Heterogeneously Initiated Transcription from the Pre-B- and B-Cell-Specific *mb-1* Promoter: Analysis of the Requirement for Upstream Factor-Binding Sites and Initiation Site Sequences

ADAM TRAVIS, JAMES HAGMAN, AND RUDOLF GROSSCHEDL*

Departments of Microbiology and Immunology and of Biochemistry and Biophysics, Howard Hughes Medical Institute, University of California, San Francisco, California 94143-0414

Received 24 June 1991/Accepted 20 August 1991

The *mb-1* gene, encoding a membrane immunoglobulin-associated protein, is developmentally regulated and expressed specifically in pre-B and mature B lymphocytes. Analysis of the TATA-less *mb-1* promoter indicated that it directs initiation of transcription from multiple sites. Promoter sequences between –68 and +70 conferred the correct pattern of cell type-specific transcription upon a heterologous gene. Two nuclear factor-binding sites that are important for promoter function were identified between –59 and –38. Both sites interacted with ubiquitous nuclear factors *in vitro*. One of these factors was identified as Sp1. Multimerized copies of both factor-binding sites augmented expression from a heterologous minimal promoter in both lymphoid and nonlymphoid cells, suggesting that additional *mb-1* promoter sequences are involved in determining the correct cell type specificity. Analysis of the heterogeneity of transcription initiation indicated that a mutation which increased the distance between upstream sequences and the region of initiation resulted in the utilization of a novel set of initiation sites. Moreover, an insertion of a TATA element into the *mb-1* promoter at –30 biased initiation of transcription to +1 but did not abolish the use of the other sites. Mutation of an initiator sequence homology encompassing one of the major initiation sites had only a minor effect on its utilization. From these data, we conclude that upstream factor-binding sites in the TATA-less *mb-1* promoter define a region in which initiation of transcription occurs at multiple sites.

B-lymphocyte differentiation turns multipotential progenitor cells into highly specialized cells which ultimately secrete antibody (reviewed in references 1 and 13). The B-cell lineage can be divided into three major stages of differentiation which have been largely defined by changes in the pattern of immunoglobulin (Ig) gene expression. Pre-B cells represent the earliest stage in which the Ig μ heavy-chain gene is expressed. Subsequently, the Ig κ or λ light-chain gene is expressed in B cells, which display antibody on the cell surface. Together, pre-B and B cells represent the early, antigen-independent stages of B-cell differentiation. After B cells encounter antigen, they terminally differentiate into antibody-secreting plasma cells.

Antibody production requires the expression of other lymphoid-specific genes that can be classified according to their stage-specific expression pattern. One set of genes is expressed in pre-B cells but not in later-stage B cells. Included in this group are the λ_5 and V_{preB} genes (28, 42), which likely encode the Ig ω and ι surrogate light-chain components (9), as well as the terminal deoxynucleotidyl transferase (TdT) gene (31) and the recombinase-activating genes (*RAG1* and *RAG2*) (39, 46) which are involved in the assembly of Ig gene segments. The other set of genes is expressed specifically in pre-B and surface IgM-positive B cells, but not in late-stage antibody-secreting plasma cells, and includes the *mb-1* and CD19 genes (43, 52). The *mb-1* gene encodes a protein which is noncovalently associated with membrane-bound IgM (6, 24) and may function as part of a complex involved in signal transduction (5). Likewise, CD19 is a membrane protein thought to participate in signal transduction (11).

Most studies of gene regulation during B-cell differentiation have focused on the lineage-specific expression of the Ig heavy- and light-chain genes. The cell type-specific function of the promoters of Ig genes was shown to be governed by the OCTA site, a nucleotide sequence which can interact with the lymphoid-specific Oct-2 protein and the ubiquitous Oct-1 protein (36, 47, 48, 51, 55). In addition to the OCTA site, the promoters of the Ig genes contain a TATA element positioned approximately 30 bp upstream of the transcription start site. TATA elements are recognized by the general transcription factor TFIID (14, 38, 45), and they have been shown to specify the site of transcription initiation (2, 20, 21) and to be critical for promoter activity *in vitro* (12, 33, 54).

To address the molecular mechanisms that underlie the common lymphoid lineage specificity and distinct temporal expression patterns of the genes involved in antibody production, we began to study the cell type-specific regulation of the *mb-1* gene. The promoter of this gene differs from those of Ig genes in two aspects. First, the *mb-1* promoter lacks an OCTA site, raising the question as to whether other cell type-specific factors are involved in its regulation. Second, the *mb-1* promoter is a representative of a group of promoters of lymphoid-specific genes that lack a discernible TATA element (25). Other members of this group include the B-cell-specific genes λ_5 (29), V_{preB} (28), and B29 (23) and the early B- and T-cell-specific TdT gene (49). Similar to other TATA-less promoters, transcription from the promoters of the *mb-1*, B29, and V_{preB} genes initiates at multiple sites (this study; 23, 52a). By contrast, transcription from the TATA-less TdT promoter initiates at a single nucleotide within an initiator sequence, which appears to substitute for a TATA element (49, 50). Therefore, the question arises as to whether heterogeneous initiation from the *mb-1* promoter

* Corresponding author.

FIG. 1. Evidence that transcription of the *mb-1* gene is initiated at multiple sites. (A) Primer extension mapping of the 5' ends of endogenous *mb-1* transcripts in RNA from two pre-B-cell lines (40E1 and 38B9) and a fibroblast cell line (NIH 3T3). A DNA sequencing ladder (G, A, T, C) and the sequence of the coding strand at the sites of initiation are shown. Large and small filled circles designate strong and weak initiation sites, respectively. The 5'-most major primer-extended product was arbitrarily designated +1. (B) Detection of the 5' ends of endogenous *mb-1* transcripts (mb5') by S1 nuclease protection analysis. RNA from pre-B-cell lines (38B9 and PD36), B-cell lines (WEHI231 and M12), and a plasmacytoma cell line (J558L) was analyzed. Below, an S1 nuclease protection assay detecting the 5' ends of endogenous H4 transcripts [H4(5')] is shown. (C) Schematic diagram of the 5'-end-labelled DNA probe used for detection of the *mb-1* transcripts by S1 nuclease mapping. The DNA probe was isolated from a gene construct containing a fusion of genomic and cDNA sequences in exon 1 (see Materials and Methods). nt, nucleotides.

differs mechanistically from that mediated by the TdT initiator, which specifies a single start site.

Here, we report that the *mb-1* promoter region between -68 and +70 directs early B-cell-specific transcription from multiple sites. Transcription from the *mb-1* promoter is dependent on two upstream factor binding sites. Both sites interact with ubiquitous nuclear factors in vitro and can augment the activity of a heterologous minimal promoter in lymphoid and nonlymphoid cells. Heterogeneous initiation of transcription from the *mb-1* promoter was found to depend on the sequence of the initiation sites and their approximate location relative to upstream sequences.

