Signature page

MOLECULAR MECHANISMS INVOLVED IN TRANSCRIPTION REGULATION OF PROTEIN-ENCODING GENES

A Dissertation

Submitted to

the Temple University Graduate Board

---

in Partial Fulfillment

of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

---

by

Wuchao Yuan

May, 1996

ABSTRACT

MOLECULAR MECHANISMS INVOLVED IN TRANSCRIPTION REGULATION OF PROTEIN-ENCODING GENES

by Wuchao Yuan

Doctor of Philosophy

Temple University, 1996

Major Advisor: Dr. Antonio Giordano

Transcription regulation plays a major role in the repertoire of gene expression from a genome in development and in many other biological processes. In eucaryotes, transcription of protein-encoding genes is catalyzed by RNA polymerase II. Regulation of this reaction requires both cis-acting elements and trans-acting factors.

I have analyzed nucleotide sequences required for the transcription regulation of the c-myb proto-oncogene. It has been reported that the expression of c-myb is regulated by a transcriptional block in intron 1 during elongation. My research shows that this block maps at position Tm, approximately 1700 bp from the start of the intron. The 5' 6.5-kb genomic sequence of c-myb including intron 1 is sufficient to regulate differential transcription of a reporter gene in cell types that express different levels of c-myb mRNA. Intron 1, rather than the 2.3-kb c-myb 5' untranslated sequence, is important in the regulation of the transcription block at Tm and in the control of differential expression of the reporter gene. A domain covering Tm is responsible for the transcription block and correlates with the differential expression. A region at the 5' end of this domain is associated with release of the transcription block in 7OZ/3B cells, in which c-myb is highly expressed. It is possible

that certain RNA-binding factors may interact with this region in transcripts and play a role in regulation of transcription elongation through Tm.

I have also analyzed the function of human p300 protein as a coactivator for MyoD. The p300 protein is a cellular target of adenovirus E1A oncoprotein and a potential transcription coactivator. E1A represses MyoD-mediated E box enhancer activity. This repression by E1A requires its interaction with p300. Further, p300 interacts with MyoD and potentiates MyoD transactivation. p300 is also able to interact with TBP and TFIIB. The ability of p300 to communicate between the basal transcription complex and DNA-binding transactivator MyoD in transcription modulation defines p300 as a transcription coactivator for MyoD. This finding suggests a mechanism by which p300 could be involved in muscle differentiation.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Antonio Giordano, for his advice during the course of my study in his lab. The portion of the dissertation researches involving the human p300 protein was supported by Sbarro Institute for Cancer Research and Molecular Medicine. The portion involving the c-myb transcription was done in Dr. E. P. Reddy's lab. I thank Dr. Reddy for his instruction and thank the other persons in his lab for their support of this research. The injection of Xenopus oocytes was done in Dr. Robin H. Chou's lab at Hahnemann University. I thank Dr. Chou for his suggestions and for offering his help in using the materials and facilities, and I thank Mr. Everest O. Lam for technical help. My dissertation research was also supported by a Grant-in-Aid of Research for 1995 from Sigma Xi, The Scientific Research Society.

Many thanks are extended to the other members of my advisory committee for their discussions, comments, advice, and other help. Specifically, I appreciate the kind help provided by Drs. Stephen H. Pilder and John J. Gartland with the text of this dissertation.

During the studies, Drs. Pierre Chambon, Gianluigi Condorelli, Richard Eckner and David M. Livingston, Armando Felsani, Stanley Fields, Michael R. Green, Nouria Hernandez, Andrew B. Lassar, Michael B. Mathews, Elizabeth Moran, Mark Ptashne, C. D. Reddy, and Roberto Weinman kindly offered gifts of published reagents necessary for the research.

I wish to acknowledge Dr. Barbara Hoffman, the Director of the Graduate Program in Molecular Biology and Genetics, Dr. Bawn B. Marks, the Assistant Dean for Graduate Studies, and the other Graduate School Office staff at the School of Medicine, especially Mrs. Shirley H. Burton. I thank Temple University for its assistance. Last but not least, I would like to thank the Jefferson Cancer Institute, and Thomas Jefferson University for their assistance and concerns during the course of my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

PROLOGUE

Differentiation and Development in Eucaryotes are Largely Controlled at the Transcriptional Level

A basic biological rule is that genes stored in the genome of the cell contain all genetic information needed for development of the whole body, and code for diverse RNAs and proteins. These gene products are involved directly in all biological activities in the organism. For control of development and other biological processes in eucaryotes, the spatial and temporal expression of various sets of genes must be regulated (Alberts et al. 1989). Expression of a gene has several steps: transcription, posttranscriptional processing, mRNA transportation, translation, and posttranslational modification. Although each of these steps has a regulatory role, gene expression is initially and largely controlled at the transcriptional level, especially at transcription initiation through a promoter of a gene.

Transcription is a process by which RNA is synthesized. Some of the RNA products, ie. messenger RNA (mRNA), serve as templates for translation of proteins. Transcription is catalyzed by RNA polymerases. In addition to RNA polymerases, transcription activation requires interactions among cis-acting elements, ie. nucleotide sequences in a genome, and a variety of transcription factors including general factors, specific transactivators, and coactivators (Wingender 1988; Gill and Tjian 1992).

Each Transcription Unit Contains Regulatory Elements for the Transcription Reaction

The transcription process is conserved in all biological organisms, from procaryotes to eucaryotes. This conservation is stored within the genome at each gene or gene cluster

(operon). The site in DNA where the transcript starts to be synthesized is called the transcription initiation site. Transcription requires a DNA sequence near the transcription initiation site, which is defined as a promoter. The promoter sequence is responsible for anchoring the RNA polymerase that is associated with the general factors to form a basal transcription complex. A promoter is required for transcription initiation, but alone it only supports a basal level of transcription. The orientation of the promoter sequence specifies the direction of transcription along the genome. The sequence of a promoter determines the specificity of the basal transcription complex formed on that promoter and influences the mode of the subsequent transcription reaction (Alberts et al. 1989).

Transcription activation requires additional cis-acting elements, either upstream or downstream of the promoter, the orientations of which are not important. Enhancers and silencers in eucaryotes are this type of regulatory elements. They are usually condition- and tissue-specific, and allow interactions with specific transactivators, and these interactions are required for the modulation of transcription initiation.

All Types of Transcription Require a Basal Machinery Formed on the Promoter

Eucaryotic RNA polymerases have at least two major differences from procaryotic ones (Sawadogo and Sentenac 1990; Zawel and Reinberg 1992). (1). Eucaryotes have three different RNA polymerases, RNA polymerase I, II, and III. RNA polymerase II (pol II) transcribes all the genes that code for mRNA. (2). Eucaryotic RNA polymerases alone do not associate with promoters. In transcription initiation, they form a basal complex on the promoter with other general factors. For RNA polymerase II, these general factors are TFIID, TFIIA, TFIIB, TFIIE, TFIIF, and TFIIH. TFIID binds to the promoter in the complex while RNA polymerase interacts with TFIID. In fact, TFIID is a protein complex

consisting of the TATA-binding protein (TBP) and many TBP-associated factors (TAFs). TBP is a common factor for all three RNA polymerases. It recognizes TATA box and cooperates with other factors to localize the basal transcription machinery on the promoter.

Transactivators Bind to Enhancer Sequences, Modulate the Rate of Transcription Initiation, and are Responsible for Specific Transactivation

The basal transcription machinery alone supports a basal level of transcription. The rate of transcription initiation is modulated by transactivators. Transactivators have two features: (1). binding specifically to DNA upstream or downstream of the promoter or enhancer, and (2). activating the promoter by increasing the rate of transcription initiation dramatically. These two functions of transactivators are carried out by their separate domains, DNA-binding domains and transactivation domains. The DNA-binding domain recognizes and anchors the transactivator onto cis-elements on the promoter or enhancer of genes. The transactivation domain may communicate with the basal transcriptional machinery and modulate transcription initiation. The binding of transactivators to DNA is required for transactivation (Pabo and Sauer 1992). These transactivators play a crucial role in control of gene activation. They may be expressed or activated conditionally, developmentally or in tissue-specific manner.

Activation of Transcription with RNA Polymerase II Requires Coactivators

During the last decade, a concerted effort has been made to identify both components, the general transcription factors and transactivators in the initiation of transcription. It is still unclear how transactivators communicate with the basal

transcriptional machinery and modulate the rate of transcription initiation. A favored hypothesis has been that in order to accelerate transcription initiation, DNA-binding transactivators are delivered to the basal transcriptional machinery through special regulatory proteins called coactivators (or adapters) (Gill and Tjian 1992). Coactivators are a new class of regulatory proteins that are required, in addition to the general factors, for DNA-binding transactivators to stimulate transcription. Identification of specific coactivators is becoming an important topic in the study of transcription regulation.

In higher eucaryotes several TAF proteins have been identified as coactivators that associate tightly with TBP (Gill and Tjian 1992; Hoey, et al. 1993). In contrast, yeast coactivators, called adapters or mediator molecules, are not associated tightly with TBP and other general transcription factors (Flanagan et al. 1991; Berger et al. 1992).

Two models have been proposed for coactivator function: (1). Coactivators may serve as bridge molecules to link the transactivator and the general factors and (2), Coactivators may not interact directly with transactivators but rather may be required to propagate the effect of a direct interaction between the transactivator and one of the general factors to the rest of the transcription complex. In addition, viral transactivators VP16, E1A and Zta interact directly with TBP (Lieberman and Berk 1991; Lee, et al. 1991; Stringer, et al. 1990). Several transactivators and different types of transactivation domains have been shown to directly contact several different general transcription factors, such as TBP, TFIIB, THIIH, and TFIIF (reviewed by Triezenberg 1995).

Regulation of Transcription Elongation is an Efficient Way to Control the Level of mRNA

Transcription termination is defined by terminator sequences where transcription stops and the sequences upstream specify the location of the termination site. Frequently,

RNA polymerases may pause or transcription may be terminated prematurely during elongation before reaching the end of the transcription unit (Yager and von Hippel 1987; Friedman 1988). This pausing or termination depends on discontinuous, inchworm-like movement of RNA polymerase (Chamberlin 1994). All RNA polymerases have two RNA-binding sites. During elongation, the polymerase changes its conformation and this change moves part of the transcript from one binding site to the other in the polymerase. Any stress that inhibits switching between these two conformation stages of the polymerase, and any road block of DNA-binding proteins during transcription elongation may cause pausing and lead to further premature termination. The transcription elongation efficiency of RNA polymerase is termed processivity (Bentley, 1995).

In eucaryotes, several transcription elongation factors have been found to stimulate transcription elongation by RNA pol II. TFIIS, a 38-kD protein, interacts with arrested Pol II, and induces it to cleave nascent RNA from the 3' end. The cleavage reaction causes pol II to move backwards, and restarts transcription to read through the block (Rudd et al. 1994; Johnson and Chamberlin 1994). TFIIF and TFSIII increase the overall rate of RNA chain elongation. The stimulation of elongation by TFIIF of a heterodimer (RAP74 and RAP30) is potentiated by phosphorylation of its RAP74 subunit (Kitajima et al. 1994). TFSIII (elongin) is a heterotrimer, and stimulates transcription elongation of certain oncogenes (Krumm and Groudine 1995). The von Hippel-Lindau tumor suppressor gene (VHL) product (Duan et al. 1995; Aso et al. 1995) down-regulates the TFSIII-stimulated transcription elongation by competing with the activating subunit A of TFSIII to bind to the other two regulatory subunit B/C complex. The mutations in VHL protein disrupting its interaction with the subunit A loss its tumor suppressor role (Kibel et al. 1995). The amino acid sequence of the smallest subunit C of TFSIII share homology with the RNA binding domain of E. coli transcription termination protein Rho (Garret et al. 1994), indicating that its function in elongation may relate to its binding to RNA.

Specific antitermination proteins, like phage N and Q proteins, and HIV tat gene product, interact with specific sites of either DNA or transcribed RNA, enhance the conformational changes in polymerase, and allow the polymerase to read through the blocks (Greenblatt et al. 1993).

Evidence reveals that some specific interactions of transactivators with enhancers not only modulate the rate of transcription initiation, but also stimulate elongation by increasing pol II's processivity (Bentley, 1995). The mechanisms by which transactivators stimulate elongation is not well understood. One possibility is that transactivators recruit elongation factors into the pre-initiation complex. Alternatively, transactivators could target a factor that modifies the transcription machinery. For example, VP16 targets TFIIH to the promoter, and decreases the rate of abortive synthesis by facilitating promoter clearance (Maldonado and Reinberg 1995).

Regulation of gene expression at the transcription elongation level has been studied extensively in procaryotes. This type of regulation has also been found frequently in the expression of viral genes. It has been observed recently that an increasing number of eucaryotic cellular genes, including many proto-oncogenes, are regulated at the transcription elongation level (Spencer and Groudine 1990). Regulation of transcription elongation offers an efficient way to control the level of mRNA in the cell. It also may be a crucial mechanism in some type of oncogenesis (Krumm and Groudine 1995).

CHAPTER 2

ANALYSIS OF NUCLEOTIDE SEQUENCES REQUIRED FOR TRANSCRIPTION REGULATION OF A PROTEIN-ENCODING GENE---THE C-MYB GENE

Introduction and Literature Review

Development and differentiation in eucaryotes are controlled essentially by spatial and temporal expression of various sets of genes. This differential expression of eucaryotic genes is regulated largely at the level of transcription initiation by DNA-binding transactivators. A mechanism of conditional block in transcription elongation by RNA pol II has been implicated recently as an additional regulatory mechanism in the differential expression of eucaryotic cellular and viral genes (reviewed by Proudfoot 1989; Spencer and Groudine 1990). The control of gene expression by conditional transcription elongation has been studied extensively in procaryotic systems (reviewed by Yager and Von Hippel 1987; Friedman 1988). This control in eucaryotes often involves the conditional use of a termination site located within either the first exon or the first intron. The recognition of such a site to block transcription elongation represses the expression of the gene. In contrast, the antiterminator protein recognizes a specific sequence in either DNA or transcribed RNA and further modulates pol II to read through the termination site resulting in the synthesis of a complete transcript (Greenblatt et al. 1993).

The mechanisms involved in the control of transcription elongation have been studied extensively in E. coli λ phage (reviewed by Greenblatt et al. 1993). Terminator sequences are required to arrest RNA polymerase to block transcription elongation. Some of these terminators function only in the presence of specific terminator proteins, such as

Rho, that recognize and interact with upstream sequences in the transcribed RNA. Antitermination leading RNA polymerase efficient enough to read through the terminator sequences usually requires additional upstream sequences. In delayed early gene expression of the lysogenic cycle of λ phage, a nut sequence in transcribed RNA is recognized by the λ N protein that communicates further with the RNA polymerase and with several other cellular factors, such as Nus proteins and the ribosomal protein S10. These interactions at the nut site change the processivity of RNA polymerase and allow the enzyme to read through the terminators. In gene expression of the lytic cycle of λ phage, a qut site located in the late gene promoter is required to mediate antitermination. The λ Q protein binds to the qut sequence and also the RNA polymerase during transcription initiation. This interaction between the Q protein and the RNA polymerase causes a conformational change in the transcription complex, and allows RNA polymerase to read through the downstream terminator. Both of these well studied examples in λ phage have been used as models for analyzing mechanisms of conditional transcription termination in eucaryotes.

The regulation of transcription elongation of eucaryotic genes has been studied mainly by the nuclear run-on assay (reviewed by Spencer and Groudine 1990). The highest resolution of this assay has been achieved in the c-myc gene by using a series of 49-nucleotide (nt) oligomers spanning the region from the P1 promoter to the 5' end of the intron 1 sequence to identify the transcription block within exon 1 (Strobl and Eick 1992). Because the terminated transcripts are unstable in mammalian cells, a Xenopus oocyte system has been used to map the 3' end of the prematurely terminated RNA in c-myc (Bentley and Groudine 1988) and adenosine deaminase (ADA) (Ramamurthy et al. 1990) because the truly terminated transcripts are stable in this system. In Xenopus oocytes the 3'-end formation of human U1 snRNA precursors (Neuman de Vagvar et al. 1986) has

been identified. The in vitro transcription assay using nuclear extracts has also been useful in studying sequence requirements for premature termination of pol II transcripts (Maderious and Chen-Kiang 1984; Moore and Sharp 1985; Gick et al. 1986; Kao et al. 1987; Kessler et al. 1989; Resnekov and Aloni 1989; Resnekov et al. 1989). Premature termination of transcription in c-myc detected in the in vitro assay using purified RNA pol II (Kerppola and Kane 1988) and in HeLa cell nuclear extracts (London et al. 1991) was consistent with the results obtained from the nuclear run-on assay and from the assay in Xenopus oocytes.

The c-myb proto-oncogene is the cellular homologue of the avian myeloblastosis virus and the avian leukemia virus (E26) transforming genes (Baluda and Goetz 1961). It encodes a sequence-specific DNA-binding protein (c-Myb) with the ability to transactivate promoters containing the specific consensus sequence (T/C)AAC(G/T)G (Biedenkapp et al. 1988; Ness et al. 1989). c-Myb plays a central role in the regulation of hematopoietic cell development (Luscher and Eisenman 1990; Shen-Ong 1990; Graf 1992). Abnormal expression of c-myb is involved in tumor progression in hematopoietic malignancies, such as leukemia (Slamon et al. 1986), and in some breast cancers (Guerin et al. 1990). Normally, c-myb is expressed primarily in immature hematopoietic cells (Westin et al. 1982; Gonda et al. 1982; Nomura et al. 1988). The expression of c-myb is down-regulated during terminal differentiation (Duprey and Boettiger 1985; Kirsch et al. 1985). The regulation of c-myb tissue-specific expression is still largely unclear.

