Initiation of Transcription at the Human Terminal Deoxynucleotidyl Transferase Gene Promoter: A Novel Role for the TATA Binding Protein†

David H. Sorscher,* Baoli Yang, Dipa Bhaumik, Theoni Trangas, Anne V. Philips, Karen E. Chancellor, and Mary Sue Coleman

Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, North Carolina 27599-7260

Received April 29, 1994; Revised Manuscript Received July 1, 1994®

ABSTRACT: Control of initiation of transcription of the human terminal deoxynucleotidyl transferase (TdT) gene was investigated by using an *in vitro* transcription assay. The precise contribution of discrete basal promoter elements to transcription initiation was determined by testing deletion and substitution mutations. The primary element, contained within the region spanning −34 to −14 bp relative to the transcription start site, accounted for 80% of basal promoter activity. TdT promoter activity required the sequence ACCCT at −24 to −20 bp since a dramatic decrease in transcription initiation was observed after mutation of this sequence, whereas mutation of the adjacent sequence from −32 to −25 bp did not alter promoter activity. The secondary element contained sequences surrounding the transcription start site and had 20% of promoter activity. Deletion of both elements completely abolished transcription initiation. Initiator characteristics of the secondary element were revealed by using the *in vitro* assay: promoter sequences at the transcription start site were sufficient to direct accurate initiation at a single site. Mutation of the sequence GGGTG spanning the transcription start site resulted in loss of transcription initiation. Both the primary and secondary elements were nonhomologous to corresponding regions from the mouse TdT gene promoter. While the human basal promoter functioned in the absence of TATA consensus sequences or GC-rich SP1 binding sites, it was dependent on active TFIID. In contrast to other TATA-less promoters, purified TATA binding protein substituted for the TFIID complex and restored promoter activity to TFIID-inactivated nuclear extracts.

Terminal deoxynucleotidyl transferase (TdT)[1] is a template-independent DNA polymerase which randomly polymerizes deoxynucleoside triphosphates onto initiator DNAs (Bollum, 1974). TdT is expressed only in pre-B- and pre-T-cells of the developing immune system, and the peak in this activity coincides with recombination of Ig and TCR gene fragments (Alt & Baltimore, 1982; Leiber et al., 1987; Blackwell & Alt, 1988; Lo et al., 1991). During recombination and joining of V, D, and J segments, the random addition of nucleotides (N regions) at both sides of the D element greatly increases the diversity of Ig and T-cell receptor molecules (Desiderio et al., 1984; Yancopoulos et al., 1986; Gu et al., 1990). The precise correlation between TdT expression and the appearance of N regions, and the ability of TdT to act as a template-independent DNA polymerase, led to the suggestion that TdT was responsible for N region addition (Blackwell & Alt, 1984). Two recent reports provide conclusive evidence that TdT activity is responsible for N region addition during the recombination process (Komori et al., 1993; Gilfillan et al., 1993). In these studies, mice deficient in TdT were created by inactivation of the TdT gene through homologous recombination. These TdT knockout mice lacked N regions between V, D, and J segments of Ig and TCR sequences, and Komori et al. (1993) have postulated that TdT is a tissue-specific component of the V(D)J recombinase.

An important step in understanding regulation of the V(D)J recombinase and how it functions during immunocyte differentiation is to identify mechanisms underlying expression of its component genes. The mouse and human TdT genes have been available for investigation. In both species, TdT is encoded by a single gene (Riley et al., 1988; Landau et al., 1984; Kallenbach et al., 1992). The murine TdT gene promoter contains an initiator (Inr) element spanning the transcription start site coupled with a TATA-less 5′ region (Smale & Baltimore, 1989; Smale et al., 1990). This element directs basal transcription of the mouse TdT gene without conferring tissue specificity (Smale & Baltimore, 1989; Lo et al., 1992; Zenzie-Gregory et al., 1993). Limited tissue specificity has been identified in another element located 60 bp upstream of the transcription start site that contains a LyF-1 binding element (D′) and an adjacent element which binds a member of the Ets family of lymphoid-specific transcription factors (Lo et al., 1991; Ernst et al., 1993).