MATERIALS AND METHODS

Isolation of the *mb-1* promoter. Genomic clones of the *mb-1* gene were isolated from a BALB/c mouse genomic library in the bacteriophage λ vector EMBL-3 SP6/T7 (Clontech), using a 32 P-labelled *mb-1* cDNA clone as the probe. Upon digestion of the *mb-1* genomic bacteriophage clone M9, a 4.5-kb *Bam*HI-*Xho*I fragment containing the *mb-1* promoter region was identified by using a 32 P-labelled synthetic oligonucleotide comprising the noncoding strand of the *mb-1* cDNA from positions +63 to +43 (coordinates of cDNA as in reference 43): 5'-ACAGGCGTATGACAAGAAGAG. This fragment was subcloned into Bluescript M13+ (Stratagene) to yield plasmid pM9-4. The nucleotide sequence of the *mb-1* promoter was determined by using this oligonucleotide, plasmid pM9-4, Sequenase (U.S. Biochemical), and published methods (44), and it was found to match the sequence previously obtained (25) at 318 of 323

nucleotide positions (Fig. 3A). Sequences 5' to the *Bam*HI site (position -252; Fig. 3A) were isolated from the M9 *mb-1* genomic bacteriophage clone as a 2.5-kb *Bam*HI fragment for use in the construction of -800 β Py and -2800 β Py (see below).

Plasmid construction and mutagenesis. To generate the β Py vector backbone (Fig. 2A), the *Nco*I site at +359 of a hybrid mouse-human β -globin gene (7) was blunt ended with Klenow enzyme, *Sal*I linkers were ligated to the blunt ends, and a 1.8-kb *Sal*I-*Xho*I fragment containing the hybrid β -globin gene sequences from +359 to +2200 (37) was isolated and ligated with the *Sal*I-*Xho*I vector fragment of plasmid p μ (19).

Isolation of the *mb-1* promoter region for expression analysis was accomplished by first introducing a *Hind*III site into the first exon (25) of the *mb-1* gene at position +70 (Fig. 3A) by site-directed mutagenesis. Single-stranded DNA was prepared from pM9-4, and mutagenesis was performed as described previously (30), using the oligonucleotide 5'-CCTGCCTCTCCTCCTGCAAGCTTCATACGCCTGTTTG, yielding plasmid pM9-4S1. The 325-bp *Bam*HI-*Hind*III fragment (Fig. 3A) was then isolated from pM9-4S1, blunt ended with Klenow enzyme, and ligated into blunt-ended *Sal*I β Py vector to yield wt(-252) β Py [abbreviated wt(-252)]. The constructs Δ -218, Δ -159, Δ -114, Δ -68, and Δ -22 were similarly generated by subcloning blunt-ended *Hph*I-*Hind*III, *Bal*I-*Hind*III, *Nde*I-*Hind*III, *Hpa*II-*Hind*III, and *Asp* 718-*Hind*III fragments, respectively, into the blunt-ended *Sal*I β Py vector. An identical set of constructs lacking polyomavirus sequences was generated for transfection of M12 B cells (Fig. 3C) by subcloning the same set of frag-

FIG. 2. Lineage- and differentiation stage-specific activity of the *mb-1* gene promoter. (A) Schematic diagram of the plasmid backbone of the *mb-1* promoter constructs used in this and subsequent transfection experiments. The constructs contain *mb-1* promoter fragments linked to the second exon of β-globin genomic sequences. The second and third exons of the β-globin gene are represented by filled bars. The gene constructs also contain the early region of the polyomavirus (PyT) and an H4 reference gene (H4) (19). (B) Detection of transcripts from the *mb-1* promoter constructs by S1 nuclease protection analysis. Constructs containing *mb-1* promoter sequences from -2800 to +70 bp or from -800 to +70 bp (shown in panel A) were transfected into fibroblasts (L), pre-B cells (PD36), B cells (M12), and plasmacytoma cells (J558L and SP2). Accurately initiated transcripts are indicated (*mbβ5'*). Higher-molecular-weight bands represent read-through transcripts from a strong upstream promoter in the vector. Transcription from this upstream promoter, however, does not interfere with the activity of the *mb-1* promoter because similar levels of specific transcripts were detected with a wt(-252) construct containing a transcription termination cassette (34) inserted upstream of the *mb-1* promoter (data not shown). An S1 nuclease assay which detects transcripts from the H4 reference

elements into a blunt-ended *SalI* β vector from which the 3.5-kb *Bam*HI fragment (containing the polyomavirus sequences) was removed. The -800βPy and -2800βPy constructs were generated by inserting a blunt-ended 840-bp *Hind*III fragment of the M9 bacteriophage (-770 to +70) and a 2.9-kb *Bam*HI-*Hind*III partial digestion fragment of the M9 bacteriophage (-2800 to +70), respectively, into a blunt-ended *SalI* βPy vector.

Linker scanning mutants of the *mb-1* upstream factor (MUF) binding site region (Fig. 4A), the initiation site region (Fig. 8A), and the TATA box-containing mutant (Fig. 8B) were generated by site-directed mutagenesis of single-stranded DNA prepared from plasmid pBluMβ. pBluMβ was generated by subcloning a 450-bp *Bam*HI fragment of pwt(-252)βPy into Bluescript M13+ cut with *Bam*HI. Mutagenesis was performed as described previously (30), using the following oligonucleotide primers: 5'-CTCACTTCCTGTTCATCTGCAGCGGAGATGGGCTCC for LS(-66/-60), 5'-CAAACCTCCGCTCACTGCTGCAGCAGCCGTGCCGAG for LS(-57/-51), 5'-GATTCCACTCCAACTCACTGCAGCTTCCTGTTCAGCC for LS(-48/-42), 5'-GGTACCAGATTCACGCTGCAGTCCGCTCACTTCC for LS(-39/-33), 5'-GTAGCCTTGAGGTACACTGCAGCACTCCAACTCCG for LS(-29/-23), 5'-GACCCCTGGCATCCTCAGCAGTGAGTCGGTTAG for LS(+20/+25), 5'-GACCCCTGGCATCCTCGAGCAGTACTTCGGTTAGTTTGGG for LS(+13/+25), 5'-GGCATCGTCTCCAGTACTTCGGTTAGTTTGGG for LS(+13/+15), and 5'-GCCTTGAGGTACAGTTTATATACTCCAACTCCGCG for TATA(-30). The deletion mutants Δ(-57/-42) and Δ(-57/-33) were generated by triple ligations of the 200-bp *Bam*HI-*Pst*I fragment of LS(-57/-51) with a *Bam*HI-*Hind*III Bluescript vector and either a 120-bp *Pst*I-*Hind*III fragment from LS(-48/-42) or a 110-bp *Pst*I-*Hind*III fragment from LS(-39/-33). The 12-bp insertion mutants LI(-18) and LI(+25) were generated by inserting an 8-bp *SalI* linker into the filled-in *Asp* 718 site of LS(-66/-60) or the filled-in *Xho*I site of LS(+20/+25) in pBluMβ. The deletion mutant Δ(-18/+20) was generated by triple ligation of a Bluescript M13+ *Eco*RI-*Bam*HI vector with a 240-bp *Eco*RI-*Asp* 718 fragment (with the *Asp* 718 site blunt ended) and a 170-bp *Xho*I-*Bam*HI fragment (with the *Xho*I site blunt ended), both isolated from the LS(+20/+25) mutant in pBluMβ. All 14 mutant promoters were isolated from the Bluescript subclones as approximately 300-bp *Bam*HI-*Hind*III fragments, then blunt-ended with Klenow, and ligated into the blunt-ended *SalI* βPy vector to produce the final constructs for transfection.

