ELSEVIER

Biochimica et Biophysica Acta 1490 (2000) 74–86

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

www.elsevier.com/locate/bba

Intron 1 rather than 5′ flanking sequence mediates cell type-specific expression of c-myb at level of transcription elongation

Wuchao Yuan [1]

Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, Philadelphia, PA 19140, USA

Received 17 August 1999; received in revised form 22 October 1999; accepted 4 November 1999

Abstract

Previous studies have shown that expression of steady-state c-myb mRNA was regulated primarily by a block in intron 1 during transcription elongation. This study shows that the block site maps approximately 1700 bp from the start of the intron. Studies based on a reporter construct containing c-myb flanking region and intron 1 suggest that the flanking region is not important in the regulation of the cell type-specific expression of c-myb. RNA splicing of intron 1 may enhance the expression in a non-cell type-specific manner. A conserved intron domain comprising the block site is required for defining this site, but this function of the domain is independent of cell type. The cell type-specific regulation of c-myb transcription elongation is mediated by a 5′ intron sequence. A mechanism for down regulation of c-myb gene expression by the block to transcription elongation has been proposed.

Keywords: c-myb; Gene expression; Intron; RNA-binding protein; Transcription elongation

1. Introduction

In eukaryotic systems, gene regulation through transcription elongation by RNA polymerase II (pol II) has been reported for a number of cellular genes, such as c-myc [1], c-fos [2], c-myb [3], adenosine deaminase (ADA) [4], α-tubulin [5], hsp70 [6] and others [7]. A newly proposed mechanism of action for the von Hippel–Lindau tumor suppressor gene (VHL) product in its interaction with transcription elongation factor elongin (TFSIII) further suggests a crucial role for the transcription elongation step in cellular control and oncogenesis [8]. The control in transcription elongation often involves the conditional use of a block site located in the proximal region, as well as within the first exon or first intron. The recognition of such a site as a block to transcription elongation by the pol II elongation complex represses gene expression, while the pol II complex reads through the site upon modulation through certain elongation factors resulting in the expression of the gene.

The c-myb proto-oncogene encodes a nuclear transactivator that plays a central role in the regulation of hematopoietic cell development [9,10]. Abnormal expression of c-myb is involved in tumor progression in hematopoietic malignancies. The synthesis, and consequently, the steady-state levels of c-myb mRNA are high in immature hematopoietic cells, and are down-regulated during terminal differentiation.

[1] Present address: Vollum Institute L-474, Oregon Health Sciences University, 3181 SW Sam Jackson Park Road, Portland, OR 97201, USA.

0167-4781/00/$ – see front matter
PII: S0167-4781(99)00222-5

Previous nuclear run-on studies in different cell types suggest that transcription of the endogenous c-myb initiates constitutively, and that differential levels of c-myb steady state mRNA correlate primarily with a block to transcription elongation within the middle of the first intron [3,11,12]. This transcription block occurred at the same region in different cell types which expressed different levels of c-myb mRNA, and in erythroid cells induced toward differentiation by dimethyl-sulfoxide (DMSO). The precise mechanisms regulating this transcription block are not yet understood.

In order to study molecular mechanisms involved in regulation of the transcription block in the c-myb gene, a reporter construct MybCAT3 was used containing 6.5 kilobases (kb) of the c-myb 5′ flanking region and intron 1 that was able to regulate cell type-specific expression at the level of transcription. MybCAT3 was used further as a basic vector to investigate the sequences required for the regulation in vivo and in vitro. Transient transfection assays were carried out in different cell types to study the roles of the flanking and intron 1 regions in the differential expression. A *Xenopus* oocyte system was used to map the transcription block site in intron 1 and the sequences responsible for the block. The intron regions required for defining the block site and for regulating the block to transcription elongation were studied both in vivo and in vitro.

2. Materials and methods

2.1. DNA constructs

MybCAT1 is a myb-CAT fusion construct containing 2.3 kb of the flanking and exon 1 regions, 4.2-kb intron 1 and the first 31 basepairs (bp) of exon 2 of the mouse c-myb gene. This 6.5-kb myb sequence was derived from a 7.5-kb c-myb *Eco*RI genomic fragment via introducing a *Sal*I site at the 5′ end by subcloning, and *Cal*I–*Avr*II sites at the 3′ end by polymerase chain reaction (PCR). The *Sal*I–*Avr*II fragment of the 6.5-kb c-myb sequence was subcloned into pCAT basic vector (Promega) between *Sal*I and *Xba*I. The CAT protein in MybCAT1 is in frame with the upstream c-myb open reading frame (34 amino acid residues), but this large fusion protein has no detectable enzymatic activity. In order to force translation to start from the CAT start codon, the myb translation start codon at *Nco*I site in exon 1 was replaced with an 18-bp insert containing *Xho*I and *Avr*II sites, and three other ATGs in exon 2 were also mutated to ATCs. This modified construct was named MybCAT3.

ΔPromoter was a deletion of the 2.3-kb *Sal*I–*Xho*I flanking region from MybCAT3. To make TKC3, 190-bp *Hin*dIII–*Bgl*II thymidine kinase (TK) promoter region of herpes simplex virus was subcloned into the pBluescript vector (Stratagene) between *Hin*dIII and *Bam*HI as pBSTK. The 2.3-kb *Sal*I–*Avr*II region of MybCAT3 then was replaced by 196-bp *Sal*I–*Spe*I fragment of pBSTK. To make AdC3, 300-bp *Xho*I–*Xba*I PCR fragment containing adenovirus major late promoter (Ad-MLP) replaced the *Sal*I–*Avr*II portion of MybCAT3. To make CMVC3, 845-bp *Stu*I–*Hin*dIII fragment containing the immediate early promoter–enhancer sequences of cytomegalovirus (CMV) was subcloned into the pBluescript vector between *Eco*RV and *Hin*dIII as pBSCMV. An 862-bp *Sma*I–*Xho*I fragment of pBSCMV then was subcloned into MybCAT3 to replace the *Eco*47III–*Xho*I c-myb region. MybpCAT, TKCAT and AdCAT were made by subcloning the 2.3-kb *Sal*I–*Avr*II region of MybCAT3, the *Sal*I–*Avr*II TK promoter region, and the *Sal*I–*Xba*I Ad-MLP region respectively into pCAT basic vector between *Sla*I and *Xba*I. The intron 1 deletions of MybCAT3 were made either by digestion with restriction enzymes and religation or by PCR and subcloning. In vitro transcription templates, CMVXhBg, CMVPmBg and CMVBsBg, were made by PCR from CMVC3 with two primers (one derived from the CMV promoter at the *Nco*I site and the other derived from the sequences at the distal *Bgl*II restriction site of c-myb intron 1) or from CMVCAT and MybCAT3 with an additional fusion primer. All constructs have been confirmed by restriction enzyme analyses and/or sequencing.

2.2. Transient transfection and chloramphenicol acetyl transferase (CAT) assay

Two million 7OZ/3B cells growing in log phase in Roswell Park Memorial Institute 1640 medium containing 10% fetal bovine serum were transfected with

20 μg of test DNA plus 1 μg of RSV-β-galactosidase (gal) expression plasmid and 330 μg/ml diethylaminoethyl (DEAE)-dextran as described [13]. Eighty percent confluent mouse fibroblast NIH3T3 cells growing in a 60-mm Petri dish in Dulbecco's modified Eagle's medium containing 5% calf serum were transfected with 10 μg of test DNA plus 1 μg of RSV-β-gal expression plasmid by the calcium phosphate precipitation method [14]. The transfected cells were harvested, normalized by β-gal activity, and assayed for CAT activity [15,16].