The sequences 5′ to the human TdT gene show limited homology to other eukaryotic promoters (Riley et al., 1988; Coleman et al., 1992). Similar to mouse TdT, the human gene is regulated at the transcriptional level (Trangas & Coleman, 1989). We have defined a basal promoter for the human TdT gene by using a deletional analysis of TdT 5′ flanking sequences *in vivo* in immature TdT$^+$ and TdT$^-$ lymphocyte cell lines (Bhaumik et al., 1994). The minimal

† This work was supported by a grant from the National Cancer Institute, RO1CA19492, and from the Tobacco Research Council, 2872R1.

* To whom correspondence should be addressed. Phone: 919-966-6292; Fax: 919-966-2852; email: dhs@med.unc.edu.

® Abstract published in *Advance ACS Abstracts,* August 15, 1994.

[1] Abbreviations: bp, base pair(s); cDNA, complementary DNA; CAT, chloramphenicol acetyltransferase; dNTP, deoxyribonucleoside 5′-triphosphates; DEPC, diethyl pyrocarbonate; HEPES, *N*-(2-hydroxyethyl)-piperazine-*N*′-(2-ethanesulfonic acid); Ig, immunoglobulin; Inr, initiator; NTP, ribonucleoside 5′-triphosphates; PEG, poly(ethylene glycol); PMSF; phenylmethanesulfonyl fluoride; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; TBP, TATA box binding protein; TCR, T-cell receptor; TdT, terminal deoxynucleotidyl transferase; TdT$^+$, cells expressing TdT; TdT$^-$, cells not expressing TdT; V, D, J, variable, diversity, joining.

0006-2960/94/0433-11025$04.50/0

promoter region for the human TdT gene contains no apparent homology to other known promoter elements, such as the TATA box or GC-rich SP1 binding sites. In contrast to promoters with a consensus TATA box, promoters lacking this element (such as the Inr-containing mouse TdT promoter) require functional TFIID for activity, and purified TBP does not, by itself, reconstitute transcription in the absence of the other TFIID components (Smale et al., 1990; Pugh & Tjian; 1991; Zhou et al., 1992; Hernandez, 1993).

We have found that analysis of transcription initiation *in vivo* is limited by the low level of transcripts synthesized by TdT promoter expression constructs after transient transfection (Bhaumik et al., 1994). We describe here a sensitive *in vitro* transcription assay which we developed to identify mechanistic features of the human TdT basal promoter. This *in vitro* analysis permitted quantitation of the role of defined sequences within the primary and secondary elements of the basal promoter to the accuracy and efficiency of transcription initiation. The cell-free system was also exploited to investigate the role of the TFIID complex and its DNA binding component, TBP, during the initiation of human TdT gene transcription. Interestingly, purified human TBP substituted for the TFIID complex during initiation of transcription from the TATA-less human TdT promoter.

EXPERIMENTAL PROCEDURES

Cell Lines and Culture Conditions. The cells (kindly provided by Dr. J. Minowada, Fujisaki Cell Center, Okayama, Japan) used in this study were established cell lines of leukemia–lymphoma origin and were found to represent different, early stages of lymphocyte maturation (Srivastava & Minowada, 1983). The human lymphocytic cell lines Nalm 6 (pre-B), KT1 (pre-T), and MOLT4 (pre-T) express human TdT mRNA. No TdT transcripts were observed in the TdT$^-$ KE37 (T) cell line. HeLa cells (TdT$^-$) were obtained from the Tissue Culture Facility at the UNC Lineberger Cancer Research Center. Cells were maintained in suspension culture in an humidified atmosphere at 37 °C and 5% CO_2. Growth media (RPMI-1640, Gibco) contained 10% fetal bovine serum and penicillin/streptomycin. Cell growth was monitored closely to maintain an actively dividing culture in a logarithmic growth phase. Transcriptionally active nuclear extracts were prepared from cells which were expanded, in spinner flasks (Bellco glass), to 3.5 or 4.5 L of actively growing cells.