Plasmids MUFwt, MUFmut1, MUFmut2, and MUFmut3 were derivatives of the *fos*-chloramphenicol acetyltransferase (CAT) reporter plasmid p301-56 (17). Three copies of double-stranded oligonucleotides containing wild-type (wt) or mutant (mut) *mb-1* promoter sequences (-59 to -38) (Fig. 7A) were introduced in the same orientation into the *SalI* site at -56 of the *fos*-CAT plasmid.

Cell culture and DNA transfection. NIH 3T3 and L cells were grown in Dulbecco modified Eagle medium containing 10% fetal bovine serum (Hyclone). Lymphoid cell lines were grown in RPMI containing 10% fetal bovine serum and 50 μM 2-mercaptoethanol.

gene (H4^{ref}) was used to control for transfection efficiency (lower panel) in this and subsequent transfection experiments. (C) Schematic diagram of the DNA probe used to detect the 5' ends of transcripts from the transfected *mb-1* promoter/β-globin gene constructs. nt, nucleotides.

FIG. 3. Serial 5' end deletions define two functional regions of the *mb-1* promoter. (A) Nucleotide sequence of the coding strand of the *mb-1* gene promoter from -253 to +70 bp. The strong and weak transcription initiation sites are designated with thick and thin arrows, respectively. The +1 position is arbitrarily assigned to the 5'-most start site. Positions of the 5' ends of the various deletion mutants are indicated. (B) Determination of the relative number of specific transcripts derived from *mb-1* promoter 5'-deletion mutants in pre-B cells by S1 nuclease protection analysis. PD36 cells were transfected with gene constructs containing *mb-1* promoter fragments from the variable 5' boundary indicated (-2800, -800, -252, -218, -159, -114, -68, or -22 bp) to +70 bp. An S1 nuclease assay with RNA from untransfected PD36 cells was performed to control for background bands (lane 9). Specific transcripts from the *mb-1* promoter are designated mbβ5'. The strong bands of lower mobility represent vector read-through transcripts. Below, specific transcripts from the H4 reference gene (H4^{ref}) are shown. (C) Determination of the relative number of specific transcripts derived from *mb-1* promoter 5'-deletion mutants in B cells by S1 nuclease protection analysis. Because of their higher transfection efficiency, M12 B cells were transfected with gene constructs analogous to those in panel B but lacking the early region of polyomavirus (see Fig. 2A for structures of plasmids).

Transfections were done by the DEAE-dextran method (19), with 2 μg of DNA per ml. For cotransfections, this consisted of 1.5 μg of CAT reporter per ml plus 0.5 μg of pRSV-luciferase (15) per ml. Cells were harvested after 48 to 50 h, and total RNA (10) or extracts for luciferase and CAT assays (18) were prepared as described previously.

DNA probes. The template for generating the probe for endogenous *mb-1* transcripts (pMb1gc) was prepared by introducing a 255-bp *StuI-XbaI* *mb-1* promoter fragment (from -217 to +38) into an *EcoRV-XbaI*-cut Bluescript subclone containing *mb-1* cDNA sequences from +38 to the poly(A) tract. To make the radiolabelled probe, a 393-bp *NdeI-PvuII* fragment (Fig. 1C) was isolated from pMb1gc

and 5' end labelled with ³²P, and the correct strand was isolated from a strand-separating gel as described previously (44). Single-stranded DNA probes for S1 nuclease mapping of the 5' ends of transcripts from the *mb-1* promoter constructs (Fig. 2C) were generated by extending a ³²P-5'-end-labelled primer comprising the mouse β-globin noncoding strand, second exon sequence beginning at the *Bam*HI site (27), 5'-GATCCACATGCAGCTTGTACAG, using a single-stranded DNA template generated from the *mb-1* promoter/β-globin gene fusion subclone pBluMβ. *Eco*RI digestion of the extension product at a polylinker site was used to generate the 3' end of the probe. For the *mb-1* initiation site mutants, DNA probes were similarly prepared by using the same ³²P-labelled primer and templates containing the various mutations (Fig. 8A and B) in pBluMβ. The DNA probe used to detect endogenous H4 [H4(5')] and reference H4 transcripts was prepared as described previously (19).

Primer extension, S1 nuclease protection, luciferase, and CAT assays. Primer extension of endogenous mRNA was performed by using 1 μg of poly(A)⁺ RNA, a ³²P-end-labelled complementary oligonucleotide (+63 to +43 of the *mb-1* sequence; see above), and murine leukemia virus reverse transcriptase (Bethesda Research Laboratories), using previously described methods (35).

The S1 nuclease protection assay was performed by using 20 μg of total RNA as described previously (19). Hybridization was at 48°C for 12 to 15 h. The hybrids were digested with 20 U of S1 nuclease (Pharmacia) at 37°C for 1 h. The protected DNA fragments were visualized by using 6% polyacrylamide-urea gels and autoradiography.

Luciferase and CAT assays were performed and quantitated as described previously (53) using a PhosphorImager (Molecular Dynamics).

Nuclear extracts and analysis of protein-DNA complexes. PD36 nuclear extract was prepared as described previously (16). Electrophoretic mobility shift and methylation interference assays were performed as described previously (22) except that MgCl₂ was not present in the binding reaction. The synthetic oligonucleotide probes and competitors used are described in the figure legends. Mobility supershift assays were performed by incubating PD36 nuclear extract with 1 μl of either normal rabbit serum or anti-Sp1 serum (a gift from Steve Jackson and Robert Tjian) for 30 min at room temperature before adding either the SV40(87/111) or the mb(-66/-23) oligonucleotide probe.

RESULTS

Early B-cell-specific transcription of the *mb-1* gene is initiated at multiple sites. The 5' ends of endogenous *mb-1* transcripts were mapped by using primer extension (Fig. 1A) and S1 nuclease protection (Fig. 1B) assays. Both methods detected multiple 5' ends of *mb-1* transcripts in poly(A)⁺ RNA from the pre-B-cell lines 40E1, 38B9, and PD36 and the B-cell lines WEHI231 and M12 (Fig. 1A and B). Specific transcripts were not detected in NIH 3T3 fibroblast and J558L plasmacytoma cell lines, representing nonlymphoid and late-stage B cells, respectively. These data confirmed the early B-cell-specific expression pattern of the *mb-1* gene (43) and further indicated that transcription initiation is heterogeneous. The majority of *mb-1* transcripts initiated within a 16-bp region. The precise positions of the 5' ends were determined by electrophoresis of the primer extension products next to a sequencing ladder (Fig. 1A). For reference, we arbitrarily assigned +1 to the position of the 5'-most major initiation site.

S1 nuclease mapping of the *mb-1* RNA 5' ends with a probe diagrammed in Fig. 1C revealed a similar set of major initiation sites and, in addition, minor initiation sites within a 32-bp upstream region (Fig. 1B, lanes 1 to 4). Together, these data indicate that transcription initiation from the *mb-1* promoter is heterogeneous but restricted primarily to a 16-bp region.