2.3. *Xenopus oocyte microinjection and total RNA preparation*

The adult female frog (*Xenopus laevis*) was anesthetized with 0.15% Tricaine (Sigma) for 45 min. The ovary was removed surgically, rinsed three times in ND96 solution (5 mM HEPES pH 7.6, 100 mM NaCl, 1 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 2.5 mM sodium pyruvate, and 100 μg/ml penicillin/streptomycin), once in OR2 solution (5 mM HEPES pH 7.6, 82.5 mM NaCl, 2.5 mM KCl, and 1 mM $MgCl_2$), and incubated for 3 h at room temperature in 5 vols. of 2.5 mg/ml Collagenase B in OR2 solution. Oocytes were then rinsed once in OR2, and five times in ND96. Healthy stage-5 and -6 oocytes were selected and incubated overnight in ND96 at room temperature. Germinal vesicles of the oocytes were microinjected with closed circular double stranded DNA (0.6–1.0 ng per oocyte). After incubation in ND96 at room temperature for 18–24 h, *Xenopus* oocytes were lysed in guanidinium solution (4 M guanidinium isothiocyanate, 20 mM sodium citrate, 0.1 mM dithiothreitol, 0.5% Sarkosyl) (40 per ml) by vigorous vortexing. The lysate was loaded onto 9 ml of $CsCl_2$ cushion (4.7 M $CsCl_2$, 100 mM sodium acetate, pH 5.0 and 5 mM ethylene diamine tetraacetic acid (EDTA)) in a centrifuge tube for Beckman SW41 rotor, and centrifuged at 35 000 rpm, 20°C for 16 h. The RNA pellet was resuspended in 10 mM Tris-Cl (pH 7.4), 5 mM EDTA and 0.5% sodium dodecyl sulfate (SDS), extracted with chloroform/butanol (4:1), and precipitated with ethanol.

2.4. *Nuclear run-on*

Isolation of nuclei and nuclear run-on reactions were performed as described [17]. Log-phase 7OZ/3B cells were lysed in 5 vols. of lysis buffer (10 mM Tris-Cl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.5% Nonidet P-40) using a B-type pestle glass Dounce homogenizer. The cell lysate was mixed with an equal volume of sucrose solution (2 M sucrose, 10 mM Tris-Cl pH 7.5, 5 mM magnesium acetate, 0.1 mM EDTA, 1 mM dithiothreitol), layered on a cushion of the sucrose solution, and centrifuged at 4°C, $30\,000 \times g$ for 45 min. The nuclear pellet (10^8 nuclei) was resuspended in 1 ml of nuclear buffer (50 mM Tris-Cl, pH 8.0, 5 mM $MgCl_2$, 0.1 mM EDTA, and 40% glycerol). For nuclear run-on, 100 μl of nuclear preparation and 100 μl nuclear run-on buffer (10 mM Tris-Cl pH 8.0, 5 mM $MgCl_2$, and 300 mM KCl) were added along with 10 μl of unlabeled rNTPs (10 mM each) including either ^{32}P-labeled or cold UTP. After incubation 30 min at 30°C, the reaction was terminated with 5 vols. of the guanidinium solution by passing through no. 26 syringe needles, RNA was prepared as described above. Hybridization of the ^{32}P-labeled nuclear run-on RNA to the complimentary oligonucleotides and the cDNA was carried out as described [11,18].

2.5. *Northern analysis*

Northern blot analysis was carried out essentially as described [19]. Briefly, RNA (10 μg) was separated on 1% agarose/5% formaldehyde gel and blotted onto a nitrocellulose filter in 100 mM Tris (pH 7.5) and 1×SSC (1.5 M NaCl, 0.15 M sodium citrate, pH 7.0). The filter was then dried, UV cross-linked, and baked at 80°C for 2 h. Prehybridization and hybridization reactions were performed at 42°C in 50% formamide, 5×SSC, 1×Denhardt's solution, 1% SDS, 0.25% non-fat dry milk and 1 mg/ml of denatured salmon sperm DNA. Hybridization was carried out overnight in the presence of 10% dextran sulfate and ^{32}P-labeled probe (10^7 cpm). After hybridization, the filter was washed at room temperature twice in 2×SSC/0.1% SDS and twice at 37°C in 0.5×SSC/0.1% SDS, air-dried, and visualized by autoradiography.

2.6. *RNase protection assay*

^{32}P-labeled RNA probe was synthesized using T3

or T7 RNA polymerase as described [20], and hybridized (10^{5-6} cpm) to 5–10 μg of total oocyte RNA or RNA extracted from two nuclear run-on reactions. Hybridization (30 μl) was carried out at 50°C for 3–18 h in 40 mM PIPES (pH 6.4), 0.4 M NaCl, 1 mM EDTA and 80% formamide. The reaction mixture was diluted with 350 μl of 10 mM Tris-Cl (pH 7.5) containing 300 mM NaCl, 5 mM EDTA, 40 μg/ml RNase A and 20 U/ml RNase T1 and incubated for 45 min at 30°C. The reaction was stopped by adding 10 μl of 20% SDS and 2.5 μl of 20 mg/ml proteinase K, incubated 15 min at 37°C, and then extracted with phenol/chloroform/isoamyl alcohol (25:24:1). RNA was precipitated in 3 vols. of ethanol, dissolved in 6–10 μl of sample buffer, and analyzed on 8 M urea/4% acrylamide gels.

2.7. Nuclear extract preparation and in vitro transcription

Nuclear extract preparation and in vitro transcription were carried out essentially as described [21]. Briefly, log-phase cells were washed in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4-$7H_2O$, and 1.4 mM KH_2PO_4) and hypotonic solution containing 10 mM HEPES (pH 7.9), 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 mM dithiothreitol (DTT), and lysed in 3 vols. of the same solution using a B-type pestle Dounce homogenizer. Nuclei were centrifuged 15 min at $500 \times g$. The pellet was resuspended in a half volume of Buffer B (20 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 20 mM KCl, 0.2 mM EDTA, 25% glycerol, 0.5 mM PMSF, and 0.5 mM DTT), mixed slowly by stirring with another half volume of Buffer C (same as Buffer B except using 20% glycerol and 1.0 M KCl), and continued to stir at 4°C for 30 min. Nuclear extract was centrifuged 30 min at $25\,000 \times g$. The supernatant was dialyzed for 3 h against 50 vols. of Buffer C with 100 mM KCl, and centrifuged again.

The in vitro transcription reaction (25 μl) included 11 μl of nuclear extract (20 μg of proteins) plus reaction buffer (20 mM HEPES pH 7.6, 100 mM KCl, 0.2 mM EDTA, 20% glycerol, 0.5 mM DTT), 1 μl of 50 mM $MgCl_2$, 1 μl of 25X rNTP mix (10 mM ATP, 10 mM GTP, 10 mM CTP, and 0.1 mM UTP), 1 μl [α-^{32}P]UTP (3000 Ci/mmol, 10 mCi/ml), 0.2 pmol gel-purified template DNA fragment, 20–40 U RNasin, and dH_2O. The reaction was incubated at 30°C for 60 min, stopped by adding 175 μl of stop mix (0.3 mM Tris-Cl pH 7.4, 0.3 M sodium acetate, 0.5% SDS, 2 mM EDTA, 3 μg/ml tRNA), extracted with 200 μl of phenol:chloroform:isoamyl alcohol (25:24:1), and precipitated in 500 μl of ethanol. RNA was analyzed on 8 M urea/4% acrylamide gel.