The regulation of c-myb expression appears to be complex and occurs at several levels. It has been reported from transient transfection assays that the promoter activity of c-myb can be regulated positively by c-Myb itself in hamster fibroblasts (Nicolaides et al. 1991) and negatively in T cell lines (Guerra et al. 1995), positively by Jun family proteins in hamster fibroblasts (Nicolaides et al. 1991, 1992), and negatively by Wilms' tumor suppressor gene product WT1 in both T and B cells (McCann et al. 1995). Similar to the

housekeeping gene promoter, the mouse c-myb promoter region is relatively GC-rich, contains multiple SP1 binding sites, does not have well defined TATA and CAAT consensus sequences, and generates multiple transcription start sites (Bender and Kuehl 1986; Watson, et al. 1987; Sobieszczuk et al. 1989). These features of the c-myb promoter question that the tissue-specific expression of c-myb is regulated mainly at the level of transcription initiation.

The results obtained by nuclear run-on assays in a mammalian system (Bender et al. 1987; Watson 1988a and b) claerly showed that transcription of c-myb was initiated constitutively in different cell types, and also suggested that the differential level of mouse c-myb steady state mRNA correlated predominantly with a block to transcription elongation in the middle of intron 1. These nuclear run-on data further demonstrated that this transcription block occurred at a same intron region , not only in different cell types which expressed different levels of c-myb but also in erythroid cells induced to differentiate by dimethyl-sulfoxide (DMSO). The precise mechanisms regulating this transcription block are not yet fully understood. The effect of NFκB family transcription factors on the expression of the myb-CAT reporter containing c-myb intron 1 (Toth et al. 1995) might be associated with their abilities to regulate the block to transcription elongation, because it has been shown that transactivators can afffect pol II processivity to read through downstream pause sites during elongation (Yankulov et al. 1994).

In the studies leading to this dissertation, I constructed a plasmid of a 5' c-myb genomic sequence (including entire intron 1) fused to chloramphenicol acetyl transferase (CAT) reporter gene (MybCAT3) (Fig. 1-1). I used this construct to study the nature of the block to transcription elongation, to map the site of the block in intron 1, and examine the nucleotide sequences required for the regulation of c-myb transcription.

Materials and Methods

DNA Constructs

MybCAT1 is a myb-CAT fusion construct containing 2.3 kb of promoter/enhancer and exon 1 sequences, 4.2-kb intron 1 and the first 31 bp of exon 2 of the mouse c-myb gene (Fig. 1-1). To make MybCAT1, a 7.5-kb c-myb Eco RI genomic fragment (Lavu and Reddy 1986) was subcloned into pUC18 as pSEE. A 590-bp PCR fragment of the 3' intron 1 sequence, the exon 2 portion, and Cla I-Avr II restriction sites introduced by the reverse primer to the 3' end of the exon 2 portion was blunt-ended and subcloned into pUC18 at its Sma I site as pI1E2. The Sal I-Spe I fragment of pI1E2 was then replaced with the 6.4-kb Sal I-Spe I fragment of pSEE to generate pSX6.6. Finally, the 6.6-kb Sal I-Avr II fragment of pSX6.6 was subcloned into the pCAT basic vector (Promega) between Sal I-Xba I. The CAT protein in MybCAT1 is in frame with the upstream c-myb open reading frame.

All the ATGs in the myb exons of MybCAT1 were mutated further in order to force translation to start from the CAT start codon. The myb translation start site (Nco I) in exon 1 was replaced with Xho I and Avr II sequences by PCR using a four primer strategy (Higuchi et al. 1988) followed by subcloning the Eco47III and Nsi I fragment into MybCAT1. For mutations of the other three ATGs in the exon 2 portion, the 4586-bp Spe I-Sal I fragment of MybCAT1 was subcloned into pBluescript SK (Strategene) as BSpeCAT1. The ATGs in exon 2 were mutated to ATCs by PCR using a three primer fusion strategy (Yon and Fried 1989) followed by subcloning the Spe I and Nco I fragment into BSpeCAT1 to generate BSpeCAT2. The 4548-bp Spe I-Sal I fragment of BSpeCAT2 was subcloned into the basic MybCAT construct. This final construct was named MybCAT3 (also called Myb6.5; Fig. 1-1).

ΔPromoter was made by digestion of MybCAT3 with Sal I and Xho I and then religation of the larger fragment.

To make TKC3 (Fig. 1-2), the 190-bp Hind III-Bgl II Thymidine kinase (TK) promoter region of Herpes simplex virus (HSV) (Miksicek et al. 1986) was subcloned into the pBluescript vector between Hind III and Bam HI sites as pBSTK. The 196-bp Sal I-Spe I fragment of pBSTK was then subcloned into MybCAT3 between the Sal I and Avr II sites. To make CMVC3 (Fig. 1-2), the 845-bp Stu I-Hind III fragment which contains the immediate early promoter-enhancer sequences of cytomegalovirus (CMV) (Cullen 1986) was subcloned into the pBluescript vector (Strategene) between the Eco RV and Hind III sites as pBSCMV. The 862-bp Sma I-Xho I fragment of pBSCMV was then subcloned into MybCAT3 to replace the Eco 47III-Xho I fragment. The MybpCAT plasmid was made by subcloning the 2331-bp Sal I-Avr II fragment from MybCAT3 into the pCAT basic vector between the Sal I and Xba I sites.

MybpCAT was made by subcloning the 2.3-kb Sal I-Avr II fragment of c-myb 5' flanking sequence containing the promoter and enhancers from MybCAT3 into the pCAT basic vector between the Sla I and Xba I sites. TKCAT was made by subcloning the Sal I-Avr II TK promoter region from pBSTK into the pCAT basic vector between the Sla I and Xba I sites. CMVCAT was made by subcloning the Pst I-Sal I CMV promoter region from pBSCMV into the pCAT basic vector.

The intron 1 deletions of MybCAT3 were made either by digestion with restriction enzymes followed by ligation or by PCR and subcloning. All constructs have been confirmed by restriction enzyme analyses and/or sequencing.

<u>In vitro</u> transcription templates, CMVXhBg, CMVPmBg and CMVBsBg fragments, were prepared by PCR from CMVC3, CMVRd and CMVTd respectively with two primers, one derived from the CMV promoter at the Nco I site and the other derived from 3' sequences at the distal Bgl II restriction site of the c-myb intron 1 (Bgl IId reverse primer). CMVRd and CMVTd are two myb intron deletion constructs derived from CMVC3. In CMVRd a 981-bp sequence between Xho I and Pml I sites was deleted. In

CMVTd a 1484-bp sequence between Xho I and Bst EII sites was deleted. The 3' deletions were made from CMVTd by PCR using the CMV Nco I primer and either of the reverse primers F3d (for CMVBsF3d) or Xbd (for CMVXbd).

Transient Transfection

Two million mouse pre-B cell lymphoma 7OZ/3B cells growing in log phase in RPMI medium containing 10% fetal bovine serum at 37°C were transfected with 20 μg of test DNA plus 1 μg of RSV-β-gal expression plasmid (Santhanam et al. 1991) and 330 μg/ml DEAE-dextran as described (Lopata et al. 1984). After incubation for 40 min at 37°C with mixing every 5 min, the cells were washed with serum-free RPMI 1640 medium and then grown in RPMI medium containing 10% fetal bovine serum at 37°C for 36 h. 80% confluent mouse fibroblast NIH 3T3 cells growing in a 60-mm petri dish were transfected with 10 μg of test DNA plus 1 μg of RSV-β-gal expression plasmid by the calcium phosphate precipitation method as described (Graham and van der Eb 1973). The cells were incubated with calcium phosphate-precipitated DNA at 37°C for 18 h, and grown in Dulbecco's modified Eagle's medium containing 5% calf serum at 37°C for 48 h.

Chloramphenicol Acetyl Transferase Assay

The transfected cells were harvested, washed twice with PBS, lysed in 0.25 M Hepes buffer (pH 7.8) by three cycles of freezing on dry-ice and thawing at 37°C. The cell extract (10 μl) was normalized by assay for β-galactosidase (β-gal) activity with 1 ml of 3 mM chlorophenol red β-D-galactopyranoside (Boehringer Mannheim Biochemicals). After incubation at 37°C for 1 h, the reaction was stopped by adding 2 μl of 1 M zinc acetate, and measured for optical density (OD) at 570 nm. Cell extract with 0.05-0.1 OD_{570} of β-gal activity was assayed for CAT activity with 20 μl of 40 mM acetyl CoA and 3 μl of ^{14}C-labeled chloramphenicol in a total volume of 150 μl and incubation at 37°C for 3 h as described (Gorman et al. 1982). The reaction was stopped by extraction with 1 ml of ethyl

acetate, dried in vacuum, resuspended in 20 µl of ethyl acetate, and immediately analyzed by thin layer chromatography in a chloroform/methanol (19:1) solvent. The reaction products were quantitated with a phosphoimage analyzer. Relative CAT activity was calculated from the average of three image readings from each independent transfection and normalized to the activity of the pCAT vector as 1.

Xenopus Oocyte Microinjection

An adult female frog (Xenopus laevis) was anesthetized with 0.15% Tricaine (Sigma) for 45 min. An ovary was removed surgically, rinsed 3 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 volumes 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 by size and color, 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) and incubated in ND96 at room temperature for 18-24 h prior to RNA extraction.

Total RNA Preparation

Xenopus oocytes (40 per ml) were lysed in guanidinium solution (4 M guanidinium isothiocyanate, 20 mM sodium citrate, 0.1 mM dithiothreitol, 0.5% Sarkosyl) by vigorous vortexing. Nuclei to be used in nuclear run-on reactions were lysed in 5 volumes of the guanidinium solution by passing through a 26 gauge needle. 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 EDTA) in a centrifuge tube for a Beckman SW41 rotor, and centrifuged at 35,000 rpm, 20°C for 16 h.

The RNA pellet was resuspended in 10 mM Tris-Cl (pH7.4), 5 mM EDTA and 0.5% SDS, extracted with chloroform/butanol (4:1), and precipitated with ethanol.

Preparation of Cold Nuclear Run-on RNA

Isolation of nuclei and nuclear run-on reactions were performed as described (Marzluff and Huang 1984). Log-phase 7OZ/3B cells were washed with PBS, lysed in 5 volumes 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 glass pestle 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 xg for 45 min. The nuclear pellet (10^8 nuclei) was resuspended in 1 ml of storage buffer (50 mM Tris-Cl pH 8.0, 5 mM $MgCl_2$, 0.1 mM EDTA, and 40% glycerol), and stored at -80°C in 100 µl aliquots. For nuclear run-on, a 100 µl of aliquot was mixed with a 100 µl nuclear run-on reaction buffer (10 mM Tris-Cl pH 8.0, 5 mM $MgCl_2$, and 300 mM KCl), and 10 µl of unlabeled rNTPs (10 mM of each rNTP) were added. After incubation of nuclei at 30°C for 30 min, the RNA was extracted as described above.

Northern Blot Analysis

Northern blot analysis was carried out essentially as described by Thomas (1986). Briefly, 10 µg of RNA was electrophoresed through 1.0% agarose/5% formaldehyde gels and blotted onto nitrocellulose filters in 100 mM Tris (pH 7.5) and 1x SSC. The filters were 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, 5x SSC, 1x Denhardt's solution, 1% SDS, 0.25% nonfat 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 twice at room temperature in 2x SSC/0.1% SDS and twice at 37°C in 0.5x SSC/0.1% SDS, air-dried, and visualized by autoradiography.

RNase Protection Assay

A ^{32}P-labeled antisense RNA probe was synthesized by T3 or T7 RNA polymerase according to Melton et al. (1984), and hybridized (10^{5-6} cpm) to 5-10 µg of total oocyte RNA or RNA extracted from 2 nuclear run-on reactions. Hybridization reactions were carried out in 30 µl reaction volumes containing 40 mM Pipes (pH 6.4), 0.4 M NaCl, 1 mM EDTA and 80% formamide at 50°C for 3-18 h. 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. After incubation 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, and incubated for 15 min at 37°C. RNA was then extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated with 3 volumes of ethanol. The precipitate was dissolved in 6 -10 µl of sample buffer and electrophoresed through an 8M urea, 4% polyacrylamide gel. The gel was then dried at 80°C for 1 h, and subjected to autoradiography.

Nuclear Extract Preparation

Nuclear extract preparation was carried out essentially as described (Dignam et al. 1983). Briefly, log-phase 7OZ/3B cells were washed in PBS and in hypotonic solution (10 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM phenylmethylsulfonyl fluoride, and 0.5 mM dithiothreitol), and lysed in 3 volumes of hypotonic solution using a B type pestle Dounce homogenizer. Nuclei were pelletted for 15 min at 500 xg. The pellet was resuspended in a half volume of low salt buffer (20 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 20 mM KCl, 0.2 mM EDTA, 25% glycerol, 0.5 mM phenylmethylsulfonyl fluoride, and 0.5

mM dithiothreitol), mixed slowly with gentle stirring with the addition of a half volume of high salt buffer (same as the low salt buffer except for 1.0 M KCl, and 20% glycerol), and stirred continuously at 4°C for 30 min. The nuclear extract was centrifuged at 25,000 xg for 30 min. The supernatant was dialyzed for 5 h against 50 volumes of dialysis buffer (same as low salt protection except 100 mM KCl and 20% glycerol), and centrifuged again. The supernatant was aliquotted and stored at -80°C.

In vitro Transcription Assay

In vitro transcription assays were carried out essentially as described (Dignam et al. 1983). Briefly, a 25 µl reaction assay contained 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 dithiothreitol), 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 (3,000 Ci/mmol, 10 mCi/ml), 0.2 pmole 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 with 500 µl of ethanol. RNA was analyzed on 8M urea, 4% polyacrylamide gel.

Electromobility Shift Assay

Mobility shift DNA-binding assays using gel electrophoresis were performed essentially as described by Fried and Crothers (1981) and by Garner and Revzin (1981). The test DNA was labeled either at 5' ends with [γ-^{32}P] ATP and T4 polynucleotide kinase or at 3' ends with [α-^{32}P]dCTP and Klenow enzyme (large fragment of DNA polymerase I from E. coli). The DNA-protein binding reaction includes 10^4 cpm of the labeled DNA probe, 2.5 µl of 10x binding buffer (10 mM Tris-Cl, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1

mM DTT and 5% glycerol), 1-3 µg poly(dI-dC) and nuclear extract with 1 µg protein in a final volume of 25 µl. The reaction was incubated at room temperature for 30 min, and then analyzed on a 4% low-ionic-strength polyacrylamide gel (containing 6.75 mM Tris-Cl, pH 7.9, 1 mM EDTA, 3.3 mM sodium acetate, pH 7.9, 0.5 mg/ml of bisacrylamide and 2.5% glycerol) in a low-ionic-strength electrophoresis buffer (6.75 mM Tris-Cl, pH 7.9, 3.3 mM sodium acetate, pH 7.9, 1 mM EDTA). Electrophoresis was performed at 100 volts for 90 min as a pre-run and at 30 mA at room temperature after sample loading. The gel was dried before autoradiography.

UV Crosslinking of Nuclear Proteins to RNA

The UV crosslinking analysis used here is modified from the method described by Chodosh et al. (1986). RNA probes were synthesized and labeled by in vitro transcription with ^{32}P-UTP as described above in "RNase protection assay". The RNA-nuclear protein binding reaction included 10^5 cpm of RNA probe, 2.5 µl of 10x binding buffer, 4 U RNasin, 1 µg tRNA, 1 µg poly(dI-dC) and nuclear extract with 2 µg protein in a final volume of 25 µl. After incubating at 30°C for 20 min, the reaction was placed on ice and exposed to UV light at a 254 nm wavelength at a 2-cm distance for 40 min. Then, RNA was digested with 10 U RNase I and 10 µg RNase A at 37°C for 2 hr. The reaction was stopped by adding 27 µl sample buffer (50 mM Tris-Cl, pH 6.8, 2% SDS, 10% glycerol and 0.01-0.1% bromophenol blue) and boiled at 100°C for 5 min before being analyzed on a 10% SDS-PAGE at 11 volts/cm constant. The gel was dried before autoradiography.

Results

Mapping the 3' Ends of Prematurely Terminated Transcripts of c-myb

Detection of Transcripts Prematurely Terminated in the Middle of the First c-myb Intron in Injected Xenopus Oocytes

It was reported earlier that the tissue-specific expression of the c-myb gene was associated with premature termination in transcription elongation in the middle of intron 1 (Bender et al. 1987; Watson 1988). The presence of the block in the intron during transcription elongation indicated the probable occurrence of short transcripts of c-myb that paused or prematurely terminated in intron 1. In order to detect and identify the potential short transcripts, the MybCAT3 plasmid was used to inject Xenopus oocytes. Total RNA isolated from injected oocytes was examined by Northern analysis using a probe derived from the 693-nt reverse strand of the 5' portion of the c-myb intron 1 sequence (see Fig. 1-1 for the position of the probe).