Plasmid Templates. Human TdT genomic and cDNA sequences were obtained from λ phage clones which contained the TdT gene (Riley et al., 1988). These TdT sequences were cloned into the Stratagene plasmid vector, pBSK$^+$, to generate templates for *in vitro* transcription. A 635 bp *Bam*HI to *Eco*RI fragment containing the first three exons of human TdT (Riley et al., 1988) was ligated into the *Bam*HI and *Eco*RI sites of pBSK (pTdT123). The numbering for the plasmid templates corresponds to the distance in bp upstream from the transcription start as determined by 5′RACE (Bhaumik et al., 1994). The −930 bp template (total size = 4.4 kb) for *in vitro* transcription was obtained by isolation of a *Bam*HI fragment containing 995 bp of sequence immediately 5′ to the TdT translation start site and ligation of this fragment to *Bam*HI-digested pTdT123. Forward (−930) and reverse (−930R) orientations of the −930 bp template were used in the *in vitro* transcription assay. The next 7 kb of sequence 5′ from −930 bp to −8 kb was isolated by digestion with *Bam*HI. The −930 plasmid was linearized by partial digestion with *Bam*HI. The −8 kb template (total size = 12 kb) for *in vitro* transcription was generated by ligation of *Bam*HI-linearized −930 plasmid to the 7 kb *Bam*HI fragment. PCR subcloning was used to generate the −157, −34, −14, and +21 bp templates. The start site mutation (SSMT) template carries a 5 bp mutation at the transcription start site. The PM1 and PM2 templates contain 8 and 5 bp mutations, respectively, in the primary TdT promoter element. PCR mutagenesis was carried out according to Higuchi et al. (1988) to create SSMT, PM1, and PM2. The mutagenic internal primers for SSMT (TdT sequences are underlined, mutagenic sequences in bold) were 5′-TCTTGTAGGA**TTTGT**GCAGTCTCCCTC-3′ (SSMTF) and its complement (SSMTR). The mutagenic primers for PM1 and PM2 were 5′-GCGGGTGAGATG**CACGA**CCCAC-CCTTCGTGTA-3′ (PM1F) and 5′-GACATCAAA-**CAAAG**TCGTGTAGGAGG-3′ (PM2F) and their complements. The −54 bp construct was used as template for PCR mutagenesis in creating the SSMT, PM1, and PM2 plasmids. The −54 bp template was created by isolation of an *Eco*RI/*Bam*HI fragment (117 bp containing sequences immediately 5′ of the ATG translation start site) from the plasmid −54CAT (Bhaumik et al., 1994), incubation of this fragment with Klenow to create blunt ends, and then ligation to pTDT123 which had been linearized by digestion with *Bam*HI and blunt-ended by incubation with Klenow. Identification and orientation of the −54 plasmid was confirmed by DNA sequencing. The sequences of −157, −34, −14, +21, SSMT, PM1, and PM2 plasmids were also confirmed by DNA sequencing.