3. Results

3.1. The 5′ flanking sequence and intron 1 of the c-myb gene are sufficient to regulate differential expression at the level of transcription elongation

To study the regulation of c-myb expression, a myb-CAT fusion plasmid (MybCAT3) was constructed, which contained the 5′ flanking region and intron 1 of c-myb (Fig. 1). After transient transfection, MybCAT3 expressed considerable CAT activity in 7OZ/3B (pre-B-cell lymphoma), but not in NIH3T3 cells (Fig. 1A, lanes 3 and 6), as is the case for expression of endogenous c-myb in each cell line [3,11]. Deletion of the flanking region in the ΔPromoter construct further demonstrated that the reporter protein was expressed from the transcript initiated from the c-myb promoter (Fig. 1A, lanes 7 and 8). These results suggest that the transcript of MybCAT3 was processed properly, and that the c-myb promoter and intron 1 are sufficient to regulate cell type-specific gene expression. To examine whether transcription from the MybCAT3 construct was initiated in both 7OZ/3B and NIH3T3 cells as observed for endogenous c-myb [3,11], nuclear run-on studies were carried out further, using two MybCAT3-specific oligonucleotides complementary to an exon 1 and a CAT sequences. As shown in Fig. 1B, after transient transfection of the MybCAT3 DNA, the transcript synthesized from the MybCAT3-specific exon 1 region was detected in both cell lines, but the transcription was elongated through the intron 1 to the reporter region only in 7OZ/3B cells. These results suggest that transcription of the MybCAT3 construct is initiated constitutively and the differential expression of CAT reporter in MybCAT3 is regulated at the level of transcription elongation. In conclusion, expression of the MybCAT3 construct is

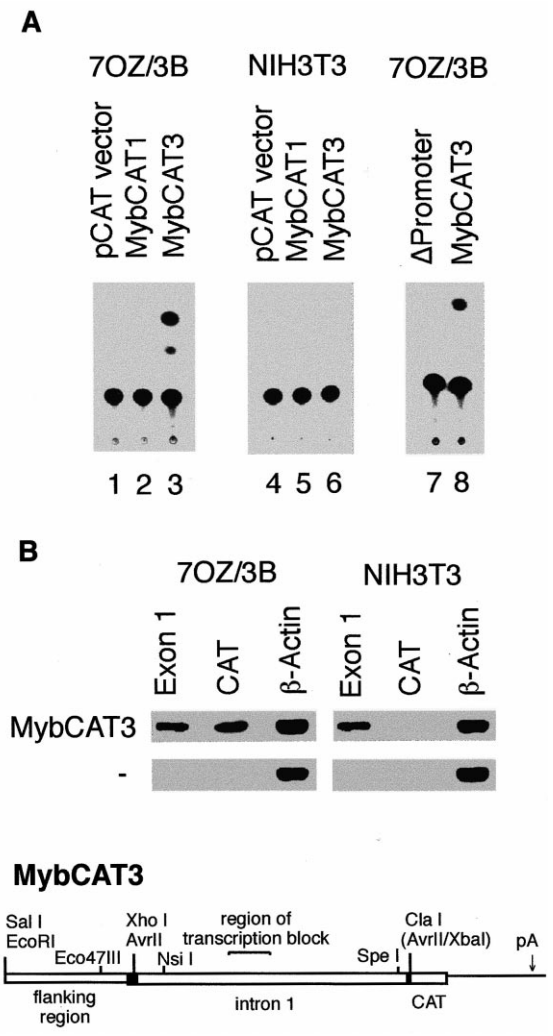


Fig. 1. The flanking region and intron 1 of c-myb is sufficient to regulate cell type-specific expression at the level of transcription. In the bottom diagram for MybCAT3, boxes represent the fusion construct of c-myb and CAT sequences, filled boxes represent c-myb exons, and horizontal line represents other vector sequence that is not drawn to scale. pA, the SV40 polyadenylation signal, is approximately 1.5 kb from the junction of the myb-CAT fusion. The region of transcription block suggested by previous nuclear run-on data is indicated. (A) DNA constructs were transiently transfected into cells and assayed for CAT activity as described in Section 2. (B) Nuclei were isolated from the cells either non-transfected or transfected for 36 h with MybCAT3 DNA. RNA probes were prepared by nuclear run-on using ^{32}P-UTP, and hybridized to slot blots containing oligonucleotides complementary to the transcription products and β-actin cDNA. 'Exon 1' represents a MybCAT3-specific oligonucleotide (38 bases) in the c-Myb exon 1 region. It contains an 18-nucleotide insert (replacing original Myb translation start codon ATG) plus upstream 11 bases and downstream 9 bases of Myb exon 1 sequences. 'CAT' represents the oligonucleotide derived from CAT sequence (bases 93–144).

regulated similarly to that of the endogenous c-myb gene.

3.2. *The c-myb intron 1 rather than the flanking region regulates cell type-specific expression by a differential repression*

To examine the role of the c-myb promoter in differential expression, the flanking region of c-myb in MybCAT3 was replaced by the TK promoter region in TKC3 and by the Ad-MLP region in AdC3. These flanking regions also were fused directly to CAT reporter in MybpCAT, TKCAT and AdCAT. These constructs were transfected into 7OZ/3B and NIH3T3 cell lines, followed by assays for transient expression of CAT reporter. As shown in Fig. 2A, MybpCAT (the flanking region alone) did not support a significant differential expression as did MybCAT3. The differential levels of the reporter expression from TKC3 and AdC3 in two cell lines were not changed when compared with MybCAT3, though strength of three promoters were different as shown in the constructs without intron 1. These results suggest that the c-myb flanking region, including promoter, does not play a regulatory role in the cell type-specific expression. Instead, c-myb intron 1 appears crucial in the regulation.

To examine whether this intron 1 function is caused by an RNA splicing event, a construct ΔIu-Id was made by deleting the majority of the intron 1 sequences from MybCAT3 except the 5′ 45 bp and the 3′ 53 bp. These remaining intron sequences in ΔIu-Id contain all the consensus sequences necessary for RNA splicing, including the 5′ splice site, the 3′ splice site and the branch point (20–50 bp upstream of the 3′ splice site) [22]. It is proposed that this small intron in the transcript of ΔIu-Id would be spliced normally. Interestingly, relative CAT activity measured from expression of ΔIu-Id was markedly higher in both cell lines in contrast with those from both MybpCAT and MybCAT3 (Fig. 2B). The remaining intron sequences in ΔIu-Id did not show any enhancer activity when inserted at the 5′ end of MybpCAT construct (data not shown). The large increase in expression of the c-myb promoter probably was caused by the splicing event in a non-cell type-specific manner.

