinsic sites of transcription termination and pausing in the c-*myc* gene. Mol. Cell. Biol. **8**:4389–4394.
27. **Kimura, A., A. Israel, O. Le Bail, and P. Kourilsky.** 1986. Detailed analysis of the mouse H-2K^b promoter: enhancer-like sequences and their role in the regulation of class I gene expression. Cell **44**:261–272.
28. **Krystal, G., M. Birrer, J. Way, M. Nau, E. Sausville, C. Thompson, J. Minna, and J. Battey.** 1988. Multiple mechanisms for transcriptional regulation of the *myc* gene family in small-cell lung cancer. Mol. Cell. Biol. **8**:3373–3381.
29. **Maderious, A., and S. Chen-Kiang.** 1984. Pausing and premature termination of human RNA polymerase II during transcription of adenovirus *in vivo* and *in vitro*. Proc. Natl. Acad. Sci. USA **81**:5931–5935.
30. **Manley, J. L., A. Fire, M. Samuels, and P. A. Sharp.** 1983. *In vitro* transcription: whole-cell extract. Methods Enzymol. **101**: 568–582.
31. **Mather, E. L., K. J. Nelson, S. Haimovich, and R. P. Perry.** 1984. Mode of regulation of immunoglobulin μ- and δ-chain expression varies during B-lymphocyte maturation. Cell **36**: 329–338.
32. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavage. Methods Enzymol. **65**: 499–580.
33. **McCachren, S. S., Jr., Z. Salehi, J. B. Weinberg, and J. E. Niedel.** 1988. Transcription interruption may be a common

mechanism of c-*myc* regulation during HL-60 differentiation. Biochem. Biophys. Res. Commun. **151**:574–582.

34. **Mulligan, R. C., B. H. Howard, and P. Berg.** 1979. Synthesis of rabbit beta globin in cultured monkey cells following infection with a SV40 beta globin recombinant genome. Nature (London) **277**:108–114.
35. **Nepveu, A., P. D. Fahrlander, J. Q. Yang, and K. B. Marcu.** 1985. Amplification and altered expression of the c-*myc* oncogene in A-MuLV-transformed fibroblasts. Nature (London) **317**:440–443.
36. **Nepveu, A., and K. B. Marcu.** 1986. Intragenic pausing and anti-sense transcription within the murine c-*myc* locus. EMBO J. **5**:2859–2865.
37. **Nepveu, A., K. B. Marcu, A. I. Skoultchi, and H. M. Lachman.** 1987. Contributions of transcriptional and post-transcriptional mechanisms to the regulation of c-*myc* expression in mouse erythroleukemia cells. Genes Dev. **1**:938–945.
38. **Neuman de Vegvar, H. E., E. Lund, and J. E. Dahlberg.** 1986. 3′ end formation of U1 snRNA precursors is coupled to transcription from snRNA promoters. Cell **47**:259–266.
39. **Piechaczyk, M., J. M. Blanchard, L. Marty, C. Dani, F. Panabieres, S. El-Sabouty, P. Fort, and P. Jeanteur.** 1984. Post-transcriptional regulation of glyceraldehyde-3-phosphate dehydrogenase gene expression in rat tissues. Nucleic Acids Res. **12**:6951–6963.
40. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.
41. **Reines, D., D. Wells, M. J. Chamberlin, and C. M. Kane.** 1987. Identification of intrinsic termination sites *in vitro* for RNA polymerase II within eukaryotic gene sequences. J. Mol. Biol. **196**:299–312.
42. **Resnekov, O., and Y. Aloni.** 1989. RNA polymerase II is capable of pausing and prematurely terminating transcription at a precise location *in vivo* and *in vitro*. Proc. Natl. Acad. Sci. USA **86**:12–16.
43. **Robert, J., S. B. Sloan, R. A. Welsberg, M. E. Gottesman, R. Robledo, and D. Harbrecht.** 1987. The remarkable specificity of a new transcription termination factor suggests that the mechanisms of termination and antitermination are similar. Cell **51**:483–492.
44. **Roberts, J. W.** 1988. Phage lambda and the regulation of transcription termination. Cell **52**:5–6.
45. **Salditt-Georgieff, M., M. Sheffery, K. Krauter, J. E. Darnell, Jr., R. Rifkind, and P. A. Marks.** 1984. Induced transcription of the mouse β-globin transcription unit in erythroleukemia cells. J. Mol. Biol. **172**:437–450.
46. **Schibler, W., O. Hagenbuchle, P. K. Wellauer, and A. C. Pittet.** 1983. Two promoters of different strengths control the transcription of the mouse alpha-amylase gene Amy-^{1a} in the parotid gland and the liver. Cell **33**:501–508.
47. **Selby, J. M., E. S. Bain, P. A. Luciw, and B. M. Peterlin.** 1989. Structure, sequence, and position of the stem-loop in *tar* determine transcriptional elongation by *tat* through the HIV-1 long terminal repeat. Genes Dev. **3**:547–558.
48. **Watson, R. J.** 1988. A transcriptional arrest mechanism involved in controlling constitutive levels of mouse c-*myb* mRNA. Oncogene **2**:267–272.
49. **Watson, R. J.** 1988. Expression of the c-*myb* and c-*myc* genes is regulated independently in differentiating mouse erythroleukemia cells by common processes of premature transcription arrest and increased mRNA turnover. Mol. Cell. Biol. **8**: 3938–3942.
50. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. Chasin.** 1979. DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells. Proc. Natl. Acad. Sci. USA **3**:1373–1376.
51. **Wright, S., and J. M. Bishop.** 1989. DNA sequences that mediate attenuation of transcription from the mouse protooncogene *myc*. Proc. Natl. Acad. Sci. USA **86**:505–509.
52. **Yang, J. Q., E. F. Remmers, and K. B. Marcu.** 1986. The first exon of the c-*myc* proto-oncogene contains a novel positive control element. EMBO J. **5**:3553–3562.