Northern analysis revealed the presence of four major transcripts that hybridized with the 5' intron probe in the injected but not in uninjected oocyte RNA (Fig. 1-3, lanes 2 and 1). These four transcripts have sense strand origin because they hybridized with the reverse strand probe. These transcripts were synthesized by pol II because coinjection of template with α-amanitin (2 ng/μl) eliminated their appearance (Fig. 1-3, lane 5). Hybridization of these transcripts with the 5' intron 1 probe implies that their 5' ends lie upstream of the intron. When the c-myb promoter in MybCAT3 was replaced with either the HSV-TK or CMV immediate early promoter (see the maps in Fig. 1-2), the four major transcripts were consistently detected with the same probe (Fig. 1-3; the three larger common transcripts synthesized from CMVC3 were distinguished in a shorter exposure). Because these common transcripts were all the same in size, and because the inserted regions containing TK and CMV promoters and the replaced myb sequences were all

different in size, initiation of transcription must occur in the promoter regions rather than in the upstream sequences. In conclusion, the mouse c-myb promoter functioned well in oocytes, and the four detected transcripts were specifically synthesized from the c-myb promoter by pol II.

Hybridization with the intron probe also indicated that these transcripts were unspliced or not completely spliced. The larger 6-kb transcript corresponds with the full length unspliced transcript to the polyadenylation sites in the vector. The other two larger transcripts have sizes of over 4 kb, similar to the length of intron 1. These two transcripts were not terminated within the intron as shown by using a riboprobe (Nco I-Cla I) covering the 3' myb intron sequences (about 530 bp) in an RNase protection assay in which no short RNA was protected (Fig. 1-4). It is possible that these transcripts were generated from the full-length transcript by oocyte-specific splicing.

The shortest major transcript detected in the Northern analysis is about 1.8 kb (Fig. 1-3, lane 2) corresponding with the region from exon 1 to the transcription block in the intron, consistent with the results from nuclear run-on assays. It is possible that this 1.8-kb short RNA represents the transcript produced by premature termination in the middle of intron 1, and the termination site is designated Tm.

Transcription of c-myb is Prematurely Terminated at an Intron Site Tm, Approximately 1710 bp from the Start of Intron 1 in Injected Xenopus Oocytes

Experiments were carried out to map the 3' end of the 1.8-kb short transcript detected in above RNase protection assays. In the first set of experiments, total RNA isolated from oocytes injected with MybCAT3 was used. The assay was carried out with several probes (riboprobes 1-4) derived from the middle of intron 1 (see Fig. 1-1 for the positions of the probes). When using riboprobe 3 and 4, no 3' ends of major transcripts were detected (Fig. 1-5). When riboprobe 1 derived from a 516-nucleotides (nt) Bst EII-Bgl II fragment was used, a distinct doublet of RNA fragments of approximately 260 and

265 nt was detected (Fig. 1-6A, lane 3). The size of these protected RNA fragments allowed definition of the 3' end of the transcript at Tm, approximately 1710-bp downstream from the exon 1/intron 1 junction (see Fig. 1-14 for the positions of Tm). To confirm that this region does represent the 3' end of the 1.8 kb RNA detected in Northern analysis (Fig. 1-3, lane 2), the 1.8 kb RNA band from 1% agarose/formaldehyde gels was purified, and hybridized with riboprobe 2 followed by RNase digestion. A distinct short RNA fragment was protected with a size of approximately 230 nt (Fig. 1-6B, 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 is responsible for the generation of this transcript.

The Transcription Termination Site (Tm) is Utilized for Premature Termination of c-myb Transcription in Mammalian Cells

In contrast to the stability of terminated transcripts in oocytes, they are unstable and difficult to detect in mammalian cells. In our RNase protection assays with total steady state RNA prepared from several mouse cell lines, the transcripts prematurely terminated at the Tm site were never detected. Because the terminated transcripts were detectable with labeled nuclear run-on probes in mammalian cells, it could be expected that a cold nuclear run-on RNA preparation containing newly extended transcripts might be useful to directly map the 3' end of prematurely terminated transcripts in the c-myb intron 1 in mammalian cells. Thus, 2x 10^7 nuclei were isolated from log-phase 7OZ/3B cells, and incubated with cold ribonucleotides to extend the RNA. This nuclear run-on reaction was terminated by the addition of guanidinium isothiocyanate, RNA was extracted immediately, and assayed by RNase protection with the ^{32}P-labeled riboprobe 1. This RNA represents endogenous c-myb RNA synthesized in 7OZ/3B cells. The assay revealed the protection of two fragments of 260 and 265 nt (Fig. 1-7, lane 5) that were identical to the ones obtained from the protection assays with RNA derived from the oocyte injected with MybCAT3 (Fig 1-7, lane

4 and Fig. 1-6A, lane 3). These results strongly suggest that the same transcription termination site, Tm, is used in 7OZ/3B for transcription termination. In addition to the prematurely terminated transcripts, the protection of the full length probe also could be seen, consistent with the fact that this cell line produces appreciable levels of mature c-myb mRNA. The intensity of the fully protected probe was approximately half of the sum of the 260- and 265-base bands, and these short protected RNAs contained about half the amount of the labeled U residues compared with the fully protected probe. These data suggest that approximately 25% of the pol II molecules read through the Tm site in this cell line.

The 6.5-kb 5' Genomic Sequence of the Mouse c-myb Gene Containing the Promoter and Intron 1 is Sufficient to Regulate the Tissue-Specific Gene Expression at Transcriptional Level

To investigate the sequence required for the transcription block in the c-myb gene, a myb-CAT fusion reporter plasmid (MybCAT1) was constructed, which contained 2.3-kb of the 5' flanking sequence, exon 1 and intron 1 of the mouse c-myb gene (Fig. 1-1). The expression of the c-myb gene is high in 7OZ/3B cells (pre-B cell lymphoma) (Bender et al. 1987) but not detected in NIH3T3 cells (fibroblast) (Watson 1988). These two cell lines were used to examine the differential expression of the CAT reporter gene after transient transfection with this construct.

Unfortunately, no CAT activity was detected after transfection of MybCAT1 into either cell line (Fig. 1-8A, lanes 2 and 5). The translation product of MybCAT1 was predicted as a fusion protein with 34 myb amino acids added to the CAT protein at the amino terminus. This myb-CAT fusion gene was probably transcribed and translated normally in 7OZ/3B cells but the Myb-CAT fusion protein product had no or very low, if any, enzymatic activity. Based on this consideration, MybCAT1 was modified by mutations of the c-myb translation start codon in the exon 1 and the other three ATGs in the c-myb exon 2 portion in order to force translation to start from the CAT start codon.

The modified construct, MybCAT3 (Fig. 1-1), after transient transfection, did express considerable CAT activity in 7OZ/3B cells but not in NIH3T3 cells as is the case for expression of endogenous c-myb in each cell line (Fig. 1-8A, lanes 3 and 6). Deletion of the 2.3-kb c-myb promoter/enhancer region in the ΔPromoter construct further demonstrated that the functional CAT reporter protein was expressed from the transcript initiated from the c-myb promoter (Fig. 1-8A, lanes 7 and 8). These results suggest that the transcript of MybCAT3 was processed properly, and that the c-myb promoter and intron 1 unit is sufficient to regulate tissue-specific gene expression. To determine if this differential expression is controlled at the transcriptional level, total RNAs were prepared from these two cell lines transfected with MybCAT3 and analyzed by a RNase protection assay with a riboprobe derived from CAT open reading frame. This riboprobe protected transcripts only in 7OZ/3B cells but not in NIH3T3 cells (Fig. 1-8B), indicating that the tissue-specific expression of CAT reporter gene in the myb-CAT fusion constructs was regulated at the transcriptional level. These results are correlated with those obtained from the nuclear run-on assay which showed that a transcription block occurred within intron 1 to shut down transcription elongation in the cell (such as NIH3T3) in which c-myb is poorly expressed (Bender et al. 1987; Watson 1988). Northern analysis of total RNA prepared from the two cell lines with c-myb cDNA probe demonstrated that the endogenous c-myb gene was highly transcribed in 7OZ/3B cells but its expression was not detected in NIH3T3 cells (Fig. 1-8C). Thus, the tissue-specific transcription of the reporter gene that was driven by c-myb promoter and intron 1 unit closely correlated with the differential expression of c-myb mRNA in 7OZ/3B and NIH3T3 cells.

The Role of the c-myb Promoter in the Regulation of c-myb Transcription

The Termination Site Tm is Independent of the c-myb Promoter

To define the sequence necessary for premature termination at the Tm site, the role of the c-myb promoter sequence was first investigated. As shown in Fig. 1-3 (lanes 2-4), when the 5' untranslated genomic sequence containing the c-myb promoter was replaced with the HSV-TK promoter in TKC3 and with the immediate early promoter/enhancer sequences of CMV in CMVC3, Northern analysis of injected oocyte RNA showed that the prematurely terminated 1.8-kb transcript was synthesized from all three promoters. Further analysis of the 1.8-kb transcripts generated from the different promoters by RNase protection assays with both total RNA from injected oocytes (Fig. 1-6A, lanes 3 and 4) and the gel-purified 1.8-kb RNA (Fig. 1-6B, lanes 4-6) confirmed their identical 3' ends at the Tm site of the intron. These results suggest that the c-myb promoter does not participate in defining the transcription termination site Tm during transcription elongation.

The c-myb Promoter Has a Weak Potential for RNA Polymerase II's Processivity

When the transcripts both prematurely terminated at and read through Tm were examined, the three promoters allowed substantially different levels of RNA pol II to read through the Tm site (Fig. 1-2). These results suggest that the c-myb promoter potentiates a much weaker processivity of RNA polymerase II than either the TK promoter in TKC3 or the CMV promoter/enhancer in CMVC3 during transcription elongation. However, this weak potantiality of the c-myb promoter would not suggest that it played an important role in the regulation of transcription elongation and tissue-specific expression of c-myb.

The c-myb Promoter may not Play an Important Role in the Regulation of Tissue-Specific Gene Expression

To examine if the c-myb promoter plays any role in the tissue-specific expression of c-myb, the 2.3-kb flanking region of c-myb containing the promoter and enhancers was

directly fused to the CAT reporter in MybpCAT. The TK promoter and the CMV promoter/enhancer were also fused to CAT in TKCAT and CMVCAT. These three constructs, and also MybCAT3, TKC3, and CMVC3, were transfected into 7OZ/3B and NIH3T3 cell lines followed by assays for transient expression of the CAT reporter. The results are shown in Table 1-1.

Table 1-1. The Roles of the Promoter/Enhancer and Intron 1 Sequences of c-myb in the Tissue-Specific Expression

	Relative CAT Activity					
	With c-myb Intron			Without c-myb Intron		
Cell Lines	MybCAT3	TKC3	CMVC3	MybpCAT	TKCAT	CMVCAT
7OZ/3B	20	27	19	3	11	94
NIH3T3	1	1	655	1	4	599

The constructs indicated in the table were transfected into 7OZ/3B and NIH3T3 cells, and CAT activities of reporter expression were assayed as described in Materials and Methods. The values for the relative CAT activities were averaged from three independent experiments and normalized to the value of pCAT basic as 1.

The c-myb promoter in MybpCAT expressed a very low CAT activity in 7OZ/3B cells, in which myb is highly expressed. This expression in 7OZ/3B cells is quite similar to that noted in NIH3T3 cells, in which the expression of c-myb was not detected. This similar weak strength of the c-myb promoter in different cell types contrasts with the differential expression of MybCAT3 and the endogenous c-myb gene, thus indicating that the c-myb promoter may not be of importance in the regulation of tissue-specific expression of c-myb.

When the c-myb promoter region is replaced with the TK promoter in TKC3, the relative CAT activities measured from the expression of TKC3 is high in 7OZ/3B cells and at the basal level in NIH3T3 cells (Table 1-1). This differential expression of the reporter gene from TKC3 is consistent with that from MybCAT3, further suggesting that the c-myb promoter may not be of importance in the regulation of tissue-specific expression of c-myb.

When the c-myb promoter region is replaced in CMVC3 with the CMV promoter/enhancer sequences that have much stronger activity than that of c-myb, the reporter activity became very high in transfected NIH3T3 cells compared with the basal level of the reporter activity for TKC3 and MybCAT3. This high expression of the reporter from CMVC3 correlates with the high processivity of the RNA polymerase II expressed from the construct in the injected oocyte and the stronger CMV promoter activity in NIH3T3 cells. In contrast, both CMVCAT and CMVC3 expressed much weaker reporter activity in 7OZ/3B cells. However, these results may not suggest that the c-myb promoter played a crucial role in the regulation of tissue-specific expression of c-myb.

The c-myb Intron 1 Plays an Important Role in the Regulation of c-myb Expression

The c-myb Intron 1 is Required for Tissue-Specific Expression of c-myb

The proposal that the c-myb promoter may not be important in tissue-specific expression suggests an important role for the c-myb intron 1. The relative CAT activities measured from three promoter-CAT fusion constructs were very different in 7OZ/3B cells, MybpCAT activity being nearly 10-fold and 30-fold lower than the activity from the TKCAT and CMVCAT respectively. In contrast, the relative CAT activities obtained from the expression of three promoter-intron-CAT fusion constructs in 7OZ/3B cells were quite similar, i.e. 20 for the c-myb promoter, 27 for the TK promoter and 19 for the CMV promoter (Table 1-1). These results demonstrate that the addition of c-myb intron 1 between each of the different promoters and the reporter brings gene expression in 7OZ/3B

cells to a similar high level regardless of the large difference in strength of these promoters. Thus, stimulation of c-myb promoter activity by enhancers may not be able to enhance c-myb expression, and the c-myb intron 1 may play a crucial role in the control of the level of c-myb gene expression in 7OZ/3B cells.

Intron 1 Controls the Tissue-Specific Gene Expression of c-myb Through Differential Repression

To further investigate the role of intron 1, a construct (ΔIu-Id) in which the majority of the intron 1 sequences in MybCAT3 were deleted except the 5' 45 bp and 3' 53 bp of the intron 1 sequences was made (Fig. 1-9). These remaining intron sequences in ΔIu-Id contain all of the consensus sequences for splicing, i.e. the 5' splice site, the 3' splice site and the branch point (20-50 bp upstream of 3' splice site) (Nelson and Green 1989). It is proposed that this small intron in the transcript of ΔIu-Id would be spliced normally in different cell lines. Interestingly, the relative CAT activities measured from expression of ΔIu-Id is high in both cell lines (49 in 7OZ/3B cells and 24 in NIH3T3 cells) (Fig. 1-9) when compared with those from the MybpCAT construct (5 and 1) (Fig. 1-9). The non-tissue-specific amplification of the expression signals by the small intron fragment is probably because of the splicing event, which may enhance posttranscriptional processing and transportation of mRNA. This splicing-amplified signal of promoter activity also helps us to understand the constitutive expression of the c-myb promoter in different cell types.

I then compared the expression of the CAT reporter gene from the constructs, ΔIu-Id and MybCAT3, the latter containing the entire myb intron 1 (see the maps of the constructs in Fig. 1-9). Compared with ΔIu-Id which expressed high levels of CAT activity in both 7OZ/3B (relative activity, 49) and NIH3T3 (relative activity, 24) cell lines, MybCAT3 expressed only one third of CAT activity in 7OZ/3B cells (relative activity, 17) and a basal level of CAT activity in NIH 3T3 cells (relative activity, 1) (Fig. 1-9). These results indicate that addition of the majority of the intron 1 sequences in MybCAT3 caused

a three-fold repression in 7OZ/3B cells and a complete repression in NIH3T3 cells. Because the differential expression of the MybCAT3 construct in these two cell lines was at transcriptional level, the effect of the entire intron 1 on transcription elongation appears to control a differential level of repression of transcription in the different cell types.

The Sequences in the Middle of the c-myb Intron 1 Regulate c-myb Gene Expression

A Domain (Im-Bgl IId, 1.5 kb) in the Middle of Intron 1 is Involved in the Regulation of Tissue-Specific Gene Expression of c-myb in Mammalian Cells

To map the sequences required for repression, a series of intron 1 deletions (from one upstream site to the Spe I site in MybCAT3) were generated (Fig. 1-10) and transfected into 7OZ/3B and NIH3T3 cell lines. Compared with the wildtype construct MybCAT3 in 7OZ/3B cells, there were two large increases in CAT activity with the following deletions: 1). when the deletion in ΔBgl IId-Spe I was extended to the 5' Bgl II site (Bgl IIu) and further to the Nhe I site, CAT activity then fell when the deletion was increased further to the Im site (Fig. 1-10); 2). when the deletion in Im-Spe I was extended to the Nsi I site and further to the Iu site (Fig. 1-10). It seems that there were two repression domains in intron 1 functioning in 7OZ/3B cells. One is located within the first 1-kb region of the intron to the Im site, and the other one distal to the first repression domain between the Im and 3' Bgl II (Bgl IId) sites, a 1.5-kb region. In NIH3T3 cells only two deletions, ΔBgl IIu-Spe I and ΔNhe I-Spe I, give rise to detectable CAT activity, i.e. only the 2nd repression domain mapped between the Im and the 3' Bgl II sites in 7OZ/3B cells. When the levels of the increased CAT activity expressed by the deletions of the 2nd domain in the middle of the intron are compared in the two cell lines, it is noted that this repression domain functions more strongly in NIH3T3 cells (more than an eight-fold increase in the level of CAT activity) than in 7OZ/3B cells (only a one-fold increase). The transcription termination site Tm resides within this middle intron domain. The differential repression by the middle

intron domain correlates with the differential expression of the MybCAT3 construct and the endogenous c-myb gene.