Lineage and differentiation stage specificity of the *mb-1* promoter. We examined the early B-cell specificity of *mb-1* gene transcription by transfecting various mouse cell lines with gene constructs containing *mb-1* promoter sequences linked to a β -globin reporter gene. The *mb-1* promoter fragments included sequences from either -2800 to $+70$ (-2800β Py) or -800 to -70 (-800β Py) (Fig. 2A). These gene constructs also contained polyomavirus sequences for replication in mammalian cells and an H4 reference gene to estimate transfection efficiency (19). Following transfection, total RNA was isolated and analyzed for the presence of specific *mb-1* transcripts by using an S1 nuclease protection assay (Fig. 2B). The DNA probe contains β -globin sequences and detects specific transcripts from the transfected *mb-1* promoter constructs but not from the endogenous *mb-1* gene (Fig. 2C). Accurately initiated transcripts from both the -2800β Py and the -800β Py *mb-1* promoter constructs were detected in RNA from transfected PD36 pre-B cells and M12 B cells (Fig. 2B, lanes 3 to 6). By contrast, specific transcripts were detected at a much lower abundance in RNA from the late-stage B cells J558L and SP2 (lanes 7 to 10), L fibroblasts (lanes 1 and 2), NIH 3T3 fibroblasts, and T-cell lines (data not shown). Transcription from the reference H4 transfection control gene was detected at a similar level in all RNA samples with the exception of PD36 cells, which were transfected less efficiently. These data indicate that *mb-1* promoter sequences from -800 to $+70$ contain regulatory information sufficient for pre-B and B-lymphocyte-specific transcription. The activity of the promoter, however, is approximately 10-fold higher in pre-B cells (PD36) than in B cells (M12).

Mapping of *mb-1* promoter elements. To delineate the sequences involved in *mb-1* promoter function, progressive 5' deletions (Fig. 3A) were tested for their effects on promoter activity in PD36 pre-B cells (Fig. 3B). An *mb-1* promoter fragment comprising sequences from -252 to $+70$ showed the same activity as the larger promoter fragments (Fig. 3B, lanes 1 to 3) and was arbitrarily designated as wild type [wt(-252)]. Deletion of sequences 5' to -159 had no significant effect on the generation of specific transcripts (lanes 3 to 5). Truncation of the *mb-1* promoter to -114 decreased its activity threefold. No further decrease was observed upon removal of additional sequences to -68 (lanes 5 to 7). However, truncation of the promoter to -22 decreased the level of expression 20-fold relative to that of the wt(-252) promoter (compare lanes 3 and 8). The strong bands of lower mobility represent vector read-through transcripts, because their migration is determined by the site of the 5'-deletion breakpoint. This 5'-deletion series localized functionally important sequences to the regions between -159 and -115 and between -68 and -23 . Transfection of the 5'-deletion series into M12 B cells indicated that sequences upstream of -68 did not contribute to promoter function (Fig. 3C, lanes 1 to 6), suggesting a pre-B-cell-specific function of sequences between -159 and -115 . When these truncated *mb-1* promoter/ β -globin gene constructs were transfected into NIH 3T3 fibroblasts or J558L plasmacytoma cells, no activity could be detected with any of the promoter fragments (data not shown). Together, these

FIG. 4. Linker scanning mutagenesis of the *mb-1* proximal promoter region. (A) Nucleotide sequence of the coding strand of the *mb-1* promoter from -68 to -22 and of the linker scanning and deletion mutants. Dashes represent bases identical to the wild-type sequence. Mutant nucleotides, shown as lowercase letters and deletions, shown as gaps, were introduced into the wt(-252) *mb-1* promoter (containing sequences from -252 to $+70$). For the linker scanning mutants, coordinates in parentheses designate the region of nucleotide changes. For the internal deletion mutants, the coordinates designate the boundaries of the region of nucleotide changes and deletion. (B) Determination of the relative number of transcripts derived from the mutant *mb-1* promoters (shown in panel A) by S1 nuclease protection assay. The gene constructs were transfected into PD36 pre-B cells, and the transcripts were analyzed as described for Fig. 3. The strong bands of lower mobility represent cleavage by S1 nuclease at mismatches between vector read-through transcripts and the wild-type DNA probe. Below, the specific H4 reference ($H4^{ref}$) transcripts are shown.

data suggest that the *mb-1* promoter region residing between -68 and $+70$ is sufficient for early B-cell-specific activity.

To analyze the proximal *mb-1* promoter region (-66 to -23) in more detail, clustered point mutations (linker scanning mutations) were introduced in the wt(-252) promoter (Fig. 4A). The constructs were transfected into PD36 pre-B cells, and specific *mb-1* transcripts were detected by S1 nuclease protection analysis (Fig. 4B). The LS($-57/-51$), LS($-48/-42$), and LS($-39/-33$) mutations each decreased the level of transcription fourfold relative to the wild-type promoter (Fig. 4B; compare lanes 3 to 5 with lane 1). Because these clustered mutations span a 25-bp region, the question arose as to whether they affect one or multiple factor binding sites. Therefore, two small internal deletions, $\Delta(-57/-42)$ and $\Delta(-57/-33)$, were introduced into the wt(-252) promoter (Fig. 4A). Both deletions decreased *mb-1* promoter activity 10- to 20-fold to a level comparable to that of the $\Delta(-22)$ deletion (Fig. 4B, lanes 8 to 11). Together, these data suggest the presence of multiple factor binding sites in the *mb-1* proximal promoter region.

FIG. 5. Sequence-specific binding of nuclear factors to the functionally important proximal region of the *mb-1* promoter. (A) Detection of MUF1 and MUF2 by electrophoretic mobility shift assays. A double-stranded oligonucleotide probe containing *mb-1* promoter sequences from -66 to -23 was used in binding reactions with nuclear extract from PD36 pre-B cells. Two major complexes are observed with the addition of extract (lane 2). Both complexes are competed for with excess unlabelled probe oligonucleotide (lanes 3 to 5). To further localize sequences recognized by MUF1 and MUF2, binding competition was performed with three staggered 25-bp oligonucleotides spanning the proximal *mb-1* promoter from -78 to -54 (lanes 6 to 8), -68 to -44 (lanes 9 to 11), and -58 to -34 (lanes 12 to 14). (B) Specificity of MUF1 binding. A double-stranded oligonucleotide probe containing *mb-1* promoter sequences from -58 to -34 was incubated with PD36 pre-B-cell nuclear extract. The sequence specificity of MUF1 binding (lane 2) was determined by including excess wild-type (lanes 3 and 4) or mutant (lanes 5 and 6) oligonucleotide competitors. Both competitors contained *mb-1* promoter sequences from -59 to -38. The nucleotide changes in the mutant oligonucleotide are shown in Fig. 7A. The SV40(87/111) competitor contained a single Sp1 binding site from the SV40 early region (nucleotide positions 87 to 111) (4). (C) Evidence that MUF1 is identical to the transcription factor Sp1. Oligonucleotide probe containing a single Sp1 binding site from SV40 (lanes 1 and 2) or the MUF1 and MUF2 binding sites from the *mb-1* promoter (lanes 3 and 4) were added to PD36 pre-B-cell nuclear extract that had been preincubated with either normal rabbit serum (NRS) or rabbit anti-Sp1 serum (anti-Sp1). Protein-DNA complexes were separated in an electrophoretic mobility shift assay. F, free probe.