Preparation of Nuclear Extracts. Nuclear extracts were prepared by a modification (Zeleznik-Le et al., 1991) of the Dignam et al. (1983) protocol. To prepare nuclei, lymphocytes were Dounce homogenized in buffer C (20 mM HEPES, pH 7.9, 0.75 mM spermidine, 0.15 mM spermine, 0.2 mM EDTA, 2.0 mM EGTA, 2.0 mM DTT, 20% glycerol, 1.0 mM PMSF) containing 1/10 volume buffer B (75% sucrose, 50 mM HEPES, 0.75 mM spermidine, 0.15 mM spermine, 0.2 mM EDTA, 10 mM KCl, 0.2 mM DTT, 2.5% glycerol). HeLa cells were Dounce homogenized in buffer C, alone. Nuclei were pelleted by rapid centrifugation (30 s) at 15K rpm (Sorvall SS34 rotor) and resuspended in 0.6–0.75 mL of buffer C/10^9 cells (1.5–2.0 mL of buffer C/10^9 HeLa cells). Nuclear proteins were extracted by the addition of NaCl to 0.4 M, followed by vortexing (30 s) and rocking for 20 min. DNA was pelleted by centrifugation at 47 500 rpm for 45 min in a Beckman SW60 rotor. Nuclear extracts were dialyzed against buffer D (20 mM HEPES, pH 7.9, 100 mM KCl, 12.5 mM $MgCl_2$, 0.2 mM EDTA, 0.2 mM EGTA, 2.0 mM DTT, 20% glycerol, 1.0 mM PMSF), aliquoted, and frozen in liquid nitrogen before storage at −135 °C.

In Vitro Transcription. The final reaction volume was either 30 or 60 μL. All plasmids were purified by PEG precipitation followed by CsCl gradient centrifugation. In the 30-μL reaction, 4 μL of plasmid template (4 μg) at 1 μg/μL in DEPC-treated H_2O was mixed with 5 or 15 μL of nuclear extract (5 μg/μL, unless indicated otherwise, in buffer D), 1 μL of 10 mM NTPs (BRL), 1 μL of 0.2 M DTT, and 1 μL of RNASIN (Promega), and buffer D (see above) was added to 30-μL total volume. In the 60-μL reaction, 40 μL of nuclear extract and 2 μL of 10 mM NTPs were used. Twelve micrograms of the −8 kb plasmid was used to achieve a concentration equimolar to that used for other constructs. The transcription reactions were incubated for 45 min at 30 °C, and then 200 μL of stop buffer (0.2 M NaCl, 0.02 M EDTA, 1% SDS, 250 μg/mL yeast tRNA) was added. Samples were extracted with phenol/chloroform, and the organic phase was back-extracted by the addition of 125 μL of stop buffer (without yeast tRNA). Samples were extracted with chloroform and

ethanol precipitated. Nucleic acid pellets were then resuspended in 50 μL of DNase I digestion buffer (20 mM Tris-HCl, pH 7.9, 6 mM $MgCl_2$, 10 mM NaCl, 3 mM DTT), 1 μL of RNASIN (Promega) and 1.5 μL of RNase-free DNase I (RQ1 DNase I from Promega at 1 unit/μL) were added, and samples were incubated at 37 °C for 20 min. DNase I stop buffer (10 mM EDTA, 150 mM NaCl, 0.2% SDS, 0.2 mg/mL yeast tRNA) was added (100 μL), and samples were extracted with phenol/chloroform and ethanol precipitated.

Primer Extension. RNA synthesized *in vitro* was analyzed by primer extension according to the following protocol. Primers were end-labeled, and unincorporated nucleotides were eliminated by ethanol precipitations (Sambrook et al., 1989). Primers used were 5′-GGCCATCAAGGCACCCGTCTGC-CGGGGTC-3′ (224–) and 5′-GGGATATATCAACGGTG-GTATATCCAGTG-3′ (CATP1–). Primers were purified by separation on 6% neutral polyacrylamide gels (Sambrook et al., 1989) before use in the primer extension reaction. The 224– primer was complementary to exon 1 sequences 72 bases downstream of the ATG translation start site and was used to detect RNA synthesized with all of the templates except for the CAT templates (Figure 4). RNA synthesized with the CAT templates (Figure 4) was detected with CATP1– which hybridizes to sequences 43 bases downstream of the CAT translation start site. RNA was ethanol precipitated in the presence of 5×10^5 to 1×10^6 cpm of ^{32}P-end-labeled primer and resuspended in 30 μL of hybridization buffer (0.25 M KCl, 0.01 M Tris-HCl, pH 7.9, 1 mM EDTA). Samples were heated at 85 °C for 8 min and immediately transferred to 50 °C for 1 h, cooled slowly to room temperature, and then ethanol precipitated. Pellets were resuspended in 20 μL containing 50 mM Tris, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM dNTPs, 20 units of RNASIN (Promega), and 200 units of MMLV reverse transcriptase (BRL). After 1 h at 37 °C samples were extracted with phenol/chloroform, ethanol precipitated, and resuspended in 10 mM Tris-HCl, pH 7.9, 1 mM EDTA (4 μL), and formamide loading dye (7 μL) before separation on 6% polyacrylamide/8 M urea sequencing gels. Primer extension cDNA products were visualized by autoradiography and sizes determined by comparison with ^{32}P-end-labeled φX174/*Hae*III markers and DNA sequencing ladders. cDNA products were quantitated by laser scanning densitometry (Molecular Dynamics).