The Sequences of the Middle Intron Domain Affect Gene Expression

To further examine the role of the middle intron domain, more deletions in this region from the construct MybCAT3 were made (Fig. 1-11) and assayed for CAT reporter activity when they were transfected into 7OZ/3B cells. It has been reported that three DNA fragments generated from the region within the middle intron domain bind to sequence-specific nuclear proteins and that the extent of this sequence-specific protein binding correlates with c-myb mRNA levels in different cell types (Reddy and Reddy 1989). These three intron DNA fragments are Fragment 4 (between Nhe I and the 1st Bgl II sites), Fragment 3 (between the 1st and 2nd Bgl II sites) and Fragment 1 (between Bam HI and 3rd Bgl II sites) (see the map in Fig. 1-11). Fragment 4 contains the transcription termination site Tm, and the deletion of this region from the wildtype in ΔNhe I-Bgl II causes an increase in reporter activity (Fig. 1-11) indicating its negative role in gene expression. Similarly, the deletion of the 5' region in the middle intron domain in ΔIm-Nhe I also increases the level of the reporter activity (Fig. 1-11). Thus, the intron region from Im to Bgl IIu also plays an important role in the repression of gene expression.

In contrast, when the 3' region of the middle intron domain (the 566-bp Bgl II region) containing the fragments 1 and 3 was deleted from the wildtype, the reporter activity was reduced by more than one fold (Fig. 1-11). This result indicated a positive role for this Bgl II region in transactivation of the reporter. The precise locations for the binding of the sequence-specific proteins to the fragments 1 and 3 have been defined by footprint analyses (Reddy and Reddy 1989). The nuclear protein-protected sequences are a 15-bp region (homologous to the recognition sequences of the Vaccinia virus DNA binding protein and Drosophila homeobox proteins) in Fragment 3, and a combined 49-bp (containing tandem repeat of the AP-1 recognition sequence) and 19-bp region in Fragment 1. In gel shift

assays with nuclear extracts prepared from 7OZ/3B and NIH3T3 cells, the double-stranded DNA oligomer containing either the 49 bp of Fragment 1 or the 15 bp of Fragment 3 bound to nuclear proteins in a tissue-specific manner (Fig. 1-12). However, the deletions of these two elements in ΔF1a and ΔF3 failed to affect the reporter activity compared with that of MybCAT3 in 7OZ/3B cells (Fig. 1-11).

Like ΔBgl II, the deletion of an 86-bp region at the 5' of the repression domain in ΔPml-Im reduced the reporter activity by over one fold indicating its positive role in transactivation of c-myb.

Function of the Middle Intron Region in Repression of Gene Expression is Associated with the Roles of A Conserved Termination Domain in This Region in Specifying the Transcription Termination Site Tm

Sequences at the Middle Intron Domain (Pml I-Bam HI) Control Transcription Termination at the Intrinsic Tm Site in Injected Oocytes

The next questions examined are how the sequences in the middle of intron 1 regulate the level of gene expression, and if the changes in the level of gene expression are associated with the regulation of transcription termination in the intron region. The promoter-independence of the transcription termination site Tm implies that potential sequences specifying termination at the Tm site are located in the intron. To investigate if the intron sequences regulating the level of transactivation also control transcription elongation, a series of myb intron deletions made from MybCAT3 (Fig. 1-13B) were injected into oocytes. Total RNA was extracted from the injected oocytes and assayed by RNase protection with riboprobe 1. For the purpose of quantitation, 5 μg of RNA was used to hybridize with an excess amount of the probe (10^6 cpm) for 18 h. The processivity of RNA polymerase II was calculated as the ratio between the transcripts read through the Tm site and the total RNA transcribed (transcripts read through plus transcripts terminated).

Two constructs, ΔPml I-Im and ΔXho I-Bst EII, have deletions at the 5' region of the intron (Fig. 1-13B). The deleted intron fragment Pml I-Im has been shown to participate positively in transactivation in 7OZ/3B cells. In the transfection assays in oocytes, neither deletion altered the Tm site but did decrease polymerase processivity by about two-fold, indicating the positive roles of these deleted intron sequences in regulation of transcription elongation (Fig. 1-13B).

Three deletions (ΔIm-Nhe I, ΔNhe I-Bgl II and ΔBgl II-Bam HI) in the middle of the intron, especially the first two deletions, promoted polymerase processivity (Fig. 1-13B) indicating that these deleted sequences are responsible for transcription termination at the Tm site. The deletion of the sequence upstream of the Tm site in ΔIm-Nhe I allowed higher polymerase processivity, and so this deleted sequence is more important in transcription termination. Consistent with these results, ΔIm-Nhe I and Nhe I-Bgl II also increased the level of transactivation in 7OZ/3B cells.

However, none of the three deletions in the middle of the intron eliminated transcription termination (Fig. 1-13A). Deletion of the region downstream of the Tm site (ΔBgl II-Bam HI) did not affect the position of termination. In the deletion of the sequence upstream from Tm (ΔIm-Nhe I), a 180-nt RNA fragment was protected, suggesting that the transcription termination site moved to a site Td, approximately 150-bp downstream of the Tm site (see Fig. 1-14 for the position). In the deletion including the Tm site (ΔNhe I-Bgl II), a 170-nt RNA fragment was protected, indicating that a new termination site Tu occurred approximately 90-bp upstream of the Tm site (see Fig. 1-14 for the position).

When the further downstream intron sequence Bam HI-Bgl II was deleted in the construct ΔBam HI-Bgl II, transcription processivity was reduced slightly in oocytes (Fig. 1-13B). As demonstrated above, the deletion of the sequence at this intron region in ΔBgl II caused a larger decrease in transactivation in mammalian cells (Fig. 1-11). The positive role

of the sequence at this region may be associated mainly with its potential enhancer activity to promoter non-processitive transcription.

In conclusion, the intron sequences in the middle intron domain, especially the sequence immediate 5' region of the Tm site (see Fig. 1-14 for the positions), is responsible for transcription termination at Tm in oocytes. This function of the middle intron domain is associated with its repression role in the tissue-specific expression of c-myb. The flanking intron sequences, especially the sequence at the 5' of the domain, participate in transcription regulation by enhancing pol II's processivity.

A Termination Domain Immediate Upstream of Tm Specifies Transcription Termination at Tm in 7OZ/3B Cell Nuclear Extracts

A question from the oocyte injection studies is whether this myb termination-responsible region functions the same way in a mammalian system. When the DNA sequences of this region of the c-myb intron 1 between mouse (Reddy and Reddy 1989) and human (Castellano et al. 1992) are compared (Fig. 1-14) , it is seen that this intron region contains a conserved sequence (about 86% identity between the positions 1560 and 2029 of mouse c-myb intron 1). Such a level of conserved DNA sequences is high for introns and implies that this intron region responsible for transcription termination at Tm in Xenopus oocytes may also function in mammalian systems.

To test this hypothesis, in vitro transcription assays with nuclear extracts prepared from 7OZ/3B cells were carried out. Transcription initiation at the c-myb promoter is weak, not well defined, and produces multiple transcription initiation sites (see Introduction). The CMV immediate early promoter is a strong, well-defined promoter and allows transcription termination to occur at the Tm site. The CMV promoter was fused at its 3' end to the myb intron region (Bst EII-Bgl IId) that is responsible for transcription termination at Tm (CMVBsBg, as shown in Fig. 1-15B), and used as a template in the assay. Transcription in this in vitro system was specific to pol II as demonstrated by addition of α-amanitin (Fig. 1-

15A, lane 1). In addition to the run-off transcript sythesized from the CMVBsBg template, two major short transcripts also were detected (Fig. 1-15A, lane 2). The size of these two short transcripts suggested that they were terminated at the Tm site.

To narrow down the myb intron sequences required for premature termination at the Tm site in the in vitro system, 5' and 3' deletions of the myb sequence from CMVBsBg DNA were made (Fig. 1-15B). When the 3' deletions (CMVBsF3d and CMVBsXbd) were used as templates, transcription was still terminated precisely at the Tm site (Fig. 1-15A, lanes 4 and 5). These results indicated that a 422-bp termination domain (Bst EII-Xbd region) including Tm at its 3' region (see the map in Fig. 1-16) was sufficient to specify transcription termination at Tm, and that the intron sequences downstream of the Tm site are dispensable for specifying termination at Tm. However, the 3' deleted sequences appear to enhance the level of transcription in the assays, being consistent with the transient transfection result for this intron region in 7OZ/3B cells.

When 230 bp of the 5' myb sequence was deleted in the CMVNhBg template, transcription termination did not occur at the Tm site, but occurred at a weak downstream site analogous to the Td site detected in oocytes (Fig. 1-15A, Lane 3). This further comfirms that the conserved intron sequence immediate upstream of the Tm site is responsible for specifying Tm in the mammalian system.

The Sequences in the Middle of the Intron Regulating Premature Termination at Tm are Responsible For the Tissue-Specific Transcription of c-myb

The region of intron 1-repression domain mapped in mammalian cells coincides with the region of the transcription termination domain mapped in Xenopus oocytes and in 7OZ/3B nuclear extracts. Thus, it is interesting to investigate if the differential repression of gene expression by the middle intron domain is because of transcription termination at the Tm site. For this purpose, in vitro transcription assays with crude nuclear extracts prepared from 7OZ/3B and NIH3T3 cells were used to investigate the role of the intron domain

sequences in differential transcription of c-myb. In these assays, three intron DNA fragments driven by the CMV immediate early promoter were used as templates: the termination domain in CMVBsBg and two 5'-extended intron sequences in CMVPmBg and CMVXhBg (see maps in Fig. 1-17B). As a control, 2 μg/ml α-amanitin was added in the reactions to make sure that the transcripts were synthesized by RNA pol II in this system (Fig. 1-17A, lanes 1 and 5).

Transcription patterns of these intron templates are significantly different in 7OZ/3B and NIH3T3 nuclear extracts (Fig. 1-17A). In NIH3T3 nuclear extracts, most of transcripts were terminated prematurely before elongating to the ends of the templates. In contrast, in 7OZ/3B cell nuclear extracts, many less short transcripts but a greater number of read-through or run-off transcripts were synthesized. This differential transcription caused by a differential level of premature termination at the Tm region correlated precisely with the tissue-specific expression of the construct MybCAT3 in the transfected cell lines (Fig. 1-8A and B), and with the differential levels of steady state endogenous c-myb mRNA in the cells (Fig. 1-8C). These results strongly suggest that the intron termination domain plays a regulatory role in the tissue-specific expression of c-myb mRNA. This regulation works mainly through control of premature termination in transcription elongation throughout this intron region. This result further suggests that 7OZ/3B nuclei contain antitermination activity to enhance processitive transcription.

The contrast in transcription patterns between two different cell types could be observed more clearly on the templates CMVPmBg and CMVXoBg containing the extended intron sequences upstream of the termination domain than that on CMVBsBg without the 5' intron sequence (Fig. 1-17A). On the templates CMVPmBg and CMVXoBg, the majority of transcripts read through Tm to the ends of the templates and less amount of prematurely terminated transcripts was detected in 7OZ/3B cell nuclear extracts (Fig. 1-16A, Lane 3). These results are consistent with those observed in Xenopus

oocytes, and indicate that the intron sequences upstream of the termination domain function positively to enhance processitive transcription.

The differential transcription pattern of the template CMVPml-Bgld containing 503 bp Pml I-Bst EII region of the 5' intron sequence was essentially the same as that of the template CMVXho-Bgld containing the entire 5' intron 1 sequences (Fig. 1-17A). These observations imply that the 503-bp region between Pml I and Bst EII sites makes a major contribution in antitermination to enhance processitive transcription through Tm. This result also correlated with the positive role of the Pml I-Im region in gene expression demonstrated in in vivo transfection assay.

Discussion

Some c-myb Transcripts are Prematurely Terminated at the Tm Sites of Intron 1, Approximately 1710 bp from the Start of the Intron

It has been shown previously by nuclear run-on assays that the tissue-specific expression of the c-myb gene is regulated at the transcriptional level (Bender, et al. 1987; Watson 1988). By using a reporter fusion gene system, we demonstrated that the 6.5-kb 5' genomic fragment of c-myb containing the promoter/enhancer sequences and the entire intron 1 is sufficient for tissue-specific expression of the reporter gene, and that this differential expression is controlled at the transcriptional level. Several lines of evidence have supported the contention that c-myb transcription is initiated constitutively in different cell types: 1). Sequences of the c-myb promoter resemble the promoter sequences of house keeping genes (Bender and Kuehl 1986). 2). Nuclear run-on assays demonstrated that exon 1 and 5' intron 1 sequences of c-myb were transcribed in immature and mature hematopoietic cells, and in fibroblasts (NIH3T3) (Watson 1988). 3). The construct ΔIu-Id with a minimum intron-splicing sequence expressed high levels of CAT activity in both myb-expressing and non-myb-expressing cells.

Consistent with the hypothesis that transcription initiation may not play an important role in the expression of c-myb, it has been shown by nuclear run-on assays that the tissue-specific expression of c-myb is correlated with a block in transcription elongation within the intron 1 (Bender et al. 1987; Watson 1988). The results I obtained from transcription assays in injected oocytes and in mammalian cell nuclear extracts further suggest that c-myb transcription was paused or prematurely terminated in the middle of intron 1. The 1.8-kb short transcript in oocytes is specifically initiated from the c-myb promoter by pol II and ends in a discrete site Tm in the middle of the intron. This short transcript and the Tm site were detected also in mammalian nuclei and their extracts. In these systems the short transcripts terminating at Tm were detected in large amounts. In the RNase protection assay with cold nuclear run-on RNA from 7OZ/3B cells, approximately 75% of the transcripts were truncated at the transcription block site. Such large amounts of transcripts truncated at the Tm site compared with the number of c-myb gene copies (2 copies) in the cell indicate that the short transcripts are probably generated by true termination with release of transcripts rather than by simply pausing.

Similarly, conditional blocks in transcription elongation have been found to also regulate the expression of several other eucaryotic cellular genes, such as human and mouse c-myc (Bentley and Groudine 1986; Strobl and Eick 1992), human and mouse ADA (Chinsky et al. 1989; Lattier et al. 1989), mouse and hamster c-fos (Fort et al. 1987; Schneider-Schaulies et al. 1987), mouse β-globin (Proudfoot and Whitelaw 1988), and Drosophila hsp70 (Rouguie and Lis 1988).

In Watson's nuclear run-on assays (Watson 1988a) the transcription signals dropped sharply from the intron 1 fragment f (Xba I-Bam HI) suggesting that a transcription block may be located near the Xba I site (see Fig. 1-14 for the position). However, our protection assays using RNA prepared from injected oocytes, from in vitro transcription with 7OZ/3B nuclear extracts and from the cold nuclear run-on nuclei showed

that the transcription block was caused by premature termination at the Tm site, approximately 1710 bp from the start of intron 1, just upstream of the Xba I site. We cannot rule out the possibility, however, that the transcription is actually terminated a short distance beyond the Tm site, followed rapidly by exonucleolytic cleavage back to the Tm site.

In most of the well-studied viral and cellular genes in eucaryotes, transcription blocking is promoter-proximal (Spencer and Groudine 1990; Strobl and Eick 1992). In contrast, premature termination in the c-myb gene occurs far downstream from the promoter. Similarly, premature transcription termination of mouse c-fos gene occurs at a distel intron site, 385 nt downstream from the promoter (Mechti et al. 1991). In E. 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 (Friedman 1988). The major transcription terminator at tL1 is located 850 bp from the transcription start, and the transcript terminator at tL3 is over 4 kb downstream.

Although the two termination sites at the exon 1/intron 1 boundary of the c-myc gene have been mapped in oocytes, they are correlated only with the transcription pattern analyzed by nuclear run-on assays in proliferating, but not differentiating, human HL60 cells (Strobl and Eick 1992). Actually, transcription elongation of c-myc in down-regulated differentiating cells is blocked in the promoter-proximal region. This transcription block at this proximal site was also found in proliferating cells and in nuclear extracts (Krumm et al. 1992; Strobl and Eick 1992) but not in oocytes (Bentley and Groudine 1988). In contrast, in the c-myb gene, the region of transcription block in the middle of intron 1 was identical in various cell types that express different levels of c-myb and in proliferating and DMSO-induced differentiating erythroid cells (Watson 1988a and b). In this research, premature termination site was mapped at Tm both in the injected Xenopus oocytes, and in mammalian cell nuclei and their extracts. This Tm site falls within the region of the transcription block

demonstrated by previous nuclear run-on assays. Therefore, the premature termination site Tm in the middle of c-myb intron 1 represents the site for the transcription block in elongation of the c-myb gene.

The Transcription Termination Site Tm is Specified by the Sequences of the termination domain in the Middle of the c-myb Intron 1

Sequence requirements for premature transcription termination have been studied in many viral and cellular genes of eucaryotes. In procaryotic and viral genes, premature termination often depends on terminator machinery, i.e. a stem-loop structure followed by a stretch of U residues in the transcribed RNA (Jakobovits et al. 1988; Selby et al. 1989; Kessler et al. 1989; Resnekov et al. 1989). An imperfect terminator may sitill work well in assistance of certain terminator protein, such as Rho in E. coli (Yager and Von Hippel 1987; Friedman 1988).