Compared with ΔIu-Id, MybCAT3 containing the

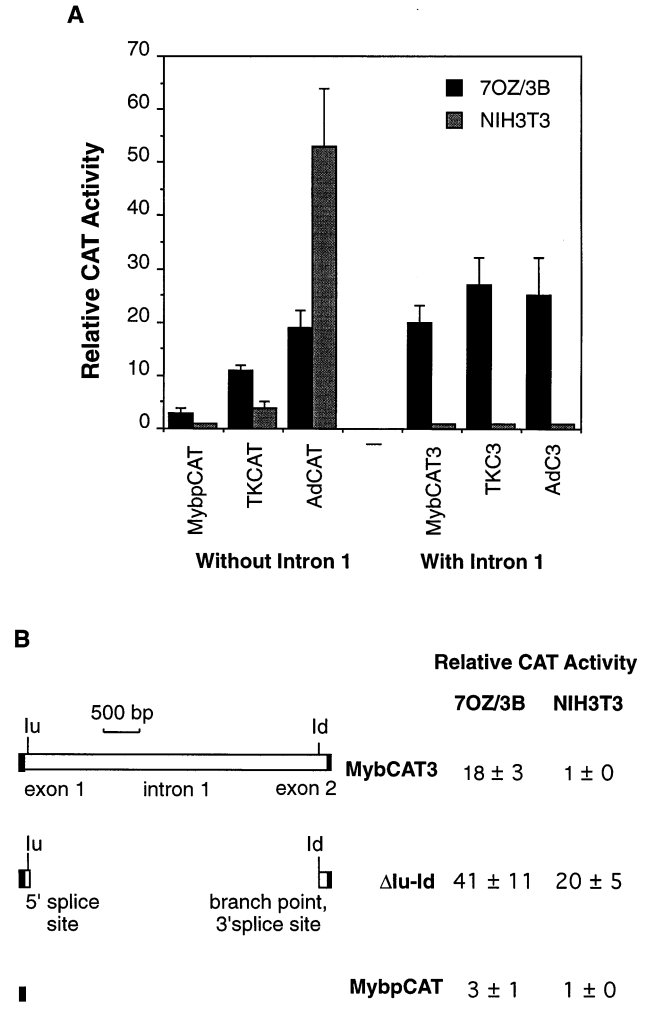


Fig. 2. Intron 1, but not the flanking region of c-myb, plays a regulatory role in the differential expression. (A) The constructs in each group differ by their 5′ flanking sequences containing promoter and are drawn on the left. Each construct was transiently transfected into cells and assayed for CAT activity. The values represent relative CAT activity compared with vector alone as 1 and standard deviation as $n=3$. (B) The portion of myb intron and exons in each construct is drawn on the left. The filled boxes represent exons.

entire intron 1 expressed near half of CAT activity in 7OZ/3B cells but a basal level in NIH3T3 cells (Fig. 2B), indicating that c-myb intron 1 regulated gene expression through a differential repression. The differential levels of repression mediated by the major intron 1 sequence is very close to the differential transcription block in the middle of intron 1 demonstrated by previous nuclear run-on assays. This intron 1-regulated repression probably is associated with the conditional block to transcription elongation in intron 1.

3.3. *Transcription elongation of c-myb is blocked at a site Tm, approximately 1700 bp from the start of intron 1*

The transcription block observed within intron 1 by previous nuclear run on assays suggested a possible occurrence of short transcripts of c-myb that arrested in intron 1. To detect the potential short transcripts, MybCAT3 was injected into oocytes for Northern analysis with a 693-nucleotide (nt) DNA probe derived from the reverse strand of the 5′ myb intron 1 sequence.

Northern analysis revealed the presence of four major transcripts in the injected, but not in the uninjected, oocytes (Fig. 3, lanes 3 and 1). Hybridization of these transcripts with the reverse strand probe indicates their sense strand origin. These transcripts were synthesized by pol II because coinjection of the template with α-amanitin (2 ng/μl) eliminated their appearance (lane 2). Hybridization of these transcripts with the 5′ intron 1 probe implied that their 5′ ends lay upstream of the intron. When the c-myb promoter in MybCAT3 was replaced by either the TK core promoter in TKC3 or the CMV immediate early promoter/enhancer regions in CMVC3, the four major transcripts were detected consistently (Fig. 3, lanes 5 and 6). Stronger expression of CMVC3 than the other two constructs probably was due to the function of enhancer sequences in the CMV flanking region. The transcripts of TKC3 and CMVC3 were also pol II-specific as demonstrated by coinjection of α-amanitin (data not shown). Because all these common transcripts are the same size, and because the inserted regions containing TK and CMV sequences and the replaced myb sequences were all different in size, initiation of transcription must occur at the promoter regions, rather than in the upstream sites.

Hybridization with the intron probe also indicated that these transcripts were either unspliced or not completely spliced. The larger 6-kb transcript corresponded with the full length of the unspliced transcript to the polyadenylation sites in the vector. The other two larger transcripts were generated from the full-length transcripts by splicing a portion of the 3′ intron region and/or a SV40 small intron in the vec-

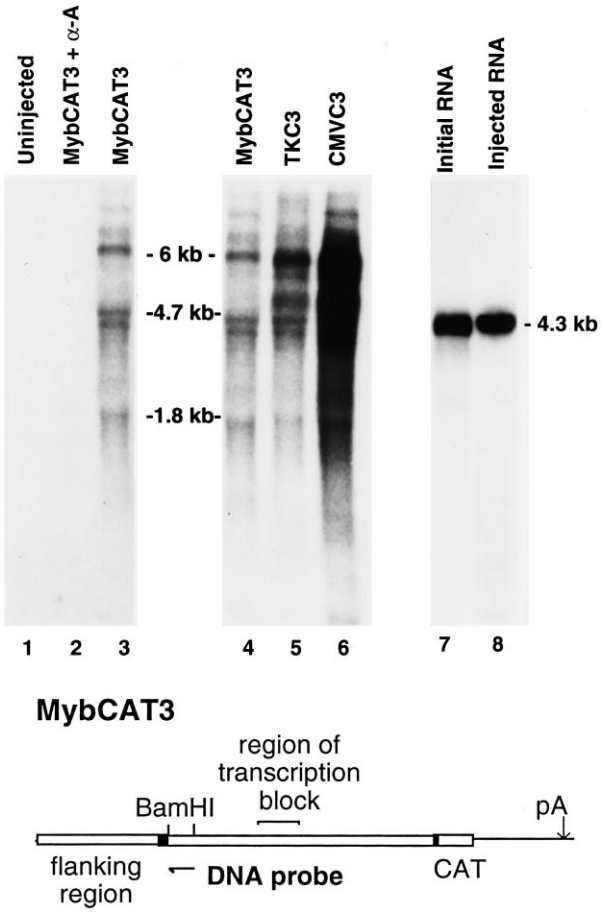


Fig. 3. Detection of short transcripts that are arrested in c-myb intron 1 in injected *Xenopus* oocytes. In lanes 1–6, RNA was extracted from oocytes either uninjected or injected with each construct, and was subject to Northern with a ^{32}P-labeled 693-nt single-stranded antisense DNA probe as indicated in the bottom diagram. Lane 2, coinjection with 2 μg/ml α-amanitin. Numbers on the side indicate approximate sizes of RNA in kb. The three larger common transcripts that were synthesized from CMVC3 (lane 6) were distinguished in a shorter exposure. Lanes 7 and 8 are processing control. The entire intron 1 RNA transcribed in vitro from the *Xho*I and *Cla*I fragment of MybCAT3 (Fig. 1) was capped, gel-purified and injected into oocytes (10^3 cpm/oocyte). After incubation for 18 h, RNA (lane 8) was extracted and analyzed on a 1% agarose/formaldehyde gel. Lane 7, gel-purified RNA before injection into oocytes.

tor region (data not shown). Because these two transcripts were not of interest in this investigation, they were not analyzed further.

The shortest major transcript detected in Northern analysis was about 1.8 kb corresponding with the region from exon 1 to the transcription block in intron 1, which was observed in the nuclear run-on assays. This short transcript probably was not generated from the longer transcripts by specific processing in the oocytes because the in vitro transcribed and capped intron 1 RNA (4.3 kb) was not processed to a stable and smaller form after injection into oocyte nuclei (Fig. 3, lane 8). It is believed that this 1.8-kb short RNA represents the transcript produced by the block during transcription elongation. The site for this block is designated Tm.

In further experiments to map the 3′ end of this short transcript with RNase protection assays, first, total RNA isolated from oocytes after injection with MybCAT3 was used. When riboprobe 1 derived from a 516-nt *Bst*EII–*Bgl*II fragment was used, a distinct doublet of RNA fragments of approximately 260 and 265 nt was detected (Fig. 4A, lane 3). The size of these protected RNA fragments allowed us to define the 3′ end of the transcript at Tm, approximately 1700-bp downstream from the start of intron 1.