Factor binding to the proximal *mb-1* promoter region. We examined the binding of nuclear factors to the functionally important proximal *mb-1* promoter sequences by electrophoretic mobility shift assays with an end-labelled oligonucleotide probe (from -66 to -23) and a pre-B-cell nuclear extract. Two major protein-DNA complexes, termed *mb-1* upstream factors 1 and 2 (MUF1 and MUF2), were detected (Fig. 5A, lane 2). The binding of both nuclear factors could be specifically competed for with excess unlabelled probe oligonucleotide (lanes 3 to 5) but not with an oligonucleotide comprising *mb-1* promoter sequences between -78 and -54 (lanes 6 to 8). The binding sites of both factors were further localized by competition with two overlapping 25-bp oligonucleotides spanning the proximal promoter region (lanes 9 to 14). These data indicate that MUF1 interacts with the -58 to -34 region and that MUF2 interacts with the -68 to -44 region.

To further characterize the MUF1-DNA complex, we used the shorter oligonucleotide probe (-58 to -34), which binds MUF1 only. The specificity of MUF1 binding to DNA was confirmed by competition with excess unlabelled wild-type or mutant *mb-1* oligonucleotide (Fig. 5B, lanes 2 to 6). The mutant oligonucleotide has the wild-type sequence GAGGCGGAG changed to GAGTATGAG. The wild-type sequence conforms to the reported consensus binding site for the transcription factor Sp1 (3). Therefore, we examined

the relationship between MUF1 and Sp1 by using a known Sp1 binding site from the simian virus 40 (SV40) early region as a competitor for MUF1 binding (lanes 7 and 8). The Sp1 binding site from SV40 was a more efficient competitor for MUF1 binding than was the *mb-1* promoter sequence. To confirm that the MUF1 complex contains Sp1, nuclear extract was incubated with either normal rabbit serum or a polyclonal anti-Sp1 serum prior to the addition of the *mb-1* (-66/-23) oligonucleotide probe (Fig. 5C, lanes 3 and 4). The anti-Sp1 serum retarded the mobility of the MUF1-DNA complex but had no effect on the MUF2-DNA complex (lanes 3 and 4). Moreover, an Sp1-DNA complex formed on the SV40 oligonucleotide comigrated with the MUF1-DNA complex (lanes 1 and 3). As expected, anti-Sp1 serum retarded the mobility of this complex as well (lane 2). Together, these data strongly suggest that MUF1 is identical to Sp1.

Nucleotides in the proximal *mb-1* promoter which are contacted by MUF2 and Sp1 were determined by methylation interference analysis (Fig. 6A). Protein-DNA complexes were formed on partially methylated oligonucleotide probes comprising the binding site of either MUF2 or Sp1. On the coding strand, two adjacent G residues at -52 and -53 interfere with MUF2 binding when methylated, and all G residues within a region between -39 and -47 interfere with Sp1 binding when methylated (Fig. 6). On the noncoding

FIG. 6. Methylation interference analysis of the *mb-1* proximal promoter region binding MUF2 and Sp1. (A) Methylation interference analysis of MUF2 and Sp1 complexes to detect specific nucleotide contacts in the *mb-1* promoter. Partially methylated and end-labelled *mb-1* DNA probes from -68 to -39 and from -58 to -34 were incubated with PD36 nuclear extract to generate MUF2 complexes (left) or Sp1 complexes (right). Bound (B) and free (F) probe were separated in an electrophoretic mobility shift assay and processed as described previously (22). Nucleotide sequences of the upper (labelled) strands are shown for reference. The regions in which methylated G residues interfered with binding are indicated with brackets. (B) Summary of methylation interference data for both the upper (from panel A) and lower (data not shown) strands of the MUF2 and Sp1 binding sites. Arrowheads indicate nucleotides which interfere with DNA binding when methylated.

strand, only a single G residue in the Sp1 site shows methylation interference, whereas no methylation interference could be detected in the MUF2 site (data not shown). These results indicate that MUF2 and Sp1 recognize adjacent and possibly nonoverlapping binding sites. Because no abundant protein-DNA complexes of lower mobility than MUF2 and Sp1 were observed in the gel mobility shift assay (Fig. 5A), their binding may be mutually exclusive.

Both MUF2 and Sp1 were detected by gel mobility shift assays in nuclear extracts from lymphoid and nonlymphoid cell lines, indicating a ubiquitous distribution (data not shown). Likewise, DNase I protection analysis of the proximal *mb-1* promoter identified a protected region from -59 to -38 with both lymphoid and nonlymphoid extracts (data not shown).

Binding sites for MUF2 and Sp1 can augment expression

FIG. 7. Evidence that proximal promoter region factor binding sites augment expression from a minimal *c-fos* promoter. (A) Structure of the *mb-1/c-fos* promoter-CAT constructs. Three tandem copies of oligonucleotides containing wild-type or mutant *mb-1* promoter sequences from -59 to -38 were inserted into the *fos*-CAT reporter plasmid upstream of c-fos promoter sequences from -56 to +109. Dashes in the *mb-1* promoter sequences of MUFmut1, -2, and -3 represent bases identical to the MUFwt sequence, and lowercase letters indicate mutations. (B) Analysis of *mb-1/c-fos* promoter activity by CAT assays. The five constructs diagrammed in panel A were transfected into various cell lines along with the pRSV-luciferase construct (15) as an internal control. The amount of cell extract used in CAT assays was normalized to luciferase activity. Representative CAT assays, separating the acetylchloramphenicol (Ac-Cm) reaction products and chloramphenicol (Cm) by thin-layer chromatography, are shown. The average relative CAT activities were determined by multiple transfections and normalized to the CAT levels obtained with the *fos*-CAT construct.

from a heterologous promoter. To examine whether the proximal *mb-1* promoter region between -59 and -38 can increase the activity of a heterologous promoter in a cell type-specific manner, three copies of oligonucleotides comprising either wild-type or mutant MUF2 and Sp1 sites were linked to the minimal *c-fos* promoter of a CAT reporter gene (Fig. 7A, *fos*-CAT) (17). The chimeric promoter constructs were transfected into various B- and non-B-cell lines, together with the control plasmid pRSV-luciferase (15), and normalized CAT activity was determined after 48 h. The wild-type *mb-1* oligonucleotide containing both MUF2 and

Sp1 binding sites (MUFwt) increased the expression of the CAT gene in all transfected lymphoid and nonlymphoid cell lines 27- to 110-fold relative to that of the *fos*-CAT gene (Fig. 7B).

The individual contributions of the MUF2 and Sp1 sites to heterologous promoter activity were examined by introducing mutations in either or both factor binding sites (Fig. 7A). These mutations completely interfered with factor binding in vitro (Fig. 5B; data not shown for MUF2). Mutation of the Sp1 site (MUFmut1) decreased expression only twofold relative to the wild-type sequence in fibroblasts, pre-B, B, and J558L plasmacytoma cells; however, the same mutation increased expression twofold in SP2 plasmacytoma cells and in BW5147 T cells (Fig. 7B). By contrast, mutation of the MUF2 binding site (MUFmut2) consistently decreased CAT levels 4- to 10-fold in all cell types analyzed. The *mb-1* oligonucleotide containing mutations in both factor binding sites (MUFmut3) was unable to stimulate expression from the *c-fos* promoter.

A single copy of the various *mb-1* oligonucleotides had a similar effect on *fos* promoter activity, although the overall CAT levels were proportionally lower than those observed with three copies (data not shown). Taken together, the binding sites for MUF2 and/or Sp1 increased the expression from a heterologous promoter in both B- and non-B-cell lines. This ubiquitous activity of the proximal *mb-1* promoter region contrasts with its contribution to the early B-cell-specific function of the intact *mb-1* promoter.