In contrast, terminator sequences for premature transcription termination in encaryotic cellular genes have been ill-defined. Usually, the sequences comprising or immediate upstream of the termination sites specify premature termination in these eucaryotic cellular genes (Bentley and Groudine 1988; Meulia et al. 1992; Lee et al. 1992; Ramamurthy et al. 1990; Mechti et al. 1991). It appears that the termination sites are intrinsic to these specifying sequences. Similarly, I identified a conserved termination domain (Bst EII-Xbd) comprising the termination site Tm in the middle of the c-myb intron 1, especially the 230-bp 5' region (Bst EII-Nhe I) of the domain, that played a crucial role in specifying Tm. It also appears that premature termination in c-myb gene is intrinsic to the intron sequence at the termination domain, since the deletions of the sequences at that region could alter the Tm site, but never eliminated premature termination. This Tm site is independent of the c-myb promoter and enhancer sequences. However, it is unclear whether any nuclear proteins, like terminator protein Rho, are involved in specifying transcription termination in c-myb, and also in the other eucaryotic genes.

The Differential Expression of c-myb mRNA is Regulated by Premature Termination at the Intron 1 Tm Site During Transcription Elongation

Previous nuclear run-on studies suggested that differential expression of steady state mRNA of c-myb in different cell types was due to a transcription block in elongation (Bender et al. 1987; Watson 1988). Here, I further demonstrated, by using nuclear extracts prepared from different cell types (7OZ/3B and NIH3T3 cell lines), that differential transcription of c-myb was caused by premature termination in elongation at the intron 1 Tm site. The sequences required for the regulation of premature termination have mapped within intron 1 in injected Xenopus oocytes and in mammalian nuclear extracts. These regulatory sequences include the 422-bp termination domain comprising termination site Tm that specifies premature termination and reduces pol II's processivity; they also include the 5' intron sequence, especially the 503-bp Oml I-Bst EII region, that contains antitermination activity and enhances processitive transcription. I have showed that these sequences required for the control of premature termination and attenuation at Tm regulated tissue-specific transactivation of the c-myb reporter in vivo, further indicating that differential expression of c-myb is regulated by premature termination in intron 1 during transcription elongation.

A well-known function of intron sequences, especially intron 1 sequences, has been their enhancer activity for modulating transcription initiation, as characterized in the murine metalloproteinase inhibitor intron 1 (Coulombe et al. 1988), the 4F2 T-cell surface antigen heavy-chain gene intron 1 (Gorman et al. 1982; Karppinski et al. 1989), the murine keratin 18 gene intron 1 (Oshima et al. 1990), the human β-globulin gene intron 1 (Yu et al 1991), and the purine nucleoside phosphorylase gene intron 1 (Jonsson et al. 1992). Similarly, by using injected oocytes and mammalian systems both in vitro and in vivo, I identified the intron sequence immediate downstream of the termination domain in c-myb that expressed a significant positive function in transactivation as well as contained a less dominant ability to

promoter processitive transcription. Appearantly, this intron sequence, at least in part, functioned as an enhancer to promote non-processitive transcription. As described in the Introduction section of this Chapter, the 5' flanking region of c-myb also contains several different enhancer activities to modulate transcription initiation. However, these enhancer activities of both the intron and the upstream sequences may not be of importance in increasing the level of c-myb mRNA in myb-expressing cells. This proposal has been supported by the data presented here that premature termination of transcription elongation regulated by the sequence in c-myb intron 1 held transactivation of the reporter at a certain level regardless of the level of the promoter activity. It is possible that the enhancer activities, like the one I studied in intron 1, may contribute for c-myb expression by promote processitive transcription, since certain transactivators did enhance transcription processivity, as reported by Yankulov et al. (1994). Toth et al. (1995) showed that the c-myb intron 1 contains two NF-κB-binding sites, which transactivated a reporter construct similar to MybCAT3 when NF-κB family proteins were cotransfected into B and T cell lines. One of the NF-κB-binding sites resides in the 5' intron 1 region that contains an antitermination activity; the other resides in the region immediate downstream of the termination domain, which contains a less dominant activity to enhance processitive transcription. Thus, the role of two NF-κB-binding sites in c-myb transactivation probably is associated with their potential to promoter transcription attenuation through the Tm site.

The Role of the c-myb Promoter in Differential Expression of c-myb mRNA

It is well known that the sequences in the promoter region could have an effect on transcription termination, even occurring far downstream (Greenblatt et al. 1993). A well-known example for this promoter-dependent transcription elongation in procaryotes is the transcription of the λ late genes. The phage λ Q protein recognizes the qut site in its pR' promoter region, interacts with the E. coli RNA polymerase and enables it to read through

the downstream terminator tR' (Greenblatt et al. 1993). In eucaryotic cells, the two most studied conditional transcription blocks (in c-myc and hsp70 genes) are mediated by the promoter sequences (Meulia et al. 1992; Lee et al. 1992). Premature termination of tubulin gene transcription in Xenopus oocytes is regulated by the promoter sequence (Hair and Morgan 1993). When c-myb promoter/enhancer sequences were replaced with CMV immediate early promoter/enhancer sequences in CMVC3, the complete block to transcription elongation within intron 1 of c-myb in NIH3T3 cells largely was released, though the level of transcription in 7OZ/3B cells was bearly affected. Similarly, in injected oocytes the CMV promoter/enhancer domain allowed a higher polymerase processivity to read through the Tm sites. It appears that the CMV promoter and enhancers played roles in transcription elongation, and that they functioned conditionally, in NIH3T3, but not 7OZ/3B, cells. It is possible that tissue-specific transcription factors present in NIH3T3 cells can interact either with the CMV enhancers or directly with the basal transcription machinery to modulate processitive transcription. It is still a question wether c-myb promoter plays a role in transcription elongation.

In the transfected mammalian cells, the c-myb promoter/enhancer sequences in MybCAT3 can be replaced by the HSV-TK promoter region without effect on the differential expression of the reporter gene in different cell types. In 7OZ/3B cells, the level of the expression of c-myb essentially was not diverse when the c-myb flanking sequence in MybCAT3 was replaced either by the TK promoter in TKC3, or by the CMV promoter/enhancer sequences in CMVC3, though the TK promoter, especially the CMV promoter/enhancer sequence, showed stronger promoter activities in this cells. It seems that the regulation of c-myb expression is independent of the c-myb promoter itself. A similar situation was also observed in mouse c-fos gene, which shares a similar mechanism to c-myb for premature termination in transcription elongation (Mechti et al. 1991).

Splicing of c-myb Intron 1 Largely Enhances the Potential of c-myb Expression in a Non-Tissue-Specific Manner

The sequences involved in RNA splicing are highly conserved in eucaryotes (McKeown 1993). The consensus required for the recognition of the specific factors in splicing include the 5' splice site, the 3' splice site, and the branch point at 20-50 bp upstream of the 3' splice site (Nelson and Green 1989; Mount et al. 1992). The small residue of the intron sequences in ΔIu-Id contains all consensus sequences sufficient for splicing of the intron. Thus, it is most likely that the non-tissue-specific increase of CAT activity in ΔIu-Id is because of the splicing event. Similarly, rat insulin II gene intron 1, human β-globulin gene intron B, or the SV40 t antigen gene intron greatly enhance the transcription level when they reside between the mouse metallothionein gene promoter and the intronless rat growth hormone gene, but not when they reside distal to the gene (Palmiter, et al. 1991).

The presence of many of introns in eucaryotic genes implies that they have their functions. The idea that splicing of introns plays a role in control of gene expression has been supported by new developments from the study of posttranscription mechanisms. An intron is recognized by splicing factors to form spliceosomes and is spliced shortly after the transcription of the intron (Rosbash and Singer 1993). Spliceosomes do not specify only splicing but also signal transcript export from nucleus (Green 1991). Upon microinjection into mammalian nuclei, intron-containing RNAs are targeted to defined loci (Wang et al. 1991), where pre-mRNA is spliced, fully processed and transported to nuclear pores along tracks (reviewed by Rosbash and Singer 1993). These results suggest that the first intron of a gene probably serves as a signal to target synthesized RNA to a defined pathway for complete splicing and further exportation. This hypothesis may explain why intron-less constructs are not usually expressed well in most eucaryotic cells and also why minimal functional sequences of c-myb intron 1 can enhance greatly gene expression in a non-

tissue-specific manner. Thus, one of the roles of the c-myb intron 1 is probably to enhance gene expression via the splicing event.

How is Transcription Attenuation of the c-myb Gene Regulated

In an increasing number of eucaryotic cellular genes, transcription by RNA pol II is regulated by conditional blocks during elongation in either exon 1, intron 1 or untranslated regions (Spencer and Groudine 1990). However, the detailed regulatory mechanism is not known yet. In expression of procaryotic genes, regulation of conditional transcription termination at a specific site is a process referred to as antitermination (Friedman 1988). During this process, antiterminator proteins interact with specific sequences of either DNA or transcribed RNA and further modify the RNA polymerase to allow it to read through termination or pausing sites.

In a previous study of the c-myb transcription block in the middle of intron 1, three DNA fragments at the transcription block region (Fragment 1, 3 and 4) interacted with sequence-specific nuclear proteins, the level of the interaction correlated with the level of c-myb mRNA in different cell types (Reddy and Reddy 1989). This correlation between the extent of the nuclear protein-DNA sequence interaction and the amount of the c-myb mRNA indicated a possible positive role of the involved nuclear proteins in regulation of the transcription block. Two of these three fragments (Fragments 1 and 3) located downstream of the Tm site in the Bgl II-Bgl II region have been shown to participate in positive transactivation of the reporter. Unfortunately, when the nuclear protein-footprint regions of Fragment 1 and 3 were deleted in ΔF1a and ΔF3, no effect on transactivation of the reporter gene was detected in 7OZ/3B cells. It is possible that other protein-DNA interactions within this Bgl II-Bgl II region are responsible for its positive function. Indeed, an NFκB-binding site resides in this Bgl II region, and the fragment containing this site binds to the NFκB-family proteins and is responsible for transactivation of a reporter gene (Toth et al. 1995). It is not clear yet whether the positive role of this fragment in gene expression is

because of either its enhancer activity or its antitermination activity. Most identified sequences regulating transcription attenuation in either procaryotic or eucaryotic genes reside upstream of the termination or pausing site.

The 5' sequence (Pml I-Bst EII) of the termination domain in the c-myb intron 1 has been shown to play a positive role in transcription attenuation in both injected Xenopus oocytes and 7OZ/3B nuclear extracts. Consistent with these in vitro results, an activity of enhancing gene expression was also detected for the Pml I-Bst EII region. These results suggest that the positive function of this region in gene expression is associated with its antitermination activity for transcription attenuation through the intron termination Tm site. In E. coli λ phage the late early genes are regulated by transcription termination at the distal termination sites (Greenblatt et al. 1993), which is similar to the termination site Tm in the c-myb gene far downstream of the promoter. The antiterminator N protein of the λ phage binds to the nut site in RNA and modulates the RNA polymerase II to read through the termination sites. The nut sequence resides between the promoter and the termination sites. Similarly, the sequence Pml I-Bst EII in the c-myb intron 1 containing an antitermination activity resides between the promoter and the termination site Tm. The function of the nut sequence relies on its interaction with the N protein on the transcribed RNA. This kind of interaction between the antiterminator proteins and RNA for transcription attenuation are observed in the expression of the HIV genome, i.e. the interaction of the tat protein and the TAR element on RNA (Cullen 1990). Also, human transcription elongation factor Elongin (SIII) regulates transcription elongation of proposed oncogenes through pausing sites by interacting with transcribed RNA (Krumm and Groudine 1995). Therefore, it is possible that the c-myb intron Pml I-Bst EII region promotes transcription attenuation at the Tm site through its transcribed RNA by recruiting certain RNA-binding transcription elongation factors. This hypothesis awaits further examination.

CHAPTER 3

FUNCTIONAL ANALYSIS OF TRANSCRIPTIONAL MODULATION OF TRANSCRIPTION FACTOR MYOD BY COACTIVATOR P300 PROTEIN

Introduction

Human p300 protein was identified initially as a cellular target for adenoviral E1A oncoprotein (reviewed by Moran 1993 and 1994; Bayley and Mymryk 1994). The region of E1A that interacts with p300 is distinct from the one that interacts with the tumor suppressor retinoblastoma protein (Rb) and Rb-related proteins (reviewed by Dyson & Harlow 1992). The p300-binding regions of E1A have been localized at the amino terminal region (amino acids 1-39) and the carboxyl portion of the conserved region I (CR1) (amino acids 61-80), while Rb-family protein-binding regions are at the amino portion of CR1 and CR2.

Function of p300 and its significance in cellular control have been largely predicted from studies of E1A mutants. Interactions of E1A with both p300 and Rb-family proteins are required for E1A-induced cellular transformation (reviewed by Boulanger & Blair 1991). The p300-binding regions of E1A are responsible also for transcriptional repression of many viral and cellular promoters and enhancers. It has been proposed that p300 protein participates in preventing the G0/G1 transition in the cell cycle, and stimulates differentiation pathways (Arany, et al. 1995). The amino acid sequence of p300 is conserved from nematode to human (Arany et al. 1994), further emphasizing the importance of its biological functions.

p300 protein is a relatively stable protein, is localized in the nucleus, is expressed ubiquitously, remains at a constant level throughout the entire cell cycle, and is fairly well

conserved among a variety of mammalian species (Yaciuk and Moran 1991). However, the p300 species are heterogenous in size, and there appear to be cell cycle specific changes in its phosphorylation, in that its level of phosphorylation appears to increase as it travels from late G1 into M phase (Yaciuk and Moran 1991). It has been reported that E1A inhibits the phosphorylation of p300 in vitro while the cyclin-dependent kinases, cdc2 and cdk2 are able to use p300 as a substrate (Banerjee et al. 1994). It is possible that the dynamics of p300 phosphorylation in the cell cycle are involved in the regulation of the cell cycle-specific functions of p300.

The sequence of p300 reveals it to be a potential transcriptional coactivator (Eckner et al. 1994). This potential is consistent with the fact that p300 has transcriptional activity (Arany et al. 1995), and forms a complex with TBP in vivo (Abraham et al. 1993). It has been reported that human protein p300 potentiates the transactivation by the cAMP-responsive element-binding protein (CREB), as does the homologous transcriptional coactivator, mouse CREB-binding protein (CBP) (Lundblad et al. 1995). However, the actual role and significance of this p300 function in cellular control and E1A-mediated transformation are not clear.

Skeletal myogenesis is a model system for study of cellular differentiation. Differentiation of skeletal myocytes proceeds through a stage of commitment/determination and one of differentiation. During this process, the cells withdraw from the cell cycle and further express muscle specific genes (reviewed by Lassar et al. 1994). An important step toward understanding this process was the finding that the single gene product, MyoD, could induce non-muscle cells to commit to a myogenic differentiation pathway (Davis et al. 1987). Since then, several myogenic regulatory proteins, such as myf5, myogenin and MRF4, have been discovered that are required for terminal differentiation of muscle cells (Olson and Klein 1994). These MyoD-family proteins belong to a large class of DNA-binding proteins that contain basic helix-loop-helix (bHLH) secondary structures in their

DNA-binding domains. These muscle-specific bHLH regulators form heteromeric complexes with ubiquitous bHLH proteins (E12 and E47), bind specifically to an E-box DNA sequence (CANNTG), and transactivate transcription of muscle differentiation-specific genes (Murre, et al. 1989; Lassar et al. 1991).

MyoD functions mainly at the commitment/determination stage to maintain myocytic lineage. MyoD, as a transactivator, stimulates transcription of itself. It promotes further expression of the G1 cyclin-dependent kinase (cdk) inhibitor p21 leading to permanent cell cycle withdrawal (Halevy et al. 1995; Parker, et al. 1995). It also stimulates expressions of other myogenic regulators (eg. myogenin) and skeletal muscle-specific genes, and further induces terminal differentiation (Sherr and Roberts 1995).

For transformation, E1A represses transactivation by MyoD-family transactivators, and thus represses skeletal muscle differentiation (Webster, et al. 1988; Caruso, et al. 1993). The amino-terminal p300-binding region of E1A is responsible for the repression of muscle differentiation (Caruso et al. 1993; Mymryk et al. 1992; Taylor et al. 1993). p300 is involved also, during keratinocyte differentiation, in the induction of the cdk inhibitor p21 (Missero et al. 1995), which, as stated above, is regulated by MyoD-dependent transactivation during muscle differentiation and is required for permanent cell cycle withdrawal. Thus, it is possible that p300 is involved in the differentiation of myocytes through modulation of MyoD-dependent transactivation.

I used E1A as an investigative tool to test this hypothesis. I studied the interaction between p300 and MyoD proteins, and demonstrated that p300 potentiates MyoD-dependent transactivation. To examine the transcription potential of p300, I also detected the regions of the p300 protein which contain transcriptional activity and which directly interact with the components of the basal transcriptional complex.

Materials and Methods

Cell Culture and Transient Transfection

The human osteogenic sarcoma cell line U-2 OS was maintained at 37°C in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS). The human rhabdomyosarcoma cell line RD was grown in DMEM with 4.5 g/L glucose and 10% FBS. The mouse muscle myoblast cell line C2C12 was maintained in growth medium (DMEM with 20% FBS) and allowed to differentiate to mature myotubes in differentiation medium (DMEM with 2% horse serum). The mouse teratocarcinoma cell line F9 was maintained in DMEM containing 15% FBS and 4.5 g/L glucose. The cells were stored in freezing medium (DMEM with 40% FBS and 10% DMSO) in liquid nitrogen.