To confirm that this region does in fact represent the 3′ end of the 1.8 kb RNA detected in Northern analysis, the 1.8-kb RNA band was purified from agarose/formaldehyde gels, and hybridized with riboprobe 2 (which is 30 nt shorter at the end than riboprobe 1), followed by RNase digestion. A distinct short RNA fragment was protected with a size of approximately 230 nt (Fig. 4B, lane 4). The size of this protected fragment from the gel-purified 1.8-kb RNA fraction indicated that the Tm site mapped with riboprobe 1 was responsible for the generation of this transcript.

Injection of a DNA template in an over-dosage amount into oocytes could cause non-specific transcriptional arrests. However, the dosage of the c-myb reporter constructs used here in the oocyte injection was not an over-dosed amount that had been determined by an initial calibration experiments (data not shown). Based on the previous publications [3,11,18,23], the conditions used in this oocyte injection study supported a specific measure. The real question would be whether the transcription block site Tm identified in oocytes occurs in vivo.

To examine whether the Tm site is used also in mammalian cells, cold nuclear run-on RNA was prepared from 7OZ/3B and NIH3T3 cells and used in RNase protection assays with labeled riboprobe 1. The results in both cells revealed the protection of

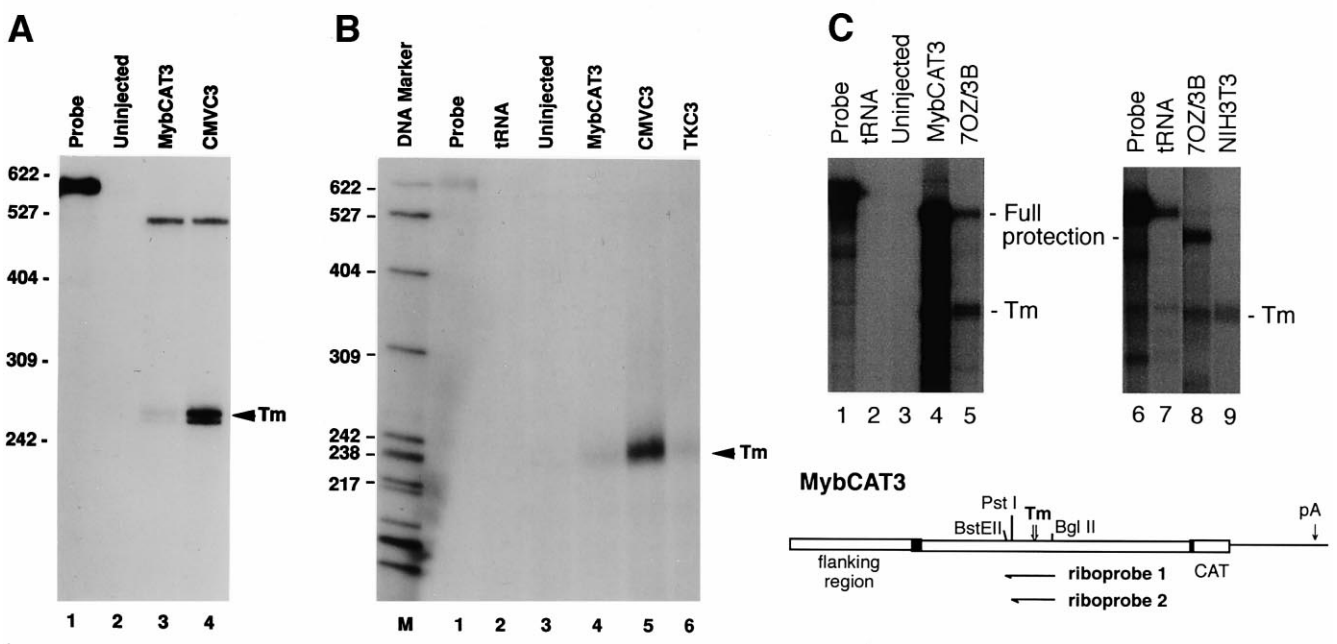


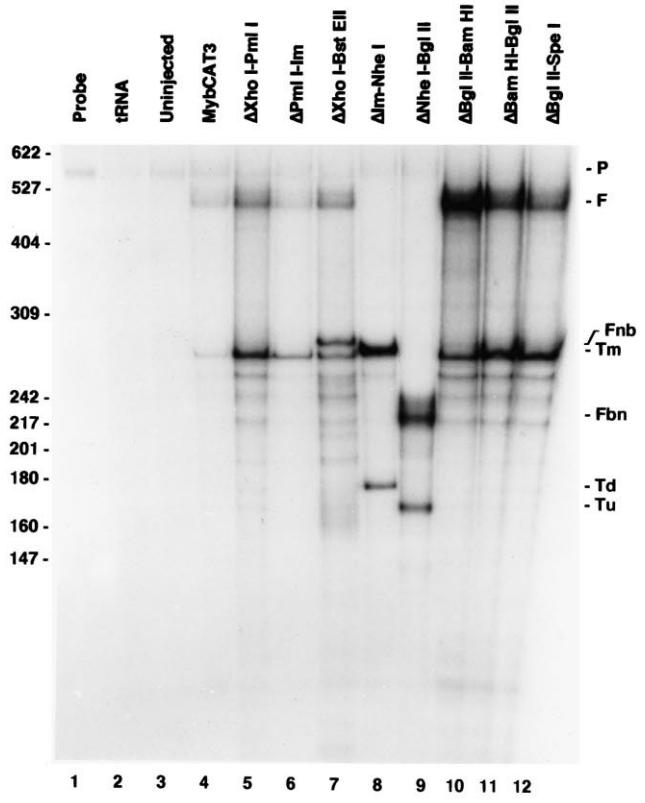
Fig. 4. Transcription block site Tm of c-myb resides approximately 1700 bp from the start of intron 1. (A) Total RNA (10 μg) extracted from oocytes either uninjected or injected with each construct was subject to RNase protection assay with 10^5 cpm ^{32}P-labeled riboprobe 1 (indicated in the bottom diagram). (B) The 1.8-kb short transcript synthesized from each construct in injected oocytes was purified from 1% agarose/formaldehyde gel with Bio101 RNaid kit, and then subject to RNase protection assay with labeled riboprobe 2 (indicated in the bottom diagram). Lane 2, a control in which yeast tRNA instead of gel-purified RNA was used. (C). 7OZ/3B (lanes 5 and 8) and NIH3T3 (lane 9) nuclei were isolated and used for preparation of cold nuclear run-on RNA followed by RNase protection assay. 10^6 cpm of ^{32}P-labeled riboprobe 1 was used in the hybridization at 50°C for 18 h. Lanes 3 and 4 are RNA prepared from the oocyte either uninjected or injected with MybCAT3. The protected transcripts ended at Tm (Tm) and the full length-protected transcripts are indicated.

fragments about 260 and 265 nt (Fig. 4C, lanes 5, 8 and 9), identical with the results obtained from the injected oocyte's RNA (lane 4). These results strongly suggest that the same transcription block site, Tm, is utilized in vivo during elongation. In addition to the short transcripts, the protection of the full-length probe was observed only in 7OZ/3B cells, but not in NIH3T3 cells, consistent with the differential expression of mature c-myb mRNA in these two cell lines and the results obtained from above transfection assays. For 7OZ/3B cells the intensity of the fully protected probe was approximately half of the sum of the 260 and 265 length bands, and the short protected RNA contained about half the amount of the labeled U residues compared with the fully protected probe. This suggest that approximately 25% of the pol II molecules read through the Tm site in this cell line.

3.4. *Sequences in intron 1 not only define the block site, but also mediate cell type-specific regulation of the block to transcription elongation*

Identification of the transcription block site Tm allows the examination of the sequences responsible for this block. The 1.8-kb short transcript as shown in Fig. 3 was synthesized from all three promoters and ended at the identical Tm site, suggesting that the c-myb promoter did not participate in defining Tm.