DNase I Protection Assay. The probe for DNase I footprinting was prepared by digestion of the –54CAT plasmid with *Eco*RI, dephosphorylation with calf intestinal alkaline phosphatase, and transfer of ^{32}P from [γ-^{32}P]ATP to the *Eco*RI site by using T4 polynucleotide kinase. After digestion with *Bam*HI, the ^{32}P-labeled 117 bp fragment was then isolated after separation on a neutral, 6% polyacrylamide gel. Incubation of nuclear extract (30 or 60 μg) with 2×10^5 cpm of ^{32}P-labeled probe in the presence of 8 mM $MgCl_2$ and 7.5 μg (with 30 μg of extract) or 15 μg (with 60 μg of extract) of poly(dI-dC)·(dI-dC) was carried out for 30 min at 24 °C in a final volume of 120 μL. $CaCl_2$ was then added to 2 mM and DNase I (Worthington Biochemicals) was added to 1.7 μg/mL and incubated at 24 °C for 4 min. In the absence of extract, samples were incubated with 0.05 μg/mL DNase I. The DNase I reaction was stopped by the addition of 240 μL of stop buffer (100 mM Tris-HCl, pH 7.9, 20 mM EDTA, 0.1% SDS, 0.1 mg/mL proteinase K, and 0.1 mg/mL yeast tRNA). Samples were incubated for 20 min at 37 °C, extracted with phenol/chloroform, and ethanol precipitated, and 40000 cpm/sample was separated on 8% polyacrylamide/8 M urea gels and visualized by autoradiography.

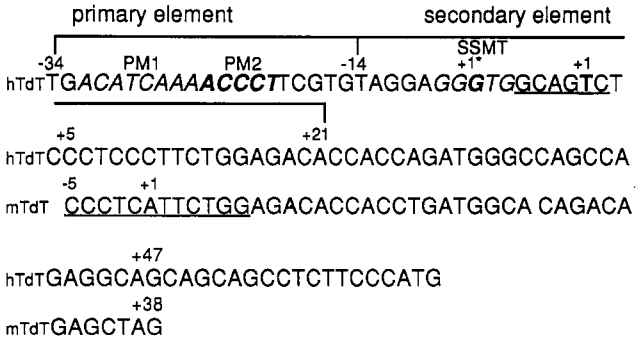

FIGURE 1: Sequence of the human TdT basal promoter. The primary element (–34 to –14 bp) and the secondary element (–14 to +21 bp) are shown. Also shown in *italics* are the sequences mutated to create the transcription start site mutation (SSMT), *GGGTG*, and primary element mutations PM1, *ACATCAAA*, and PM2, *ACCCT*. The transcription start site observed by *in vitro* transcription (this report) and *in vivo* transient transfection analyses (Bhaumik et al., 1994) is shown (+1* **G**). The start determined by 5′RACE (+1 **T**) (Bhaumik et al., 1994) is slightly different, and all constructs were numbered relative to this putative start site. Underlined is the homology (5 out of 6) to the Inr consensus sequence described by Javahery et al. (1994). The mouse TdT promoter 5′ sequences are aligned to homologous sequences in human TdT (83%, as determined using the BLAST program). The mouse TdT Inr sequence is underlined. Note that this homology required alignment of the –5 to +38 in the mouse with +5 to +42 in the human sequence. Another homology (90%) to mouse TdT is not shown and requires aligning –2 to –32 in the mouse with –20 to –50 in the human sequence.