The region of heterogeneous transcription initiation is determined by upstream sequences. In addition to its early B-cell-specific activity, the *mb-1* promoter differs from the Ig gene promoters in that it directs transcription initiation from multiple sites. In this respect, the *mb-1* promoter also differs from the TATA-less TdT promoter, which initiates transcription at a single nucleotide within the initiator element. The initiator was shown to be necessary and sufficient for specifying initiation from a single site (49, 50). Alignment of the 3'-most initiation site at +16 of the *mb-1* promoter with the initiator of the TdT promoter revealed a 5-of-7 match (CTCANTGG) and a 6-bp identity in the adjacent transcribed region (GGAGAC; Fig. 8A).

We examined the importance of these sequences for specifying transcription initiation by introducing clustered mutations in the wt(-252) *mb-1* promoter construct (Fig. 8A). For S1 nuclease analysis of RNA from transfected PD36 pre-B cells, each RNA sample was hybridized with a complementary probe containing the same base changes as the transfected gene construct (see Materials and Methods). The mutation LS(+13/+15) of the initiator homology and the LS(+13/+25) double mutation of both the initiator and transcribed-region homologies decreased the efficiency of utilization of the 3'-most pair of major initiation sites only by a factor of 2 and had no effect on the utilization of the initiation sites further upstream (Fig. 8A; Fig. 8C, lanes 3 and 4). The LS(+20/+25) mutation of the transcribed-region homology alone had no detectable effect on the utilization of multiple initiation sites (Fig. 8C, lane 2). Together, these data suggest that neither of these homologies in the *mb-1* promoter is required for the generation of the heterogeneous pattern of transcription initiation.

To examine the role of upstream or downstream regulatory sequences in the positioning of the multiple sites of transcription initiation, we inserted a 12-bp linker either at -18 or at +25 (Fig. 8B). The upstream insertion mutant LI(-18) utilized an additional set of initiation sites which extended 21 nucleotides 5' of +1. The utilization of the

normal start sites (+1 to +16) was reduced threefold (Fig. 8C, lane 5). Thus, this mutation results in transcription initiation over a broader region compared with the wild type. Moreover, the center of the initiation site region is shifted 5' by approximately 12 nucleotides. In contrast, the downstream insertion mutant LI(+25) had a very small effect on the utilization of the RNA start sites, but, as expected, it extended all transcripts by 12 nucleotides (Fig. 8C, lane 6). Together, these results suggest that upstream sequences determine the region within which multiple sites can be used for transcription initiation.

This interpretation for the functional role of upstream sequences for the positioning of transcription start sites was supported by the analysis of another promoter deletion mutant, $\Delta(-18/+20)$, which lacks the region of normal transcription initiation. A novel set of initiation sites was utilized within a region downstream of the deletion breakpoint. The location of this novel initiation site region, relative to the upstream factor binding sites, corresponds to that of the wild-type promoter (Fig. 8B; Fig. 8C, lane 7). However, the pattern of multiple initiation sites of this mutant promoter is distinct from those of both the wild-type and LI(-18) mutant promoters, indicating that the nucleotide sequences within the region of transcription initiation determine the exact positions of the start sites.

To determine whether the heterogeneous initiation of transcripts could be attributed solely to the absence of a TATA element, we examined the effect of an insertion of the sequence TATAAA in the *mb-1* promoter between -31 and -26 [Fig. 8B, TATA(-30)]. The number of transcripts initiating from the predicted site downstream of the TATA element (+1), was increased threefold relative to that of the wild-type promoter. However, initiation from the other sites was still observed, although at a reduced level (Fig. 8C, lane 8). Thus, the multiple initiation sites in the *mb-1* promoter are utilized even in the presence of an introduced TATA element, suggesting an intrinsic TATA-independent mechanism of transcription initiation.

DISCUSSION

Role of ubiquitous factors in cell type-specific regulation. The function of the *mb-1* promoter in early B cells depends on sequences between -59 and -38, which bind the ubiquitous factors MUF2 and Sp1 in vitro. The proximal region of the *mb-1* promoter can also augment the activity of a heterologous minimal promoter. This increase in promoter activity is dependent upon both binding sites for MUF2 and Sp1 and can be detected in all analyzed lymphoid and nonlymphoid cell lines. The ubiquitous activity of these factor binding sites in a heterologous context contrasts with their functional role for the cell type-specific *mb-1* promoter. Although our current experimental data do not yet provide an insight into the mechanisms underlying the cell type-specific function of the *mb-1* promoter, several possible regulatory schemes can be considered.

First, negatively acting elements in the *mb-1* promoter may interfere with the function of the proximal promoter region in late-stage B cells and in non-B cells. Negative regulation has been implicated in the B-cell-specific function of the Ig heavy-chain enhancer (41). In this case, a putative negative regulator acting through the μ E5 site can antagonize the positive function of both the ubiquitous factor ITF-1, which binds at the μ E5 site, and the ubiquitous factor TFE3, which binds at the adjacent and nonoverlapping μ E3 site (41). Because the minimal *mb-1* promoter fragment (-68

FIG. 8. Evidence that positioning of transcription initiation sites in the *mb-1* promoter is specified by upstream and initiation site sequences. (A) Aligned nucleotide sequences of the TdT and *mb-1* promoters and of three *mb-1* promoter mutants disrupting the homologies with the TdT promoter. Mutations were introduced into the wt(-252) construct containing *mb-1* promoter sequences from -252 to +70. The effect of each mutation (based on results shown in panel C, lanes 1 to 4) is summarized, with thick and thin arrows indicating the positions of major and minor initiation sites, respectively. (B) Nucleotide sequences of the initiation site regions of the wild-type and mutant *mb-1* promoters. LI(-18) and LI(+25) contain 12-bp linkers inserted at -18 and +25, respectively. Δ(-18/+20) has a deletion of sequences from -18 to +20. TATA(-30) has the sequence TATAAA substituted for sequences from -31 to -26. Thick and thin arrows indicate the positions of major and minor initiation sites, respectively (see panel C, lanes 5 to 8). (C) S1 nuclease protection analysis of RNA from PD36 pre-B cells transfected with constructs containing the various mutant *mb-1* promoters diagrammed in panels A and B. Specific 5' ends of *mb-1* transcripts were detected with single-stranded DNA probes similar to the one diagrammed in Fig. 2C except that for each sample, a probe complementary to the specific mutant *mb-1* promoter sequence (shown in panel A or B) was used. The eight probes were generated by extending the same 5'-end-labelled oligonucleotide against the various mutant template DNAs (see Materials and Methods), ensuring they had identical specific radioactivity. +1 designates the predicted position of the 5'-most initiation site for all constructs except LI(+25) and Δ(-18/+20). +1 for LI(+25) designates the predicted position of the 5'-most initiation site for LI(+25).

to +70) displays the correct cell type-specific activity, a putative negatively acting element would most likely be present in the region downstream of -38. However, a preliminary analysis of deletion mutants spanning the region between -22 and +70 did not reveal any detectable promoter activity in non-B or late-stage B cells (data not shown).