To examine the intron sequences required for the transcription block, first, it was briefly checked if any changes in transcription elongation could be detected when intron 1 deletion constructs were injected into oocytes. Results from these experiments showed that two deletions (ΔIm–*Nhe*I and Δ*Nhe*I–*Bgl*II) altered the block site and also increased polymerase read-



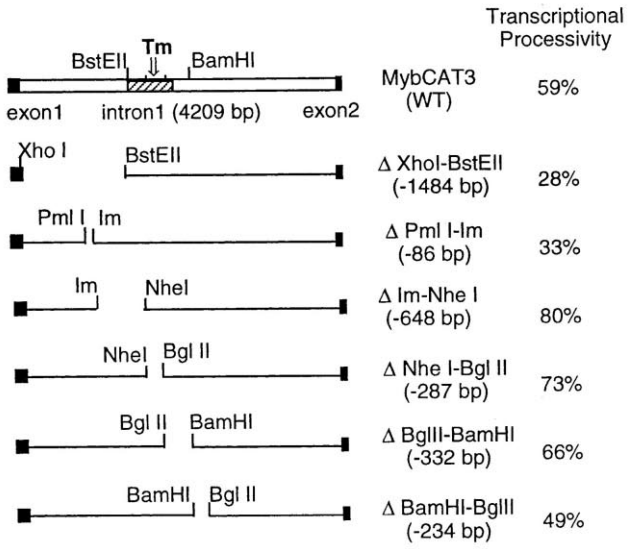


Fig. 5. The intron 1 sequences are involved in transcription block at Tm in oocytes. Total RNA (5 μg) prepared from oocytes either uninjected or injected with each construct was analyzed by RNase protection with ^{32}P-labeled riboprobe 1 (Fig. 4). For the purpose of quantitation, a molar amount of each construct was injected into oocyte's nuclei, an equal amount of each input RNA used in the assay was confirmed based on the amount of 18S and 28S RNA, and an excess amount of the probe (10^6 cpm) was used in hybridization. DNA marker is shown on the left in nt. The major bands are indicated on the right. P, undigested probe; F, the full protection (516 nt) of intron RNA by the probe; Fnb, the full protection (285 nt) of Δ*Nhe*I–*Bgl*II transcripts; Fbn, the full protection (234 nt) of ΔIm–*Nhe*I transcripts; Tm, the protected transcripts ended at Tm; Td, the protected RNA ended downstream from Tm for ΔIm–*Nhe*I (lane 8); Tu, the protected RNA ended upstream of Tm for Δ*Nhe*I–*Bgl*II (lane 9). In maps of intron deletion constructs generated from MybCAT3, Im is a site for a primer used in making the deletions, striped box represents the intron region conserved between human and mouse. 'Readthrough frequency' is measured by the ratio between the amount of the transcripts readthrough Tm and that of both arrested and readthrough Tm, and this ratio has been normalized to the ratio of the labeled U residues in both full-length and the short RNA-protected probes. Standard deviation is based on three independent experiments.
←

through frequency, i.e. percent transcripts synthesized beyond Tm (Fig. 5). This result suggested that the deleted regions participated in defining Tm. These intron regions contain a 500-bp sequence that is conserved between mouse and human (about 86% identity, high for introns) [24,25] confirming their important function. Surprisingly, none of the deletions eliminated the transcription block, which suggested that the transcription block of c-myb may be intrinsic to a larger region of the intron though the identified conserved intron region may define the specificity of the block site.

Deletion of the 5′ intron 1 regions as in Δ*Xho*I–*Bst* EII, Δ*Xho*I–*Pml*I, and Δ*Pml*I–Im decreased readthrough frequency though it did not alter the Tm site (Fig. 5). The positive role of the deleted intron sequences in transcription elongation may be caused by an anti-block activity within the 5′ intron regions.

To investigate if these active intron sequences actually function in a mammalian system, assays were carried out both in vivo and in vitro. Transfection assays in 7OZ/3B and NIH3T3 cells with reporter constructs containing various intron 1 deletions are shown in Fig. 6A. In 7OZ/3B cells, deletion of portions at the conserved intron region in Δ*Nhe*I–*Bgl*II and ΔIm–*Nhe*I increased the reporter expression, while deletion of a small 5′ intron region in Δ*Pml*I–Im decreased the expression significantly. Deletion of nearly 5′ half of the intron, including the conserved region in Δ*Nsi*I–*Bgl*IId, released repression

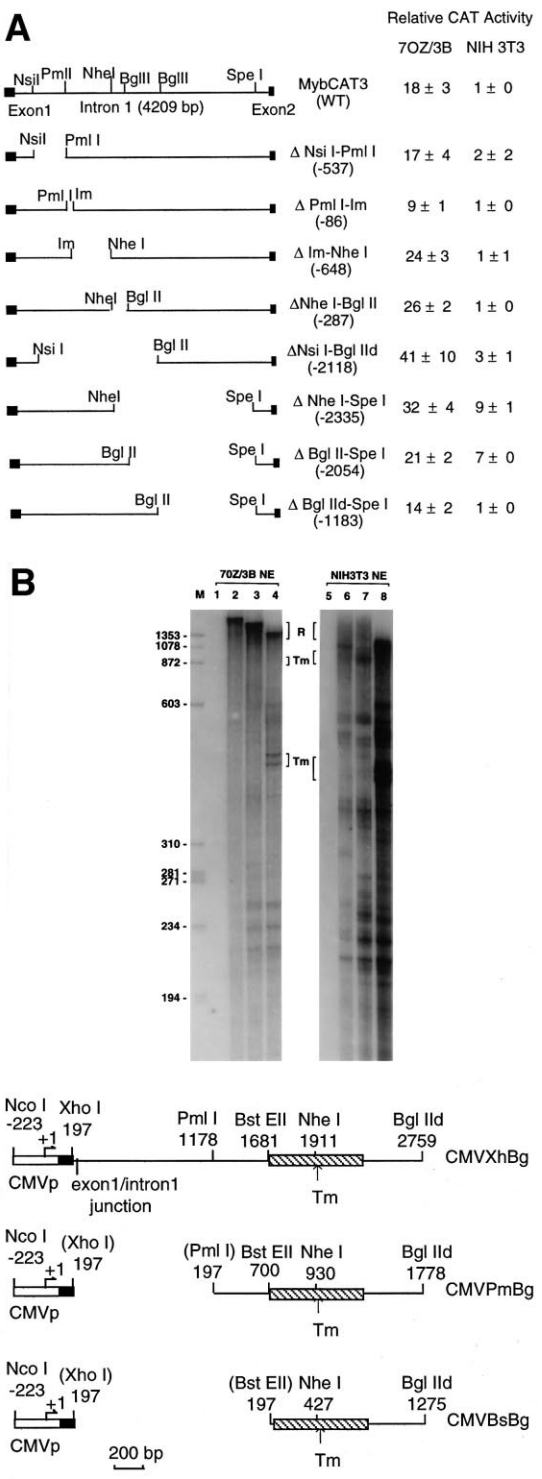


Fig. 6. The intron 1 sequences are involved in regulation of c-myb expression at the stage of transcription elongation. (A) Analysis of c-myb intron 1 function in vivo. The diagram shows the portion of c-myb exons and intron 1 in the constructs. The size of deletion is indicated in bp in parentheses. Each construct was transfected into 7OZ/3B and NIH3T3 cells and assayed for CAT activity. The values listed on the right are relative CAT activity compared with vector alone as 1 and standard deviation as $n=3$. (B) Analysis of the intron 1 functions in in vitro transcription assays. In vitro transcription was carried out in crude nuclear extracts (NE) on templates CMVXhBg (lanes 1, 2, 5 and 6), CMVPmBg (lanes 3 and 7), and CMVBsBg (lanes 4 and 8). Lanes 1 and 5, 2 μg/ml of α-amanitin was added to the reactions. Lane M, DNA marker indicated in nt on the left. Maps of the templates used in assays are diagrammed on the bottom. Open box represents CMV immediate early core promoter. The flag with an arrow indicates the transcription start site +1 and the direction of transcription. Filled box represents the spacer sequence that is derived from pBluescript vector. The horizontal line represents c-myb intron 1 sequence. Striped box represents the conserved intron region. Numbers indicate the nucleotide positions relative to +1. ←