Electrophoretic Mobility Shift Assay. The DNA probe used was a doubled-stranded oligonucleotide (27 bp) end-labeled with [γ-^{32}P]ATP and polynucleotide kinase. In each assay, 2×10^4 to 3×10^4 cpm of probe (0.02–0.05 pmol) and 5 μg of Nalm 6 extract were incubated in 40 mM Tris-HCl, pH 7.5, 200 mM NaCl, 0.2 mM DTT, 5 mM $MgCl_2$, 0.05% NP-40, 4% glycerol, and 2.5 μg of poly(dI-dC)·(dI-dC) in a final volume of 20 μL for 30 min at 24 °C. In competition reactions, samples were preincubated with unlabeled oligonucleotides for 5 min before labeled probe was added to initiate the reaction. DNA–protein complexes were analyzed on 8% polyacrylamide gels at 4 °C and visualized after autoradiography.

RESULTS

An in Vitro Transcription Assay for the Human TdT Gene Promoter. We have established an *in vitro* transcription system to identify human TdT promoter sequences which are necessary for efficient initiation of transcription and to test factors important for accurate gene transcription (see Figure 1 for promoter region sequence). To optimize the *in vitro* transcription assay, we investigated the effects of template length, source of extract, and physical form of the DNA. Nuclear extracts from TdT^+ and TdT^- cell lines were used, and comparable results were obtained with all types of nuclear extract. The templates contained varying lengths of 5′ flanking sequences, and each included the first 3 exons of the TdT coding region. Detection of the RNA products synthesized *in vitro* was performed using a primer extension assay employing an antisense primer. The transcription reaction was dependent on both template and nuclear extract (Figure 2A). The reaction product, the length of which also predicts the transcription start site, is indicated by an arrow. Both the templates containing 157 or 930 bp upstream from the transcription start functioned in the assay. These analyses indicated that transcription was initiated at 69 bp upstream from the ATG. The efficiency of linear *vs* supercoiled templates was also investigated (Figure 2B). Extracts from

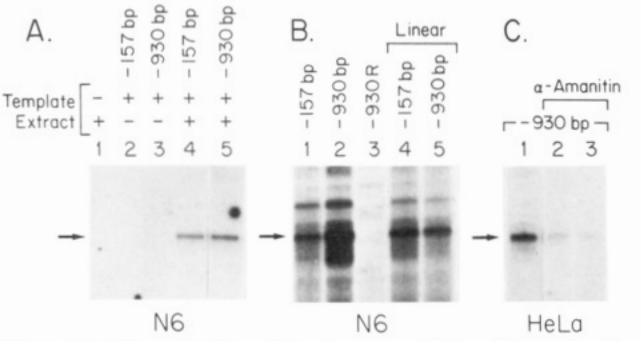

FIGURE 2: Characteristics of *in vitro* transcription directed by the human TdT promoter. (A) 200 μg of TdT+, Nalm 6 (N6) nuclear extract with 4 μg of the –157 (lane 4) or the –930 (lane 5) bp templates. No template was included with 200 μg of the Nalm 6 extract (lane 1); and no extract with 8 μg of the –157 (lane 2) or –930 (lane 3) bp templates. (B) Conditions were similar to (A): 4 μg of –157 (lanes 1 and 4) or –930 (lanes 2 and 5) bp template were used. Linear templates (lanes 4 and 5) and the –930 bp sequences in the reverse orientation (930R, lane 3) were tested. (C) 200 μg of TdT-, HeLa nuclear extract with 4 μg of –930 bp template (lanes 1–3). α-Amanitin was included at 1 μg/mL (lane 2) or 10 μg/mL (lane 3). Plasmid templates were supercoiled. The arrows point to cDNA products from primer extension.