U-2 OS and F9 cells were transiently transfected by the calcium phosphate precipitation method (Ausubel et al. 1989). Total amounts of transfected DNA for each condition were balanced by pUC19 DNA. After a 12-13 h incubation in medium containing the CaPO4-precipitated DNA, the cells were allowed to grow in the same medium for 48 h before harvesting.

CAT Assay

CAT assays were carried out as previously described (Ausubel et al. 1989). The cells were lysed in 200 µl of 250 mM Tris-Cl (pH 8.0) with three cycles of freezing on dry-ice and thawing at 42°C. The lysate was centrifuged for 5 min before use in assays. Cell lysates were normalized to β-galactosidase activity by cotransfection with 0.5-1 µg pON260 (an expression vector encoding β-galactosidase driven by the CMV promoter). In the β-gal assay, one twentieth of the transfected cell lysate (10 µl out of 200 µl) was incubated with 1 ml of 3 mM chlorophenol red β-D-galactopyranoside (CRGP, Boehringer Mannheim

Biochemicals) at 37°C for 1 h; the reaction was stopped by chilling on ice, and the optical density at 570 nm was then measured. In the CAT assay, the cell lysate was incubated at 37°C for 2-6 h with 2 µl of C^{14}-labeled chloramphenicol (NEN), 5-50 µl of 5 mg/ml acetyl CoA (Sigma), and 250 mM Tris-Cl (pH 8.0) in a final volume of 150 µl. The reaction was then extracted with 1 ml of ethyl acetate and dried for 45 min in a speed vacuum. The samples were resuspended in 20 µl of ethyl acetate, loaded on a 20x20 thin layer plate of Silica gel (IB2-F, J.T.Baker Inc.) (2 cm above the bottom and 1.5 cm apart), and developed for 1 hr in a solvent of chloroform/methanol (19:1). After air drying, the plate was exposed to X-ray film at -80°C for 1-2 days. The results were quantitated as % conversion in a phosphorimage analyzer. The relative CAT activity was calculated from the average of three independent experiments compared with the vector alone or with wildtype. Standard deviation (S.D.) was calculated by the equation: $[\Sigma(\chi-\Sigma\chi/N)^2/(N-1)]^{1/2}$. In this equation, χ is the value obtained in an experiment, and N is the number of independent experiments performed.

Plasmid Constructs

The pGal4p300 fusion vector was made from pCMVβp300 (Eckner et al. 1994) by inserting the Gal4 1-147 sequence into p300 cDNA at the position of the 19th amino acid, followed by subcloning this fusion unit into pcDNA3 (Invitrogen). The p300 deletions were made from pGal4p300 by digestion with restriction enzymes and by religation of the resulting deleted plasmids. The junctions in each construct were checked by sequencing. p4RTKCAT is a CAT reporter construct driven by the TK core promoter and four copies of the E box enhancer sequences upstream (Weintraub et al. 1990). pCMVβp300, pCMVMyoD (Halevy et al. 1995), pON260, and pCMV12S (Kannbiran et al. 1991) are expression vectors for p300, MyoD, β-galactosidase, and E1A 12S respectively, and are driven by the CMV immediate early promoter. dl 2-36 is derived from pCMV12S bearing

an amino-terminal deletion in the E1A 12S product between amino acids 2 and 36, which are required for the binding to p300. pG5e1bCAT is a CAT reporter driven by the E1b TATA box and by five copies of the GAl4-binding site upstream (Fields and Jang 1990). pGEX-MyoD, pGEX-hTBP, and pGST-IIB (Roberts et al. 1993) are bacterial expression vectors encoding the glutathione-S-transferase (GST) fusions to MyoD, to the TATA-binding protein (TBP), and to another general transcription factor, TFIIB.

Immunoprecipitation/Immunoblot

Whole cell lysates of C2C12 myoblasts were prepared and immunoprecipitated with a polyclonal anti-p300 antiserum, basically as previously described (Yaciuk and Moran 1991). The precipitations or whole cell lysates with equal amounts of total protein were suspended in 20 µl of Laemmeli buffer without boiling (because the molecular weights of MyoD and the immunoglobulin heavy chain are so close), separated on 10% SDS-polyacrylamide gel and electrotransferred onto PVDF membranes (Millipore) in a buffer containing 10 mM CAPS and 0-20% methanol, pH 11, for 1-8 h at 30-70 constant volts. The membrane was blocked with 5% nonfat dry milk in TBS-T buffer (20 mM Tris-Cl, pH 7.6, 137 mM NaCl, and 0.5% Tween 20), for 1 h at room temperature, incubated for 1 h with an anti-MyoD antibody (Santa Cruz), diluted in TBS-T with 3% nonfat dry milk, washed 3 times for 10 min each in TBS-T, further incubated for 1 h with a secondary antibody conjugated to horseradish peroxidase in TBS-T, and washed 3 times in TBS-T and once in TBS for 10 min each. The immunoreaction was visualized by incubating for 1 min in a chemiluminescent reagent (NEN) followed by exposure to X-ray film.

In vitro Transcription and Translation

Labeled p300 was produced from the pGal4p300 vector and its p300 deletion derivatives in a TNT-coupled reticulocyte system with T7 RNA polymerase and ^{35}S-labeled methionine (NEN), as described by the manufacturer (Promega).

GST Binding Assay

Expression from the pGEX-MyoD, pGEX-hTBP, and pGST-IIB plasmid in E. coli and immobilization of the GST fusion proteins are described by the manufacturer (Pharmacia). Binding assays were carried out basically as previously described (Kwok et al. 1994; Ha et al. 1993). Labeled proteins (roughly equal molar amount) were pre-incubated with glutathione-agarose beads in 0.5 ml of binding buffer (20 mM Tris-Cl, pH 8.0, 5 mM EDTA, 100 mM or 250 mM NaCl, and 0.1% or 0.5% NP-40) with 10 mg/ml BSA for 1 h at room temperature. Then, the affinity-purified GST fusion proteins (roughly equal molar amount) bound to glutathione-agarose beads were added, and incubated for another 1 h at room temperature. After washing the beads in the binding buffer 5 times, 1 ml each time,bead-bound proteins were eluted in 20 µl SDS sample buffer by boiling for 5 min, and analyzed by SDS-polyacrylamide gel electrophoresis followed by fluorography.

Results

p300 Protein is Involved in MyoD-Dependent Transactivation

An E box (MyoD consensus) reporter system (p4RTKCAT) was used to test the role of p300 in MyoD-activated transcription in U-2 OS human osteosarcoma cells. Ectopic MyoD can induces expression of other MyoD-family myogenic transactivators and differentiate further into myocytes in U-2 OS cells under a low serum level (Davis et al. 1987). In order to distinguish the E box activity mediated by MyoD from the activity mediated by other MyoD family transactivators, the experiments in this study were carried out in a high serum, non-differentiation condition. In this condition, as shown in Fig. 2-1, MyoD transactivated an appreciable level of E box reporter activity (lanes 2-4) although higher levels of MyoD transfected into the cell showed a reduced ability to transactivate the reporter construct. This transactivation for E box reporter is MyoD-specific because no

activity was detected without cotransfection with the MyoD vector (lane 1). This result indicates that endogenous p300 cannot transactivate the E box enhancer without MyoD and other MyoD-family transactivators. Consistent with this, ectopic p300 did not induce the reporter without co-transfection with MyoD, either (lane 5).

After checking the suitability of the system and the conditions to be used, E1A was used to probe whether p300 is involved in transactivation of E box reporter by MyoD. When an E1A 12S vector was cotransfected with the reporter and MyoD constructs into U-2 OS cells, E1A repressed MyoD-dependent E box activity in a dose-dependent manner (Fig. 2-2, lanes 1-4). This repression of the E box enhancer by E1A is consistent with the previous result that E1A repressed expression of muscle-specific genes. The E1A vector in this experiment was replaced with an amino terminal mutant (dl 2-36) in the E1A construct that loses its ability to bind to p300 but not to bind to Rb family proteins (Kannbiran et al. 1993; Giordano et al. 1991). This E1A mutant repressed E box reporter activity much less than the wild type E1A 12S (Fig. 2-2, lane 5-7). This result indicates that repression by E1A may require its binding to p300.

To confirm this result, the p300 cDNA construct was co-transfected into U-2 OS cells with the reporter, MyoD and E1A 12S vectors. Ectopic p300 largely released the E1A repression of MyoD-dependent transactivation (Fig. 2-3, lane 2 compared with lane 4), indicating the involvement of p300 in this E1A function. Further, when a p300 mutant 1-1737 without the E1A-binding domain and carboxyl terminal sequence replaced the wild type construct of p300 in the experiment, no repression of E box activity by E1A was detected (Fig. 2-3, lane 6). Thus, it appeares that association between E1A and p300 is required for the repression, and implies that p300 may be involved directly in MyoD transactivation. In addition, p300 mutant, in the absence of E1A, slightly reduced the level of MyoD transactivation (Fig. 2-3, lane 5). This reduction is probably related to defect in p300 function caused by deleting the carboxyl terminal region, and suggests that p300 plays

a role in MyoD transactivation. In the presence of E1A, the mutant 1-1737 not only bypassed the repression by E1A but even supported a higher level of MyoD transactivation (Fig. 2-3, lane 6). This stimulation by p300 mutant is probably associated with E1A's ability to indirectly disturb transcriptional activity of the p300 mutant that loses its E1A-binding domain (Arany et al. 1995).

Direct Interaction Between p300 and MyoD Proteins

Based on its proposed coactivator role, p300 may interact with MyoD directly to stimulate MyoD-dependent transactivation. To test this possibility, C2C12 cells were lysed in lysis buffer containing either 250 mM or 150 mM NaCl, immunoprecipitated with anti-p300 antibodies, and then immunoblotted with anti-MyoD antibodies. An interaction between p300 and MyoD proteins was detected under the low stringency salt condition (Fig. 2-4, lane 5). Further, in an in vitro binding assay, the in vitro translated p300 protein also bound to the purified bacterial GSTMyoD fusion protein linked to glutathione agarose beads (Fig. 2-5B, lane 3). This binding was associated with a specific affinity between p300 and MyoD because p300 did not show any affinity with GST alone (Fig. 2-5B, lane 2).

Through deletion analysis, the portion of p300 containing the specific affinity to MyoD was detected in a fragment 1514-1922 (Fig. 2-6, lane 1) that covers the third cysteine/histidine-rich conserved domain (C/H3), because it did not interact with GST alone and the control protein did not show any interaction with MyoD (Fig. 2-6, lane 6). The p300 fragment 1737-2414 containing a carboxyl terminal portion of the C/H3 domain was still able to bind to MyoD (Fig. 2-6, lane 2), while the most carboxyl terminal fragment of p300 1945-2414 was not (Fig. 2-6, lane 3). This affinity of p300 for MyoD was observed also in the fragment 1-1737 containing an amino portion of the C/H3 domain (Fig. 2-5C), which was able to bypass E1A repression of MyoD-dependent transactivation. The MyoD-binding domain of p300 is separate from the amino terminal portion (1-663) that binds to

CREB protein (Lundblad et al. 1995). Additionally, the amino terminal portions (1-743 and 1-596) of p300 do not have a specific binding affinity to MyoD (Fig. 2-6, lanes 4 and 5).

p300 Protein Modulates MyoD-Dependent Transactivation

The direct interaction between p300 and MyoD proteins supports the role of p300 as a coactivator in MyoD transactivation. To test this potentiality, transient transfection assays were performed using p300 and MyoD expression constructs to measure E box reporter activity in F9 teratocarcinoma cells. F9 cells can tolerate relatively high levels of p300 and MyoD, easily allowing the study of the role of p300 in MyoD-dependent transactivation. As shown in Fig. 2-7A and B, p300 potentiated the transcriptional activity of MyoD when increasing amounts of MyoD expression vector were cotransfected with and without p300 vector. As also shown in Fig. 2-8, p300 potentiated the transcriptional activity of MyoD when increasing amounts of the p300 vector were cotransfected with and without the MyoD vector. In both experiments, this function of p300 required the presence of MyoD. Similar results also were obtained in U-2 OS cells (Fig, 2-9A, lanes 1-3), although a large amount of ectopic p300 relative to the level of MyoD repressed the MyoD-dependent transactivation.

To confirm the role of p300 in MyoD-dependent transactivation, the p300 mutants that affected MyoD-dependent transactivation that had been potentiated by endogenous p300 were selected by using transient transfection assays in U-2 OS cells. Two mutants markedly affected MyoD-dependent transactivation, one (dl 242-1737) positively, and the other (1514-1922) negatively (Fig. 2-9A). The effects of these p300 mutants are associated with their disturbing the endogenous p300 function, because they did not show relevant effects on the activity of the promoter driving the ectopic MyoD expression or the MyoD level in the transfected cells (Fig. 2-9B).

p300 Protein Communicates with the Basal Transcription Complex Through Its Two Separate Transactivation Domains

The transactivation domain of the p300-related CBP protein resides at the carboxyl terminal glutamine-rich region (Kwok et al. 1994). Deletion of the corresponding region of p300 as in the mutant 1-1737 did not abolish its transactivation ability. To determine which portion of p300 allowed transactivation, p300 and its deletion mutants were fused to the Gal4 DNA-binding domain (1-147) (Fig. 2-10), and used for transient transfection assays in U-2 OS cells. As shown in Fig. 2-4, p300 wild type transactivated an appreciable level of Gal4-consensus reporter activity. This induction by the fusion protein must relate to p300 function because Gal4 1-147 alone did not show any transactivation. The most carboxyl terminal glutamine-rich portion of p300 (1945-2414) contributed a weak transactivation. When this portion was extended, as in the construct 1737-2414, the induction was increased eight-fold. The further addition of an amino terminal sequence 19-242 to the construct (dl 242-1737) allowed a nearly one-fold higher induction, but alone this amino terminal sequence in the construct (1-242) had no transactivation activity. The middle portions (964-1922 and 1514-1922) did not show any transactivation activity. The amino terminal portion of p300 (1-596) induced a level of transactivation similar to the carboxyl terminal portion (1737-2414), and about three-fold higher than the wild type. The addition of 148 amino acids in the construct (1-743) allowed an increase in transactivation by more than one-fold.

An additional question is whether the amino and carboxyl terminal transactivation domains of p300 interact directly with the basal transcription complex. Previous reports showed that p300 formed a complex with TBP in vivo (Abraham et al. 1993), and CBP interacted with TFIIB at the conserved C/H3 region (1680-1812) (Kwok et al. 1994). By using an in vitro binding assay, I detected an interaction between p300 protein and either TBP or TFIIB protein fused to GST (Fig. 2-11A, lanes 3 and 4). These interactions were specific because p300 did not interact with GST alone (Fig. 2-11A, lane 2). This specific

affinity of p300 to both TBP and TFIIB was further detected at its carboxyl terminal transactivation domain (1737-2414) that covered major sequence of the C/H3 region (Fig. 2-11B). The amino terminal portion of p300 (1-743), inducing a very high level of transactivation, showed a specific binding to TBP but not to TFIIB (Fig. 2-11C).

Discussion

A new type of transcriptional regulator has been classified recently as the transcriptional coactivator that communicates between transactivators binding on distal enhancers and the basal transcriptional complex formed on the promoter near the transcription initiation site (Gill and Tjian, 1992). This communication is required for transcriptional activation by the transactivators. p300 protein is a potential transcriptional coactivator because its sequence contains a bromodomain which is a known indicator for transcription adapter proteins (Eckner et al. 1994). This potentiality is supported by several lines of evidence demonstrating that p300 can interact with the transactivator CREB and potentiate its transactivation (Lundblad et al. 1995), that a fusion of the E2 DNA-binding domain to p300 transactivates an E2 binding site reporter (Arany et al. 1995), and that p300 can form a complex in vivo with TBP, a component of the basal transcriptional complex (Abraham et al. 1993).

I further demonstrated that p300 is a transcriptional coactivator for MyoD. The analysis with E1A and the potentiation of MyoD-dependent transactivation by p300 indicate that p300 participates in transactivation mediated by the interaction between the E box enhancer and MyoD. The interactions between MyoD and p300 proteins, and between p300 and either TBP or TFIIB, suggest that p300 can communicate directly between the E box-binding MyoD and the basal transcriptional complex. This feature of p300 defines it as a transcriptional coactivator for MyoD.

The association of p300 with MyoD was only detected with low stringency salt conditions in vivo. This result is consistent with the fact that interactions of transcription factors are generally weak and transient, and that it is difficult to detect them in vivo except by use of a two hybrid system. The association of MyoD with Rb was detected also using a similar low stringency salt condition (Gu et al. 1993). This condition is created by a lysis buffer containing a 150 mM (rather than a 250 mM) salt which is closer to the physiological salt condition in the cell.

A transcription factor communicates with the basal complex through its transactivation domain. Separate transactivation activities of p300 were detected in the amino and carboxyl terminal regions. The interactions between the amino and carboxyl terminal regions of p300 and either TBP or TFIIB in vitro indicate that p300 can associate with the basal complex directly. However, these results cannot rule out an indirect interaction of p300 with the basal transcriptional complex simultaneously, as was reported by Abraham et al. (1993) when they found that the interaction of p300 with TBP may be mediated by two 64- and 59-kD proteins. It is also possible that p300 interacts with and transactivation is mediated by other transcriptional factors (Goodrich et al. 1993; Lee et al. 1995).