caused by the intron 1 to a level of expression fully equivalent with ΔIu-Id's. These results are consistent with the observation in oocytes suggesting that the conserved region in the middle of intron 1 functions in c-myb expression by repression, and that the 5′ intron region functions by derepression in 7OZ/3B cells. However, in NIH3T3 cells, these deletions at the 5′ half of intron 1 caused nearly no change in expression, instead deletion of the conserved region with the 3′ half of intron 1 as in Δ*Nhe*I–*Spe*I and Δ*Bgl*II–*Spe*I released the repression markedly. It appears that the conserved region comprising Tm is important in the repression of c-myb expression but this property of the conserved region is not cell type-specific even though its requirement for additional intron regions in this repression may vary with cell types.

Transcription assays were carried out in crude nuclear extracts prepared from 7OZ/3B and NIH3T3 cells using templates of c-myb intron 1 sequences fused at the 5′ ends to the CMV immediate early promoter. In both nuclear extracts, transcription was pol II-specific demonstrated by adding α-amanitin into the reaction (Fig. 6B, lanes 1 and 5). Short transcripts ending at the Tm region (confirmed by RNase protection assays, data not shown) were detected on the template CMVBsBg containing the intron sequence only covering the conserved region in the middle of intron 1 (lanes 4 and 8). In both nuclear extracts, considerable amounts of run-off transcripts were detected, and a majority of the short

transcripts ended at the Tm region although the ends of the transcripts were found at additional other sites in NIH3T3 nuclear extracts. In agreement with the in vivo data, these results suggested that the conserved intron region defines transcription block at Tm during elongation, and this function is independent of cell type.

When the intron portion in the template was extended at the 5′ end as in CMVPmBg and CMVXhBg, almost all pol II read through Tm and reached the ends of the templates in 7OZ/3B nuclear extracts (Fig. 6B, lanes 2 and 3). Conversely, almost all pol II was arrested before reaching the ends of the templates in NIH3T3 nuclear extracts (lanes 6 and 7). It appears that the intron region between *Pml*I and *Bst*EII sites is necessary to induce this cell type-specific transcription elongation. This regulatory intron region has been shown to function in anti-block to transcription elongation in oocytes and in derepression of the reporter expression in 7OZ/3B cells. The difference in transcription elongation in different extracts observed in this study is consistent with the differential expression for MybCAT3, and with the differential levels of c-myb mRNA between these cell lines, suggesting that this regulatory intron region plays a crucial role in expression of c-myb in immature hematopoietic cells.

4. Discussion

It is believed that differential expression of c-myb gene is regulated mainly at the transcription elongation stage because of constitutive initiation of its transcription and observation of the transcription block in intron 1. Constitutive transcription initiation of c-myb has been speculated because of resemblance of the promoter sequences between the c-myb and the house-keeping genes [26], and has been supported by the observations here and others via nuclear run-on assays of transcripts of exon 1 and the 5′ intron 1 region in immature and mature hematopoietic cells and fibroblasts (NIH3T3) [3,11]. Expression of construct ΔIu-Id with minimum intron sequences sufficient for RNA splicing in different cell types supports this contention also.

A differential density of pol II observed in c-myb intron 1 in nuclear run-on assays [3,11] suggests that a fraction of c-myb transcripts are arrested in intron 1 during elongation. In *Xenopus* oocytes, a 1.8-kb short transcript was detected that was initiated from the c-myb promoter by pol II and ended at Tm in the intron. Similar short pol II transcripts were also detected in 7OZ/3B and NIH3T3 nuclear extracts suggesting that this short transcript arrested at the intron Tm site is synthesized in vivo. The short transcripts ending at Tm are detected in large amounts both in vitro and in vivo. In 7OZ/3B nuclei, approximately 75% of the detected transcripts are arrested at Tm. This amount of the short transcript detected by RNase protection assays is consistent with the density of pol II at the block site observed in nuclear run on analysis.

In Watson's nuclear run-on assays [11] density of pol II dropped sharply from the intron fragment f (*Xba*I–*Bam*HI), suggesting that the transcription block may occur near the *Xba*I site. The transcription block site Tm mapped in this study is located 30 bp upstream of the *Xba*I site. In contrast, the transcription block studied in many eukaryotic cellular genes during elongation has been promoter-proximal. Similar to the distal transcription block occurring in the c-myb gene, a unique premature arrest site in the c-fos gene resides within its intron 1, 385 bp downstream from the promoter [2]. In *Escherichia coli* λ-phage, expression of the delayed early genes from the promoter pL is regulated by transcription blocks at the promoter-distal terminators tL1, tL2 and tL3 [27]. The major transcription terminator at tL1 is located 850 bp from the transcription start, and the transcript terminated at tL3 is over 4 kb.

The two termination sites that mapped at the exon 1/intron 1 boundary of the c-myc gene in *Xenopus* oocytes correlated only with the transcription block detected by nuclear run-on assays in proliferating, but not in differentiating, human HL60 cells [18]. In fact, transcription elongation of c-myc in down-regulated differentiating cells was blocked in the promoter-proximal region found in proliferating cells and in nuclear extracts [18,28], but not found in oocytes [23]. However, only one transcription block region has been detected in c-myb gene, and this region in the middle of intron 1 is identical both in various cell types that express differential levels of c-myb and in proliferating and DMSO-induced differentiating erythroid cells [11,12]. The Tm site mapped here in

various systems is consistent with the region of the transcription block demonstrated by the nuclear run-on assays, suggesting strongly that the site Tm represents the site for block to transcription elongation of c-myb.

In many cases of prokaryotic genes, the sites for block to transcription elongation are defined by terminator sequences, and the most typical one contains a stem-loop structure followed by a stretch of U residues at the 3′ end of arrested transcripts [27]. It is not yet clear, however, whether this type of secondary structure plays a role in transcription block during elongation in eukaryotic cellular genes. It is found in this study that the conserved domain in the middle of c-myb intron 1 plays a crucial role in defining Tm, but none of this type of RNA structure correlated with Tm. Other portions of the intron also showed participation in the blockage. This situation was also found in the c-myc and ADA genes in deletion analyses [4,23]. Thus, although the transcription block site occurring in c-myb appears intrinsic to a larger intron region, the defined and conserved intron region may be responsible for specifying the block site.

Conditional block to transcription elongation usually is not regulated by the terminator sequences in prokaryotes. Similarly, the conserved region in the middle of c-myb intron 1 participates in defining the transcription block site in different cell types, but it does not regulate the cell type-specific usage of the block site. Conditional block to transcription elongation in eukaryotic cellular genes is often mediated by sequences at the promoter, such as in the c-myc gene [23,29] and in the α-tubulin gene [5]. However, in the c-myb gene, the 5′ flanking region containing promoter is not only dispensable for the transcription block site Tm but neither is it important for regulation of the cell type-specific gene expression, as shown in Fig. 2. Instead, intron 1 of c-myb plays a crucial role in this cell type-specific regulation.