NALM6 cells were used in these reactions. The supercoiled templates, either –157 or –930 bp (lanes 1 and 2) supported transcription more efficiently than linear templates (lanes 4 and 5). The –930 bp template (supercoiled) in the reverse orientation was ineffective in the reaction (lane 3). Nuclear extracts from the nonlymphoid cell line, HeLa, also supported transcription with the TdT promoter templates (Figure 2C). α-Amanitin inhibited transcription in HeLa cell nuclear extracts (Figure 2C, compare lane 1 with lanes 2 and 3), demonstrating that RNA polymerase II is required for transcription directed by the TdT promoter. Transcripts initiated by lymphoblast nuclear extracts were also sensitive to inhibition by α-amanitin (data not shown). While transcripts were observed when as little as 0.5 μg of supercoiled template DNA was added to the reaction, the optimal amount of template plasmid DNA was found to be 4 μg (data not shown).

Identification of Active Promoter Elements. Additional 5′ sequences were tested *in vitro* to define the boundaries of 5′ and 3′ elements of the human TdT gene promoter. Deletion analysis demonstrated that basal activity *in vitro* of the human TdT promoter resided within the 54 bp upstream from the transcription start site (Figure 3, lane 3). The transcript synthesized by the –54 construct is longer than that observed with the other constructs because of the deliberate addition of 4 bases during cloning (see Experimental Procedures). Deletion of sequences to –14 bp relative to the transcription start site reduced transcription by more than 4-fold in HeLa nuclear extracts (Figure 3B, lane 4). When additional sequences were deleted (to +21 bp), accurate transcription was abolished. Multiple, nonspecific primer extension products were observed with this +21 bp construct, and all of these cDNA products were from transcription initiated in the vector sequence (Figure 3B, lane 5).

In lymphoid nuclear extracts (in contrast to HeLa extract), the –8 kb template had dramatically greater promoter activity than the –930 and –54 bp (compare lanes 1 with 2–3 in Figure 3A *vs* Figure 3B). However, this apparent activity was an artifact of the reaction conditions. When we adjusted the amount of total DNA in lymphoid extracts to 12 μg by addition of 8 μg of promoterless plasmid DNA (pBSK+) to the reactions containing 4 μg of the –54 or –930 templates, the extent of

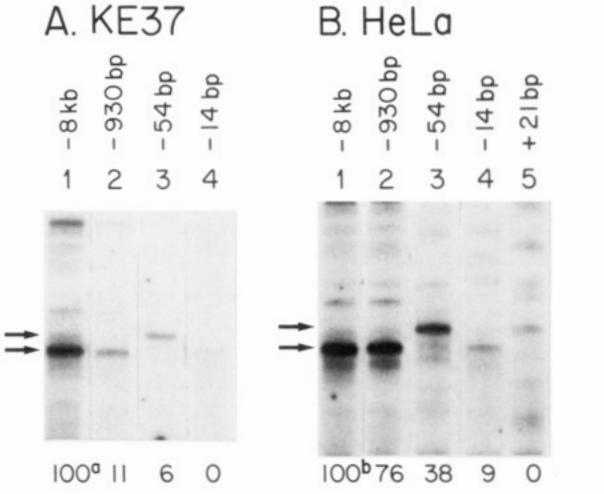

FIGURE 3: Transcription initiation from TdT promoter constructs with differing lengths of 5′ flanking sequence. Nuclear extracts were at a protein concentration of 5 μg/μL. (A) 5 μL of TdT-, KE37 nuclear extract with 4 μg of template as indicated. 12 μg of the –8kb construct was used. (B) 40 μL of HeLa nuclear extract with template as indicated. [a,b]Transcripts were quantitated by laser scanning densitometry and were expressed as a percent of the highest signal observed.