The dominant negative and dominant positive mutants of p300 are shown, in this study, to affect endogenous p300 function in MyoD-dependent transactivation of the E box reporter. Because the dominant negative mutant 1514-1922 is able to bind to MyoD but has no transcriptional activity, its negative role may be associated with sequestering MyoD from the wild type p300 in vivo. The dominant positive mutant dl 242-1737 has a high transactivation ability and lacks a large middle portion, including the bromodomain and the second cysteine/histidine-rich region. Results of other Gal4 reporter assays imply a negative role for this middle portion of p300 in transactivation (unpublished data). Thus, the mutant dl 242-1737 could bypass repressor involvement. It also contains the MyoD-

binding domain, and could therefore deliver its higher transactivation potential to the E box via MyoD. These data further support the model that p300 serves as a coactivator for the myogenic transactivator MyoD.

All MyoD family transactivators can transactivate the E box enhancer. Under muscle differentiation conditions, MyoD induces the expression of myogenin. In order to distinguish the E box activity mediated by MyoD from that of other myogenic transactivators, I carried out the experiments in nonmuscle cells under nondifferentiation conditions. However, I discovered that the dominant positive and negative mutants of p300 affected E box reporter activity during MyoD-induced differentiation in a way similar to that detected in a nondifferentiation condition (unpublished data). The effect of p300 mutants on E box activity without ectopic MyoD was also observed in C2C12 myoblasts and human rhabdomyosarcoma cells under low serum conditions (unpublished data). It appears that p300 modulates transactivation by MyoD, and possibly by other MyoD family transactivators, under myogenic differentiation conditions. Consistent with this, I observed that the p300 mutant 1-1737 bypassed E1A repression of myogenin-induced E box reporter activity (unpublished data), indicating a potential role of p300 in myogenin-dependent transactivation.

The function of p300 as a coactivator for the MyoD family transactivators suggests one of the mechanisms by which it might be involved in the regulation of muscle differentiation. This mechanism probably plays an important role because, for efficient repression of muscle-specific gene expression and differentiation, the E1A oncoprotein must have the ability to attack both p300 and Rb family proteins. This does not rule out the possibility that repression of muscle-specific gene expression by E1A is correlated also with its direct interaction with myogenin (Taylor et al. 1993), which might be another pathway for E1A to inhibit the myogenic process. It has been shown that p300 protein is involved during keratinocyte differentiation in the induction of the cyclin-dependent kinase

inhibitor p21 (Missero et al. 1995), which is regulated by MyoD and is required for permanent cell cycle withdrawal during muscle differentiation (Halevy et al. 1995; Parker et al. 1995). This result further supports the proposition that p300 participates in myogenesis through the modulation of MyoD-dependent transcription.

CHAPTER 4

EPILOGUE

For many years, a major question about eucaryotic gene expression has been how transcription of a specific gene is activated and regulated by cis-regulatory elements and trans-activating factors (Buratowski 1995). Transcription regulation of eucaryotic genes is controlled largely and has been intensively studied, at the level of initiation. The subsequent step of transcription elongation offers an additional way of regulating transcription and gene expression. The importance and the detailed mechanisms of transcription control at the level of elongation are well known in procaryotes. It was found recently that an increasing number of viral and cellular genes in eucaryotes are regulated mainly by conditional arrest during transcription elongation. These genes include many important proto-oncogenes, such as c-myc, c-fos, c-myb, and so on.

Previous nuclear run-on studies suggested that the expression of c-myb was controlled by a block in the first intron during transcription elongation (Bender et al. 1987; Watson 1988a and b). The nature of the block was further defined as premature termination in this study. The precise site Tm for premature termination of transcription elongation of c-myb was mapped to a site Tm in the middle of intron 1, approximately 1710 bp from the start of the intron by using Xenopus oocyte system. The same Tm site was used also in 7OZ/3B cells, and it correlated with the region of the transcription block mapped previously by nuclear run-on assays in different cell types (Bender et al. 1987; Watson 1988a) and in the hematopoietic cells that had been induced to terminally differentiate (Watson 1988b). In contrast to the situations found in procaryotic and viral genes, I did not find a stem-loop structure at the Tm site that might specify transcription

termination. Instead, Tm is intrinsic to the intron sequence, and is independent of the promoter sequence.

It is believed that conditional arrest in transcription elongation can control the level of mRNA in eucaryotic cells efficiently. This kind of control can cause the differential expression of a gene, such as c-myc, c-myb, or c-fos etc., under different physiological conditions and in different tissues (Spencer and Groudine 1990). It was observed previously, by using nuclear run-on assays, that the tissue-specific expression and down-regulation of c-myb correlated with the transcription block during elongation. This study demonstrated that a transcription block in elongation caused differential transcription of the c-myb sequences in nuclear extracts prepared from different cell types. This differential transcription was regulated by the intron sequences that controlled premature termination at, and attenuation through, the Tm sites. The regulatory sequences that directly control the premature termination at Tm have mapped to a domain in the middle of intron 1. A region upstream of the domain showed an antitermination activity in the different systems used in this study. Thus, the sequence requirement for transcription regulation of the c-myb gene is analogous to that in the late early genes of the bacteriophage λ, in which transcription elongation is blocked at sites far downstream from the promoter, and the sequence for antitermination (nut site) resides between the promoter and the terminators.

However, many questions remain. How does the intron termination domain specify transcription termination at Tm? Are the nature of the DNA sequence and/or secondary structures of transcribed RNA sufficient for this specification? Are there any nuclear proteins, functionally homologous to the procaryotic termination factor Rho, involved in this event? What kind of nuclear factors are involved directly in antitermination in c-myb transcription regulation? Are they DNA-binding proteins or RNA-binding proteins (like λ N protein for the function of the nut site), or both? Because promoter sequences have been found to play an important role in transcription attenuation in many eucaryotic genes, does

the c-myb promoter play a role in the regulation of transcription elongation at the far downstream Tm sites, and if so, how?

When I replaced the 5' untranslated sequence of c-myb, including the promoter, with either TK or CMV promoter in TKC3 or CMVC3 construct, CMVC3, but not TKC3, broke the transcription block that was controlled directly by the intron regulatory sequences and allowed a high level of the reporter expression in NIH3T3 cells, in which c-myb expression is normally repressed. There are two potential causes for this situation. First, certain transactivators in NIH3T3 cells binding to the CMV enhancers largely increase transcription processivity to allow pol II to read through the downstream Tm site. This potentiation has been supported by evidence that certain transactivators stimulate transcription elongation by pol II (Yankulov et al. 1994). Second, and more importantly, the nature of the basal transcription complex formed on the CMV promoter, including pol II, general transcription factors, and other components of the holoenzyme, may be different from those formed on the c-myb and TK promoters. Thus, the transcription machinery, pre-settled on certain promoters, could have laid a foundation for the possibility of certain events to occur not only in initiation but also in subsequent steps of transcription process. Study of this mechanism is crucial to understand clearly how a specific transcription could be regulated.

Another problem in studying specific transcription activation is the involvement of cofactors, either positive coactivators or negative corepressors. The role of cofactors in communication between transactivators and the basal transcription machinery during transcription activation and repression is only beginning to be investigated now. In this study, I analyzed the coactivator function of human p300 protein in transactivation by the myogenic transactivator MyoD. I detected a direct interaction between p300 and either MyoD, TBP or TFIIB, and proposed that p300 functions as a coactivator for MyoD by allowing MyoD to communicate with the basal complex during transcription modulation.

However, it is probable that p300 faces a much more complicated situation in vivo. Its presumed huge surface with multiple conserved domains may interact with many other components in the nucleus. The C/H3 domain that has an affinity for MyoD was shown to interact also with several other factors, such as YY1 (Lee, et al. 1995), E1A (Eckner, et al. 1994), TFIIB and TBP. It was reported previously that two other major polypeptides also existed in the p300-TBP complex in vivo (Abraham et al. 1993). Therefore, it is quite possible that the p300 coactivator function in MyoD-dependent transactivation may be mediated or regulated by other nuclear factors. We know so little about both the structure and function of p300. How do the conserved sequences of p300 contribute to its function? What is the role of the bromodomain which is conserved in many other transcription cofactors? How important is the coactivator function of p300 in muscle differentiation? Are there any other mechanisms that may be used by p300 to stimulate the differentiation process? All these questions await further investigation.

REFERENCES CITED

Abraham, S. E., S. Lobo, P. Yaciuk, H.-G. H. Wang, and E. Moran (1993). p300, and p300-associated proteins, are components of TATA-binding protein (TBP) complexes. Oncogene 8, 1639-1647.

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson (1989). Control of gene expression. In Molecular Biology of the Cell, Chapter 10, 2nd Edition, p. 551-612, Garland Publishing, Inc., New York.

Arany, Z., W. R. Sellers, D. M. Livingston, and R. Eckner (1994). E1A-associated p300 and CREB-associated CBP belong to a conserved family of coactivators. Cell 77, 799-800.

Arany, Z., D. Newsome, E. Oldread, D. M. Livingston, and R. Eckner (1995). A family of transcriptional adaptor proteins targeted by the E1A oncoprotein. Nature 374, 81-84.

Aso, T., W. S. Lane, J. W. Conaway, and R. C. Conaway (1995). Elongin (SIII): a multisubunit regulator of elongation by RNA polymerase II. Science 269, 1439-1443.

Ausubel, F. M., R. Brent, R. E. Kingston, D. Moore, J. G. Seidman, J. A. Smith, K. Struhl (eds) (1989). in Current Protocols in Molecular Biology, 2nd Edition, p. 9.1 and 9.6, Green and Wiley-Interscience, New York.

Baluda, M. A., and I. E. Goetz (1961). Morphological conversion of cell cultures by avian myeloblastosis virus. Virology 15, 185-199.

Bayley, S. T., and J. S. Mymryk (1994). Adenovirus E1A proteins and transformation. Int. J. Oncology 5, 425-444.

Bender, T. P., and W. M. Kuehl (1986). Murine myb protooncogene mRNA: cDNA sequence and evidence for 5' heterogeneity. Proc. Natl. Acad. Sci., USA 83, 3204-3208.

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.

Bentley, D. L. (1995). Regulation of transcriptional elongation by RNA polymerase II. Curr. Opin. Genet. Dev. 5, 210-216.

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 321, 702-706.

Bentley, D. L., and M. Groudine (1988). Sequence requirements for premature termination of transcription in the human c-myc gene. Cell 53, 245-256.

Berger, S. L., B. Pina, N. Silverman, G. A. Marcus, J. Agapite, J. L. Regier, S. J. Triezenberg, and L. Guarente (1992). Genetic isolation of ADA2: a potential

transcriptional adapter required for function of certain acidic activation domains. Cell 70, 251-265.

Biedenkapp, H., U. Borgmeyer, A. E. Sippel, and K. H. Klempnauer (1988). Viral m-b oncogene encodes a sequence-specific DNA-binding activity. Nature 335, 835-837.

Buratowski, S. (1995). Mechanisms of gene activation. Science 270, 1773-1774.

Caruso,M., F. Martelli, A. Giordano, and A. Felsani (1993). Regulation of MyoD gene transcription and protein function by the transformation domains of the adenovirus E1A oncoprotein. Oncogene 8, 267-278.

Castellano, M., J. Goley, A. Mantovani, and M. Introna (1992). Detection of a transcriptional block in the first intron of the human c-myb gene. Int J Clin Lab Res 22, 159-164.

Chamberlin, M. (1994). New models for the mechanism of transcription elongation and its regulation. Harvey Lectures 88, 1-21.

Chinsky, J. M., M.-C. Maa, V. Ramamurthy, and R. E. Kellems (1989). Adenosine deaminase gene expression. J. Biol. Chem. 264, 14561-14565.

Chodosh, L. A., R. W. Carthew, and P. A. Sharp (1986). A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. Mol. Cell. Biol. 6, 4723-4733.

Coulombe, B., A. Ponton, L. Daigneault, B. R. G. Williams, and D. Skup (1988). Presence of transcription regulatory elements within an intron of the virus-inducible murine TIMP gene. Mol. Cell. Biol. 8, 3227-3234.

Cullen, B. R. (1986). Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell 46, 973-982.

Davis, R. L., H. Weintraub, and A. B. Lassar (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51, 987-1000.

Dignam, J. D., R. M. Lebovitz, and R. G. Roeder (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucl. Acids Res. 11, 1475-1489.

Duan, D. R., A. Pause, W. H. Burgess, T. Aso, D. Y. T. Chen, K. P. Garrett, R. C. Conaway, J. W. Conaway, W. M. Linehan, and R. D. Klausner (1995). Inhibition of transcription elongation by the VHL tumor suppressor protein. Science 269, 1402-1406.

Duprey, S. P., and D. Boettiger (1985). Developmental regulation of c-myb in normal myeloid progenitor cells. Proc. Natl. Acad. Sci., USA 82, 6937-6941.

Eckner, R., M. E. Ewen, D. Newsome, M. Gerds, J. A. DeCaprio, J. B. Lawrence, D. M. Livingston (1994). Molecular cloning and functional analysis of the adenovirus E1A-

associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. Genes Dev. 8, 869-884.

Fields, S., and S. K. Jang (1990). Presence of a potent transcription activating sequence in the p53 protein. Science 24, 1046-1049.

Flanagan, P. M., R. J. Kelleher, M. H. Sayre, H. Tschochner, and R. D. Kornberg (1991). A mediator required for activation of RNA polymerase II transcription in vitro. Nature 350, 436-438.

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. Nucl. Acids Res. 15, 5657-5667.

Fried, M., and D. M. Crothers (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucl. Acids Res. 9, 6505-6525.

Friedman, D. I. (1988). Regulation of phage gene expression by termination and antitermination of transcription. In The Bacteriophages Vol. 2 (ed. Calendar, R.), pp 263-318, Plenum, New York.

Garner, M. M., and A. Revzin (1981). A gel electrophoresis method for qualifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. Nucl. Acids Res. 9, 3047-3060.

Garret, K. P., S. Tan, J. N. Bradsher, W. S. Lane, J. W. Conaway, and R. C. Conaway (1994). Molecular cloning of an essential subunit of RNA polymerase II elongation factor SIII. Proc. Natl. Acad. Sci. USA 91, 5237-5241.

Gick, O., A. Kramer, W. Keller, and M. L. Birnstiel (1986). Generation of histone mRNA 3' ends by endonucleolytic cleavage of the pre-mRNA in an snRNP-dependent in vitro reaction. EMBL J. 5, 1319-1326.

Gill, G., and R. Tjian (1992). Eukaryotic coactivators associated with the TATA box binding protein. Curr. Opin. Genet. Dev. 2, 236-242.

Giordano, A., C. McCall, P. Whyte, and B. R. Franza, Jr. (1991). Human cyclin A and the retinoblastoma protein interact with similar but distinguishable sequences in the adenovirus E1A gene product. Oncogene 6, 481-485.

Gonda, T. J., D. K. Sheiness, and J. B. Bishop (1982). Transcripts from cellular homologs of retroviral oncogenes: distribution among chicken tissues. Mol. Cell. Biol. 2, 617-624.

Goodrich, J. A., T. Hoey, C. J. Thut, A. Admon, and R. Tjian (1993). Drosophila TFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell 75, 519-530.

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.

Graf, T. (1992). Myb: a transcriptional activator linking proliferation and differentiation in hematopoietic cells. Cur. Opin. Gen. Dev. 2, 249-255.

Graham, F. L., and A. J. van der Eb (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52, 46.

Green, M. R. (1991). Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Ann. Rev. Cell Biol. 7, 559-599.

Greenblatt, J., J. R. Nodwell, and S. W. Mason (1993). Transcriptional antitermination. Nature 364, 401-406.

Gu, W., J. W. Schneider, G. Condorelli, S. Kaushal, V. Mahdavi, and B. Nadal-Ginard (1993). Interaction of myogenic factors and the retinoblastoma protein mediates muscle cell commitment and differentiation. Cell 72, 309-324.

Guerin, M., Z.-M. Sheng, N. Andrieu, and G. Riou (1990). Strong association between c-myb and oestrogen-receptor expression in human breast cancer. Oncogene 5, 131-135.

Guerra, J., D. A. Withers, and L. M. Boxer (1995). Myb binding sites mediate negative regulation of c-myb expression in T-cell lines. Blood 86, 1873-1880.

Ha, I., S. Robert, E. Maldonads, X. Sun, L.-U. Kim, M. Green, and D. Reinberg (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes Dev. 7, 1021-1032.

Hair, A., and G. T. Morgan (1993). Premature termination of tubulin gene transcription in Xenopus oocytes is due to promoter-dependent disruption of elongation. Mol. Cell. Biol. 13, 7925-7934.

Halevy, O., B. G. Novitch, D. B. Spicer, S. X. Skapek, J. Rhee, G. J. Hannon, D. Beach, A. B. Lassar (1995). Correlation of terminal cell cycle arrest of skeletal muscle with induction of p21 by MyoD. Science 267, 1018-1021.

Higuchi, R., B. Krummel, and R. K. Saiki (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucl. Acids Res. 16, 7351-7367.

Hoey, T., R. O. J. Weinzierl, G. Gill, J.-L. Chen, B. D. Dynlacht, and R. Tjian (1993). Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivator. Cell 72, 247-160.

Jacobs, S. M., K. M. Gorse, and E. H. Westin (1994). Indentification of a second promoter in the human c-myb proto-oncogene. Oncogene 9, 227-235.

Jakobovits, A., D. H. Smith, E. B. Jakobovits, and D. J. Capon (1988). A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. Mol. Cell. Biol. 8, 2555-2561.