By using the construct ΔIu-Id, it was shown that splicing of c-myb intron 1 could greatly enhance gene expression in a non-cell type-specific manner. Enhancement of gene expression by splicing of 5′ introns from transcripts has been observed in many eukaryotic genes [30]. Splicing factors bind to the consensus sequences in introns to form spliceosomes shortly after the transcription of introns [31], which not only specify splicing, but also signal, transcripts to be processed and exported from the nucleus efficiently [31,32,33]. Thus, one of the probable roles of c-myb intron 1 is to enhance gene expression via the post-transcriptional events.

Previously, cell type-specific interactions between nuclear factors and three c-myb intron sequences at the transcription block site and its downstream region were observed [25,34]. However, the conserved intron region comprising the block site did not function in cell type-specific regulation of the block, and deletions of the downstream nuclear factor-binding sites from MybCAT3 failed to affect the level of reporter activity in transient transfection assays (data not shown). Toth et al. [35] showed that c-myb intron 1 contained two NF-κB sites, and a reporter similar to MybCAT3 was activated when NF-κB family proteins were cotransfected into a thymoma cell line. Both NF-κB sites reside outside the regulatory region mapped here, and their role in cell type-specific expression of c-myb is not known. Their function in activation of the reporter may be associated with their enhancer activity, rather than with their direct influence in conditional transcription block at Tm.

The in vivo and in vitro data suggest that the regulatory intron region upstream of the conserved region is directly responsible for the cell type-specific regulation of transcription elongation of c-myb gene. This cell type-specific anti-block activity of the regulatory region may be associated with certain positive nuclear factors in c-myb-expressing cells. In *E. coli* λ-phage, the antiterminator N protein binds to a nut site in RNA between the promoter pL and the distal termination sites and modulates the RNA polymerase to read through the termination sites [36]. The HIV tat protein binds to the TAR element on transcribed RNA to modulate transcription initiation and elongation through the downstream termination site [36]. It has been suggested that eukaryotic cellular transcription elongation factor elongin (TFSIII) probably regulates transcription elongation of certain oncogenes through its interacting with nascent RNA and polymerase, and this function of elongin can be disrupted by the VHL tumor suppressor [8,37,38]. The model for regulation of c-myb transcription elongation tells us that the regulatory intron region of c-myb may be responsible for recruiting the 7OZ/

3B-specific nuclear factors to the pol II elongation complex to allow transcription to elongate through the downstream Tm site resulting in expression of c-myb in this cell line. Loss of activity of these anti-block factors may cause blocking of the pol II complex at Tm resulting the down-regulation of c-myb expression.

Acknowledgements

I thank E.P. Reddy, Robin H. Chou and Antonio Giordano for their assistance to this study. I also thank John J. Gartlandand, Everest O. Lam, and C.D. Reddy for valuable discussions and other help during the study and manuscript preparation.

References

[1] D.L. Bentley, M. Groudine, Nature 321 (1986) 702–706.
[2] N. Mechti, M. Piechaczyk, J.-M. Blanchard, P. Jeanteur, B. Lebleu, Mol. Cell. Biol. 11 (1991) 2832–2841.
[3] T.P. Bender, C.B. Thompson, W.M. Kuehl, Science 237 (1987) 1473–1476.
[4] V. Ramamurthy, M.-C. Maa, M. Harless, D. Wright, R.E. Kellems, Mol. Cell. Biol. 10 (1993) 1484–1491.
[5] A. Hair, G.T. Morgan, Mol. Cell. Biol. 13 (1993) 7925–7934.
[6] J. Lis, C. Wu, Cell 74 (1993) 1–4.
[7] C.A. Spencer, M. Groudine, Oncogene 5 (1990) 777–785.
[8] A. Krumm, M. Groudine, Science 269 (1995) 1400–1401.
[9] T. Graf, Genes Dev. 2 (1990) 249–255.
[10] G.L.C. Shen-Ong, Biochim. Biophys. Acta 1032 (1990) 39–52.
[11] R.J. Watson, Oncogene 2 (1988a) 267–272.
[12] R.J. Watson, Mol. Cell. Biol. 8 (1988b) 3938–3942.
[13] M.A. Lopata, D.W. Cleveland, B. Sollner-Webb, Nucleic Acids Res. 12 (1984) 5707.
[14] F.L. Graham, A.J. van der Eb, Virology 52 (1973) 46.
[15] C.M. Gorman, L.F. Moffat, B.H. Howard, Mol. Cell. Biol. 2 (1982) 1044–1051.
[16] W. Yuan, C. Condorelli, M. Caruso, A. Felsani, A. Giordano, J. Biol. Chem. 271 (1996) 9009–9013.
[17] W.F. Marzluff, R.C.C. Huang, in: B.D. Hames, S.J. Higgins (Ed.), IRL Press, Washington, DC, 1984, pp. 89–101.
[18] L. Strobl, D. Eick, EMBO J. 11 (1992) 3307–3314.
[19] P.S. Thomas, Proc. Natl. Acad. Sci. USA 77 (1980) 5201.
[20] D.A. Melton, P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn, M.R. Green, Nucleic Acids Res. 12 (1984) 7035–7056.
[21] J.D. Dignam, R.M. Lebovitz, R.G. Roeder, Nucleic Acids Res. 11 (1983) 1475–1489.
[22] M. McKeown, Curr. Opin. Cell Biol. 5 (1993) 448–454.
[23] D.L. Bentley, M. Groudine, Cell 53 (1988) 245–256.
[24] M. Castellano, J. Goley, A. Mantovani, M. Introna, Int. J. Clin. Lab. Res. 22 (1992) 159–164.
[25] C.D. Reddy, E.P. Reddy, Proc. Natl. Acad. Sci. USA 86 (1989) 7326–7330.
[26] T.P. Bender, W.M. Kuehl, Proc. Natl. Acad. Sci. USA 83 (1986) 3204–3208.
[27] T.D. Yager, P. Von Hippel, in: N.C. Neidhardt (Ed.), Vol. 2, *Escherichia coli* and salmonella: cellular and molecular biology, American Society for Microbiology, Washington, DC, 1987, pp. 1241–1275.
[28] A. Krumm, T. Meulia, M. Brunvand, M. Groudine, Genes Dev. 6 (1992) 2201–2213.
[29] T. Meulia, A. Krumm, C. Apencer, M. Groudine, Mol. Cell. Biol. 12 (1992) 4590–4600.
[30] R.D. Palmiter, E.P. Sandgren, M.R. Avarbock, D.D. Allen, R.L. Brinster, Proc. Natl. Acad. Sci. USA 88 (1993) 478–482.
[31] M. Rosbash, R.H. Singer, Cell 75 (1993) 399–401.
[32] M.R. Green, Annu. Rev. Cell Biol. 7 (1991) 559–599.
[33] J. Wang, L.G. Cao, Y.-L. Wang, T. Pederson, Proc. Natl. Acad. Sci. USA 88 (1991) 7391–7395.
[34] S. Dooley, T. Seib, C. Welter, N. Blin, Leukemia Res. 20 (1996) 429–439.
[35] C.R. Toth, R.F. Hostutler, A.S. Baldwin, T.P. Bender, J. Biol. Chem. 270 (1995) 7661–7671.
[36] J. Greenblatt, J.R. Nodwell, S.W. Mason, Nature 364 (1993) 401–406.
[37] T. Aso, W.S. Lane, J.W. Conaway, R.C. Conaway, Science 269 (1995) 1439–1443.
[38] D.R. Duan, A. Pause, W.H. Burgess, T. Aso, D.Y.T. Chen, K.P. Garrett, R.C. Conaway, J.W. Conaway, M. Linehan, R.D. Klausner, Science 269 (1995) 1402–1406.