A second possibility is that the ubiquitous activity of the proximal *mb-1* promoter results from its artificial juxtaposition to the TATA-containing *fos* promoter. We consider it unlikely that the TATA element itself accounts for this loss of specificity because the insertion of a TATA element into the *mb-1* promoter does not abrogate its cell type specificity (data not shown). Nevertheless, juxtaposition of the proximal *mb-1* promoter to *fos* promoter sequences may generate a context in which either or both of the ubiquitous factors MUF2 and Sp1 can activate transcription in all cell types.

In a third regulatory scheme, MUF2 and/or Sp1 could participate in the cell type-specific *mb-1* promoter function by synergizing with an early B-cell-specific factor which binds to flanking promoter sequences. Consistent with such a model, we identified early B-cell factor (EBF), a nuclear factor which has the same expression pattern as the *mb-1* gene and binds to the distal *mb-1* promoter region between -181 and -160 (22). Because clustered point mutations within the EBF site decreased *mb-1* promoter activity in pre-B cells only threefold (22), EBF is unlikely to synergize with factors bound at the proximal promoter region.

Instead, we favor the view that the binding of MUF2 and Sp1 to *mb-1* promoter sequences in vitro and the ubiquitous activity of these factor binding sites in vivo do not necessarily reflect the participation of MUF2 and Sp1 in the regulation of the intact promoter. Possibly, the *mb-1* proximal promoter sequences are normally recognized by a less abundant cell type-specific factor, the binding of which is

avored in the context of the intact promoter by an interaction with another protein. The MUF2 binding site sequence shares seven contiguous base pairs (CAGGAAG) with the π box of the Ig heavy-chain enhancer, which binds a lymphoid-specific factor of unknown identity (32). Indeed, preliminary experiments indicate that the MUF2 binding site in the *mb-1* promoter can interact with the B-cell- and myeloid-specific factor PU.1 (26) present in crude nuclear extract (21a). This factor, together with EBF and possibly other cell type-specific factors, may participate in the early B-cell-specific regulation of the intact *mb-1* promoter.

Multiple initiation sites at a TATA-less promoter. The TATA-less *mb-1* promoter initiates transcription from multiple sites and thus differs from the TATA-less promoter of the TdT gene, which initiates transcription from a single site (49). Although two homologies (5 of 7 bp) with the TdT initiator are found in the *mb-1* promoter centered around the major start sites at +2 and +16, our data suggest that these initiation sites are functionally distinct from the TdT initiator. First, a mutation of the initiator homology around +16 of the *mb-1* promoter decreased expression from the 3'-most pair of initiation sites only by a factor of 2. Second, utilization of multiple initiation sites in the *mb-1* promoter is dependent on their location relative to upstream regulatory elements, whereas the TdT initiator can function autonomously (49, 50). The importance of upstream sequences in the approximate positioning of *mb-1* initiation sites is demonstrated by the analysis of two promoter mutations. An insertion of a 12-bp linker at -18 reduced initiation from the normal sites between +1 and +16 and activated additional initiation sites between -21 and +1. Moreover, an internal deletion removing the entire region of normal transcription initiation resulted in the use of a new set of initiation sites within a downstream region which was brought close to upstream promoter elements. The exact positioning of the multiple initiation sites in the *mb-1* promoter, however, is not determined by upstream sequences because the patterns of start sites differs among the wild-type and mutant promoters. Instead, the nucleotide sequences in the region of initiation appear to determine the exact transcription start sites. Taken together, these results suggest that sequences 5' of the initiation sites in the *mb-1* promoter specify a "window" in which multiple sites can be used for initiation. A similar model has been proposed for RNA polymerase II initiation in yeast cells (8).

Although the multiple initiation sites of the TATA-less *mb-1* promoter are reminiscent of the heterogeneous transcription observed with promoters that have a mutated TATA element (2, 20, 21), the molecular basis for the heterogeneity may be different. Mutations that generate a TATA element in the *mb-1* promoter at -30 favor initiation at a single site (+1); however, they do not eliminate initiation from downstream sites. Thus, the initiation of transcription at multiple sites appears to be an intrinsic property of the *mb-1* promoter and cannot be overridden simply by introducing a TATA element.

In the absence of a TATA element, what upstream element(s) and factor(s) determine the region of initiation in the *mb-1* promoter? Recent experiments using synthetic TATA-less promoters that contain an initiator and upstream Sp1 binding sites have shown that a TFIID fraction is required for transcriptional activation by Sp1 in vitro (40, 50). Experiments with highly fractionated TFIID suggested that an additional activity, termed a tethering factor, is required for activation of this TATA-less promoter. From these data, a model was proposed in which TFIID and the

transcription complex are anchored through the tethering factor to Sp1 bound at the TATA-less promoter (40). The functionally important Sp1 site in the *mb-1* proximal promoter region could fulfill an analogous role and facilitate the formation of transcription complexes in the absence of a TATA element.

ACKNOWLEDGMENTS

We are grateful to Steve Jackson and Robert Tjian for providing the anti-Sp1 serum and to Michael Z. Gilman for the *fos*-CAT (p301-56) reporter plasmid. We thank Gail Mosely for assistance with manuscript preparation. We also thank Steve Hardy and Steve Yoshinaga for critical comments on the manuscript.

J.H. is a postdoctoral fellow of the Leukemia Society of America. This study was supported by funds from the Howard Hughes Medical Institute.