Jonsson, J. J., M. D. Foresman, N. Wilson, and R. S. McIvor (1992). Intron requirement for expression of the human purine nucleoside phosphorylase gene. Nucl. Acids Res. 20, 3191-3198.

Johnson, T., and M. Chamberlin (1994). Complexs of yeast RNA polymerase II and RNA are substracts for TFIIS-induced RNA cleavage. Cell 77, 217-224.

Kannbiran, C., G. F. Morris, C. Labrie, and M. B. Mathews (1993). The adenovirus E1A 12S product displays functional redundancy in activating the human proliferating cell nuclear antigen promoter. J. Virol. 67, 507.

Kao S.-Y., A. F. Calmen, P. A. Luciw, and B. M. Peterlin (1987). Antitermination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature 330, 489-493.

Karpinski, B. A., L.-H. Yang, P. Cacheris, G. D. Morle, and J. M. Leiden (1989). The first intron of the 4F2 heavy-chain gene contains a transcriptional enhancer element that binds multiple nuclear proteins. Mol. Cell. Biol. 9, 2588-2597.

Kerppola, T. K., C. M. Kane (1988). Intrinsic sites of transcription termination and pausing in the c-myc gene. Mol. Cell. Biol. 8, 4389-4394.

Kessler, M., E. Ben-Asher, and Y. Aloni (1989). Elements modulating the block of transcription elongation at the adenovirus 2 attenuation site. J. Biol. Chem. 264, 9785-9790.

Kibel, A., O. Iliopoulos, J. A. DeCaprio, and W. G. Kaelin Jr. (1995). Binding of the von Hippel-Lindau tumor suppressor protein to elongin B and C. Science 269, 1444-1446.

Kitajima, S., T. Chibazzakura, M. Yonaha, and Y. Yasukochi (1994). Regulation of the human general transcription initiation factor TFIIF by phosphorylation. J. Biol. Chem. 269, 29970-29977.

Kirsh, I. R., V. Bertness, J. Silver, and G. F. Hollis (1985). In Leukemia: Recent Advances in Biology and Treatment (eds. R.P. Gale and D.W. Golde), pp. 91-98, Liss, New York.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. 47, 1-45.

Krumm, A., T. Meulia, M. Brunvand, and M. Groudine (1992). The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. Genes Dev. 6, 2201-2213.

Krumm, A. and M. Groudine (1995). Tumor suppression and transcription elongation: the dire consequences of changing partners. Science 269, 1400-1401.

Kuhn, A., and I. Grummt (1989). 3’-end formation of mouse pre-rRNA involves both transcription termination and a specific processing reaction. Genes Dev. 3, 224-231.

Kwok, R. P. S., J. R. Lundblad, J. C. Chrivia, J. P. Richards, H. P. Bachinger, R. G. Brennan, S. G. E. Roberts, M. R. Green, R. H. Goodman (1994). Nuclear protein CBP is a coactivator for the transcription factor CREB. Nature 370, 223-226.

Lassar, A. B., S. X. Skapek, and B. Novitch (1994). Regulatory mechanisms that coordinate skeletal muscle differentiation and cell cycle withdrawal. Curr. Opin. Cell Biol. 6, 788-794.

Lattier, D. L., J. C. States, J. J. Hutton, and D. A. Winginton (1989). Cell type-specific transcriptional regulation of the adenosine deaminase gene. Nucl. Acids Res. 17, 1061-1076.

Lavu, S., and E. P. Reddy (1986). Structure organization and nucleotide sequence of mouse c-myb oncogene: activation in ABPL tumors is due to viral integration in an intron which results in the deletion of the 5' coding sequences. Nucl. Acid Res. 14, 5309-5320.

Lee, J.-S., K. M. Galvin, R. H. See, R. Eckner, D. M. Livingston, E. Moran, Y. Shi (1995). Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. Genes Dev. 9, 1188-1198.

Lee, H., K.W. Kraus, M. F. Wolfner, and J. T. Lis (1992). DNA sequence requirements for generating paused polymerase at the start of hsp70. Genes Dev. 6, 284-295.

Lee, W. S., C. C. Kao, G. O. Bryant, and X. A. J. Liu (1991). Adenovirus E1A activation domain binds the basic repeat in the TATA box transcription factor. Cell 67, 365-376.

Lieberman, P. M. and A. J. Berk (1991). The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction. Genes Dev. 5, 2441-2454.

London, L., R. G. Keene, and R. Landick (1991). Analysis of premature termination in c-myc during transcription by RNA polymerase II in a HeLa nuclear extract. Mol. Cell. Biol. 11, 4599-4615.

Lopata, M. A., D. W. Cleveland, and B. Sollner-Webb (1984). High-level expression of a chloramphenicol acetyltransferase gene by DEAE-dextran-mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. Nucl. Acid Res. 12, 5707.

Lundblad, J. R., R. P. S. Kwok, M. E. Laurance, M. L. Harter, R. H. Goodman (1995). Adenoviral E1A-associated protein p300 as a functional homologue of the transcriptional co-activator CBP. Nature 374, 85-88.

Luscher, B., and R. N. Eisenman (1990). New light on Myc and Myb. Part II. Myb. Genes Dev. 4, 2235-2241.

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.

Marzluff, W. F., and R. C. C. Huang (1984). Transcription of RNA in isolated nuclei. In Transcription and translation: A practical approach. (eds. B.D. Hames and S.J. Higgins), pp 89-101, IRL Press, Washington DC.

McCann, S., J. Sullivan, J. Guerra, M. Arcinas, and L. M. Boxer (1995). Repression of the c-myb gene by WT1 protein in T and B cell lines. J. Biol. Chem. 270, 23785-23789.

McKeown, M. (1993). The role of small nuclear RNAs in RNA splicing. Cur. Opinion in Cell Biol. 5, 448-454.

Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acid Res. 12, 7035-7056.

Meulia, T., A. Krumm, C. Spencer, and M. Groudine (1992). Sequences in the human c-myc P2 promoter affect the elongation and premature termination of transcripts initiated from the upstream P1 promoter. Mol. Cell. Biol. 12, 4590-4600.

Mikscek, R., A. Heber, W. Schmid, U. Danesch, G. Posseckert, M. Beato, and G. Schutz (1986). Glucocorticoid responsiveness of the transcriptional enhancer of Moloney Murine Sarcoma Virus. Cell 46, 283-290.

Missero, C., E. Calautti, R. Eckner, J. Chin, L. H. Tsai, D. M. Livingston, and G. P. Dotto (1995). Involvement of the cell-cycle inhibitor Cip1/WAF1 and the E1A-associated p300 protein in terminal differentiation. Proc. Natl. Acad. Sci. USA 92, 5451-5455.

Moore, C. L., and P. A. Sharp (1985). Accurate cleavage and polyadenulytion of exogenous RNA substrate. Cell 41, 845-855.

Moran, E. (1993) DNA tumor virus transforming protein and the cell cycle. Curr. Opin. Genet. Dev. 3, 63-70.

Moran, E. (1994). Mammalian cell growth controls reflected through protein interactions with the adenovirus E1A gene products. Semin. Virol. 5, 327-340.

Moscovici, C. (1975). Leukemic transformation with avian myeloblastosis virus: present status. Curr. Top. Microbiol. Immunol. 71, 79-101.

Mount, S. M., C. Burks, G. Hertz, G. D. Stormo, O. White, and C. Fields (1992). Splicing signals in Drosophila: intron size, and consensus sequences. Nucl. Acids Res. 20, 4255-4262.

Mymryk, J. S., R. W. H. Lee, and S. T. Bayley (1992). Ability of adenovirus 5 E1A proteins to suppress differentiation of BC3H1 myoblasts correlates with their binding to a 300 kDa cellular protein. Mol. Biol. Cell. 3, 1107-1115.

Nelson, K. K., and M. R. Green (1989). Mammalian U2 snRNP has a sequence-specific RNA-binding activity. Genes Dev. 3, 1562-1571.

Ness, S. A., A. Marknell, and T. Graf (1989). The v-myb oncogene product binds to and activates the promyelocyte-specific mim-1 gene. Cell 59, 1115-1125.

Neuman de Vagvar, H. E., E. Lund, and J. E. Dahlberg (1986). 3' end formation of U1 snRNA precursors is coupled to transcription from snRNA promoters. Cell 24, 259-266.

Nicolaides, N. C., R. Gualdi, C. Casadevall, L. Manzella, and B. Calabretta (1991). Positive autoregulation of c-myb expression via Myb binding sites in the 5' flanking region of the human c-myb gene. Mol. Cell. Biol. 11, 6166-6176.

Nicolaides, N. C., I. Correa, C. Casadevall, S. Travali, K. J. Soprano, and B. Calabretta (1992). The Jun family members, c-Jun and JunD, transactivate human c-myb promoter via an AP-1 like element. J. Biol. Chem. 267, 19665-19672.

Nomura, N., M. Takahashi, M. Matsui, S. Ishii, T. Date, S. Sasamoto, and R. Ishizaki (1988). Isolation of human cDNA clones of myb-related genes, A-myb and B-myb. Nucl. Acid Res. 16, 11075-11089.

Olson E. N., and W. H. Klein (1994). bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. Genes Dev. 8, 1-8.

Oshima, R. G., L. Abrams, and D. Kulesh (1990). Activation of an intron enhancer within the keratin 18 gene by expression of c-fos and c-jun in undifferentiated F9 embryonal carcinoma cells. Genes Dev. 4, 835-848.

Pabo, C. O. and R. T. Sauer (1992). Transcription factors: structural families and principles of DNA recognition. Ann. Rev. Biochem. 61, 1053-1095.

Palmiter, R. D., E. P. Sandgren, M. R. Avarbock, D. D. Allen, and R. L. Brinster (1991). Heterologous introns can enhance expression of transgenes in mice. Proc. Natl. Acad. Sci., USA 88, 478-482.

Parker, S. B., G. Eichele, P. Zhang, A. Rawls, A. Sands, A. Bradley, E. N. Olson, J. W. Harper, and S. J. Elledge (1995). p53-independent expression of p21 in muscle and other terminally differentiating cells. Science 267, 1024-1027.

Proudfoot, N. J., and E. Whitelaw (1988). Transcription and splicing. In Frontiers in Molecular Biology (eds. B.D. Hames and D.M. Glover), Vol. 7, pp. 97-129, IRL Press.

Ramamurthy, V., M.-C. Maa, M. Harless, D. Wright, and R. E. Kellems (1990). Sequence requirements for transcriptional arrest in exon 1 of the mouse adenosine deaminase genes. Mol. Cell. Biol. 10, 1484-1491.

Reddy, C. D., and E. P. Reddy (1989). Differential binding of nuclear factors to the intron 1 sequences containing the transcriptional pause site correlates with c-myb expression. Proc. Natl. Acad. Sci. USA 86, 7326-7330.

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, 12-17.

Resnekov, O., M. Kessler, and Y. Aloni (1989). RNA secondary structure is an integral part of the in vitro mechanism of attenuation in simian virus 40. J. Biol. Chem. 264, 9953-9959.

Ring, B. Z., and J. W. Roberts (1994). Function of a nontranscribed DNA strand site in transcription elongation. Cell 78, 317-324.

Roberts, S. G. E., I. Ha, E. Maldonado, D. Reinberg, and M. R. Green (1993). Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. Nature 363, 741-744.

Rosbash, M., and R. H. Singer (1993). RNA travel: tracks from DNA to cytoplasm. Cell 75, 399-401.

Rougvie, A. E., and J. T. Lis (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell 54, 795-804.

Roussel, M., S. Saule, S. Langrow, H. Rommens, H. Berg, T. Graf, and D. Stehelin (1979). Three new types of viral oncogene of cellular origin specific for hematopoietic cell transformation. Nature 281, 452-455.

Rudd, M. D., M. G. Izban, and D. S. Luse (1994). The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. Proc. Natl. Acad. Sci. USA 91, 8057-8061.

Santhanam, U., A. Ray, and P. B. Sehgal (1991). Repression of the interleukin 6 gene promoter by p53 and the retinoblastoma susceptibility gene product. Proc. Natl. Acad. Sci., USA 88, 7605-7609.

Sawadogo, M. and A. Sentenac (1990). RNA polymerase B (II) and general transcription factors Ann. Rev. Biochem. 59, 711-754.

Schneider-Schaulies, J., A. Schimpl., and E. Wecker (1987). Kinetics of cellular oncogene expression in mouse lymphocytes. II. Regulation of c-fos and c-myc gene expression. Eur. J. Immunol. 17, 713-718.

Selby, M. J., 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.

Shen-Ong, G. L. C. (1990). The myb oncogene. Biochim. Biophys. Acta 1032, 39-52.

Sherr, G. L., and J. M. Roberts (1995). Inhibitors of mammalian G1 cyclin-dependent kinases. Genes Dev. 9, 1149-1163.

Slamon, D. J., T. C. Boone, D. C. Murdock, D. E. Keith, M. F. Press, R. A. Larson, and L. M. Souza (1986). Studies of human c-myb and its products in human acute leukemias. Science 233, 347-351.

Sobieszczuk, P. W., T. J. Gonda, and A. R. Dunn (1989). Structure and biological activity of the transcriptional initiation sequences of the murine c-myb oncogene. Nucl. Acid Res. 17, 9593-9611.

Spencer, C. A., and M. Groudine (1990). Transcription elongation and eucaryotic gene regulation. Oncogene 5, 777-785.

Stringer, K. F., C. J. Ingles, J. Greenblat (1990). Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature 345, 783-786.

Strobl, L., and D. Eick (1992). Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-myc in vivo. EMBO J. 11, 3307-3314.

Taylor, D. A., V. B. Kraus, J. J. Schwarz, E. N. Olson, W. E. Kraus (1993). E1A-mediated inhibition of myogenesis correlates with a direct physical interaction of E1A12S and basic helix-loop-helix proteins. Mol. Cell. Biol. 13, 4714-4727.

Thomas, P. S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77, 5201.

Toth, C. R., Hostutler, R. F., Baldwin, A. S., and Bender, T. P. (1995). Members of the nuclear factor B family transactivate the murine c-myb gene. J. Biol. Chem. 270, 7661-7671.

Triezenberg, S. J. (1995). Structure and function of transcriptional activation domains. Curr. Opin. Genet. Dev. 5, 190-196.

Wang, J., L. G. Cao, Y.-L. Wang, and T. Pederson (1991). Localization of pre-messenger RNA at discrete nuclear sites. Proc. Natl. Acad. Sci., USA 88, 7391-7395.

Watson, R. J. (1988a). A transcriptional arrest mechanism involved in controlling constitutive levels of mouse c-myb mRNA. Oncogene 2, 267-272.

Watson, R. J. (1988b). Expression of c-myb and c-myc genes is required independently in differentiating mouse erythroleukemia cells by common processes of premature transcription arrest and increased mRNA turnover. Mol. Cell. Biol. 8, 3938-3942.

Watson, R. J., P. J. Dyson, and J. McMahon (1987). Multiple c-myb transcript cap sites are variously utilized in cells of mouse hematopoietic origin. EMBO J. 6, 1643-1651.

Webster, K. A., G. E. Muscat, and L. Kedes (1988). Adenovirus E1A products suppress myogenic differentiation from muscle-specific promoters. Nature 332, 553-557.

Weintraub, H., R. Davis, D. Lockshon, and A. Lassar (1990). MyoD binds cooperatively to two sites in a target enhancer sequence: occupancy of two sites is required for activation. Proc. Natl. Acad. Sci. USA 87, 5623-5627.

Westin, E. H., R. C. Gallo, S. K. Arya, A. Eva, L. M. Souza, M. Baluda, S. A. Aronson, and F. Wong-Staal (1982). Differential expression of the AMV gene in human hematopoietic cells. Proc. Natl. Acad. Sci., USA 79, 2194-2198.

Wingender, E. (1988). Compilation of transcription regulating proteins. Nucl. Acids Res. 16:1879.

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.

Xing, Y. Y., and A. Worcel (1989). A 3' exonuclease activity degrades the pseudogene 5S RNA transcript and processes the major oocyte 5S RNA transcript in Xenopus oocytes. Genes Dev. 3, 1008-1018.

Yaciuk, P., and E. Moran (1991). Analysis with specific polyclonal antiserum indicates that the E1A-associated 300-kDa product is a stable nuclear phosphoprotein that undergoes cell cycle phase-specific modification. Mol. Cell. Biol. 11, 5389-5397.

Yager, T. D., and P. Von Hippel (1987). Transcription elongation and termination in Escherichia coli. In Escherichia coli and Salmonella typhimurium: cellular and molecular biology (ed. F.C. Neidhardt), Vol. 2. pp 1241-1275, American Society for Microbiology, Washington DC.

Yankulov, K., J. Blau, T. Purton, S. Roberts, and D. L. Bentley (1994). Transcriptional elongation by RNA polymerase II is stimulated by transactivators. Cell 77, 749-759.

Yon, J., and M. Fried (1989). Precise gene fusion by PCR. Nucl. Acids Res. 17, 4895.

Yu, X. M., G. W. Gelembiuk, C. Y. Wang, W. S. Ryu, and J. E. Mertz (1991). Expression from herpesvirus promoters does not relieve the intron requirement for cytoplasmic accumulation of human beta-globin mRNA. Nucl. Acids Res. 19, 7231-7234.

Zawel, L. and D. Reinberg (1992). Advances in RNA polymerase II transcription. Curr. Opin. Cell Biol. 4, 488-495.