REFERENCES

1. Blackwell, T. K., and F. W. Alt. 1988. Immunoglobulin genes, p. 1-60. In B. D. Hames and D. M. Glover (ed.), *Molecular immunology*. IRL Press, Washington, D.C.
2. Breathnach, R., and P. Chambon. 1981. Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50:349-383.
3. Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* 234:47-52.
4. Buchman, A. R., L. Burnett, and P. Berg. 1982. The SV40 nucleotide sequence, p. 799-841. In J. Tooze (ed.), *DNA tumor viruses*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. Cambier, J. C. 1991. Capturing antigen receptor components. *Curr. Biol.* 1:25-27.
6. Campbell, K. S., and J. C. Cambier. 1990. B lymphocyte antigen receptors (mIg) are non-covalently associated with a disulfide-linked, inducibly phosphorylated glycoprotein complex. *EMBO J.* 9:441-448.
7. Chao, M. V., P. Mellon, P. Charnay, T. Maniatis, and P. Axel. 1983. The regulated expression of β -globin genes introduced into mouse erythroleukemia cells. *Cell* 32:483-493.
8. Chen, W., and K. Struhl. 1985. Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. *EMBO J.* 4:3273-3280.
9. Cherayil, B. J., and S. Pillai. 1991. The ω/λ_5 surrogate immunoglobulin light chain is expressed on the surface of transitional B lymphocytes in murine bone marrow. *J. Exp. Med.* 173:111-116.
10. Chirgwin, T. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294-5299.
11. Clark, E. A., and P. J. L. Lane. 1991. Regulation of human B-cell activation and adhesion. *Annu. Rev. Immunol.* 9:97-127.
12. Concino, M. F., R. F. Lee, J. P. Merryweather, and R. Weinmann. 1984. The adenovirus major late promoter TATA box and initiation site are both necessary for transcription in vitro. *Nucleic Acids Res.* 12:7423-7433.
13. Cooper, M. D., and P. D. Burrows. 1989. B-cell differentiation, p. 1-22. In T. Honjo, F. W. Alt, and T. H. Rabbits (ed.), *Immunoglobulin genes*. Academic Press, London.
14. Davison, B. L., J.-M. Egly, E. R. Mulvihill, and P. Chambon. 1983. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. *Nature (London)* 301:680-686.
15. DeWet, J. R., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani. 1987. Firefly luciferase gene: structure and expression in mammalian cells. *Mol. Cell. Biol.* 7:725-737.
16. 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. *Nucleic Acids Res.* 11:1475-1489.
17. Gilman, M. Z., R. N. Wilson, and R. A. Weinberg. 1986. Multiple protein-binding sites in the 5'-flanking region regulate *c-fos* expression. *Mol. Cell. Biol.* 6:4305-4316.
 18. 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.
 19. Grosschedl, R., and D. Baltimore. 1985. Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. *Cell* 41:885-897.
 20. Grosschedl, R., and M.L. Birnstiel. 1980. Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vivo. *Proc. Natl. Acad. Sci. USA* 77:1432-1436.
 21. Grosveld, G. C., C. K. Shewmaker, P. Jat, and R. A. Flavell. 1981. Localization of DNA sequences necessary for transcription of the rabbit β -globin gene in vitro. *Cell* 25:215-226.
 - 21a. Hagman, J. Unpublished data.
 22. Hagman, J., A. Travis, and R. Grosschedl. 1991. A novel lineage-specific nuclear factor regulates mb-1 gene transcription at the early stages of B cell differentiation. *EMBO J.*, in press.
 23. Hermanson, G. G., M. Briskin, D. Sigman, and R. Wall. 1989. Immunoglobulin enhancer and promoter motifs 5' of the B29 B-cell-specific gene. *Proc. Natl. Acad. Sci. USA* 86:7341-7345.
 24. Hombach, J., T. Tsubata, L. Leclercq, H. Stappert, and M. Reth. 1990. Molecular components of the B-cell antigen receptor complex of the IgM class. *Nature (London)* 343:760-762.
 25. Kashiwamura, S.-I., T. Koyama, T. Matsuo, M. Steinmetz, M. Kimoto, and N. Sakaguchi. 1990. Structure of the murine mb-1 gene encoding a putative sIgM-associated molecule. *J. Immunol.* 145:337-343.
 26. Klemsz, M. J., S. R. McKercher, A. Celada, C. Van Beveren, and R. A. Maki. 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. *Cell* 61:113-124.
 27. Konkel, D. A., S. M. Tilghman, and P. Leder. 1978. The sequence of the chromosomal mouse β -globin major gene: homologies in capping, splicing and poly A sites. *Cell* 15:1125-1132.
 28. Kudo, A., and F. Melchers. 1987. A second gene, V_{preB} in the $\lambda 5$ locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. *EMBO J.* 6:2267-2272.
 29. Kudo, A., N. Sakaguchi, and F. Melchers. 1987. Organization of the murine Ig-related $\lambda 5$ gene transcribed selectively in pre-B lymphocytes. *EMBO J.* 6:103-107.
 30. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 82:488-492.
 31. Landau, N. R., T. P. St. John, I. L. Weissman, S. C. Wolf, A. E. Silverstone, and D. Baltimore. 1984. Cloning of terminal transferase cDNA by antibody screening. *Proc. Natl. Acad. Sci. USA* 81:5836-5840.
 32. Libermann, T. A., and D. Baltimore. 1991. Transcriptional regulation of immunoglobulin gene expression. *Mol. Aspects Cell. Regul.* 6:385-407.
 33. Mathis, D. J., and P. Chambon. 1981. The SV40 early region TATA box is required for accurate in vitro initiation of transcription. *Nature (London)* 290:310-315.
 34. Maxwell, I. H., G. S. Harrison, W. M. Wood, and F. Maxwell. 1989. A DNA cassette containing a trimerized SV40 polyadenylation signal which efficiently blocks spurious plasmid-initiated transcription. *BioTechniques* 7:276-280.
 35. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217:316-324.
 36. Muller, M. M., S. Ruppsert, W. Schaffner, and P. Matthias. 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature (London)* 336:544-551.
 37. Myers, R. M., K. Tilly, and T. Maniatis. 1986. Fine structure genetic analysis of a β -globin promoter. *Science* 232:613-618.
 38. Nakajima, N., M. Horikoshi, and R. G. Roeder. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol. Cell. Biol.* 8:4028-4040.
 39. Oettinger, M. A., D. G. Schatz, C. Gorka, and D. Baltimore. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248:1517-1523.
 40. Pugh, B. F., and R. Tjian. 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* 61:1187-1197.
 41. Ruezinsky, D., H. Beckmann, and T. Kadesch. 1991. Modulation of the IgH enhancer's cell type specificity through a genetic switch. *Genes Dev.* 5:29-37.
 42. Sakaguchi, N., C. N. Berger, and F. Melchers. 1986. Isolation of a cDNA copy of an RNA species expressed in murine pre-B cells. *EMBO J.* 5:2139-2147.
 43. Sakaguchi, N., S.-I. Kashiwamura, M. Kimoto, P. Thalmann, and F. Melchers. 1988. B lymphocyte lineage-restricted expression of *mb-1*, a gene with CD3-like structural properties. *EMBO J.* 7:3457-3464.
 44. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 45. Sawadogo, M., and R. G. Roeder. 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* 43:165-175.
 46. Schatz, D. G., M. A. Oettinger, and D. Baltimore. 1989. The V(D)J recombination activating gene, RAG-1. *Cell* 59:1035-1048.
 47. Scheidereit, C., J. A. Cromlish, T. Gerster, K. Kawakami, C. Balmaceda, R. A. Currie, and R. G. Roeder. 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeo-box protein. *Nature (London)* 336:551-557.
 48. Singh, H., J. H. LeBowitz, A. S. Baldwin, and P. A. Sharp. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature (London)* 319:154-158.
 49. Smale, S. T., and D. Baltimore. 1989. The "initiator" as a transcription control element. *Cell* 57:103-113.
 50. Smale, S. T., M. C. Schmidt, A. J. Berk, and D. Baltimore. 1990. Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc. Natl. Acad. Sci. USA* 87:4509-4513.
 51. Staudt, L. M., R. G. Clerc, H. Singh, J. H. LeBowitz, P. A. Sharp, and D. Baltimore. 1988. Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory octamer DNA motif. *Science* 241:577-580.
 52. Tedder, T. F., and C. M. Isaacs. 1989. Isolation of cDNAs encoding the CD19 antigen of human and mouse B lymphocytes. A new member of the immunoglobulin superfamily. *J. Immunol.* 143:712-717.
 - 52a. Travis, A. Unpublished data.
 53. Travis, A., A. Amsterdam, C. Belanger, and R. Grosschedl. 1991. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor α enhancer function. *Genes Dev.* 5:880-894.
 54. Waslylyk, B., R. Derbyshire, A. Guy, D. Molko, A. Roget, R. Teoule, and P. Chambon. 1980. Specific *in vitro* transcription of conalbumin gene is drastically decreased by single-point mutation in TATA box homology sequence. *Proc. Natl. Acad. Sci. USA* 77:7024-7028.
 55. Wirth, T., L. Staudt, and D. Baltimore. 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature (London)* 329:174-178.