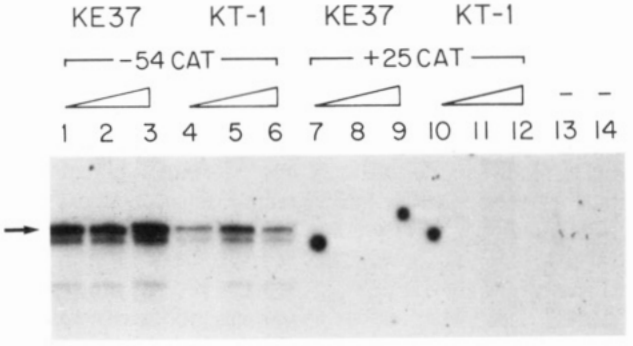

FIGURE 4: Deletion to +25 bp completely abolished transcription. 5, 15, or 40 μL of nuclear extract (5 μg/μL) was incubated with –54 or +25 bp templates. No template (lane 13) and no extract (lane 14) were included as negative controls. The triangles above the lanes in this figure and in Figure 6 indicate an increasing concentration of nuclear extract.

transcription initiation by nuclear extracts from lymphocytes and from HeLa cells was similar from the –8 kb, the –930 bp, and the –54 bp templates (data not shown).

Identification of the 3′ Boundary for Promoter Activity. The experiments above identified two elements of the basal promoter. Since a third downstream control element was detected (+25 to +40 bp) in a transient transfection assay (Bhaumik et al., 1994), we tested alternate vectors for *in vitro* activity from the third element. In these reactions, the primer was complementary to CAT mRNA, and nuclear extracts were from TdT+ and TdT- cells. Templates contained either the basal promoter elements (–54CAT) or no elements (+25CAT). The concentration of nuclear extract was varied from 0.83 to 3.3 μg/μL. Consistent with previous results, a 178 bp cDNA primer extension product reflecting accurate initiation of transcription from the TdT promoter was observed only in reactions containing the –54 template (Figure 4, lanes 1–6). In reactions containing the +25CAT construct (lanes 7–12) no reaction products were detected. Taken together (Figures 3 and 4), these results map the human TdT promoter sequences necessary for basal transcription between 54 bp 5′ and 21 bp 3′ of the transcription start site. Also, two promoter elements existed within this region; the primary element was

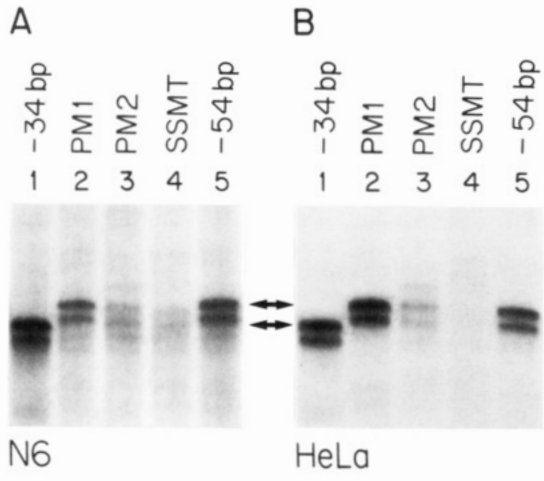

FIGURE 5: *Mutagenesis of human TdT promoter elements.* Fifteen μL of Nalm 6 (5A) or HeLa (5B) nuclear extract was incubated with the –34 (lanes 1) or –54 bp (lanes 5) templates. Templates carrying mutations in the primary promoter elements (PM1, lanes 2; PM2, lanes 3) and at the transcription start site (SSMT, lanes 4) were also incubated with Nalm 6 and HeLa nuclear extract.

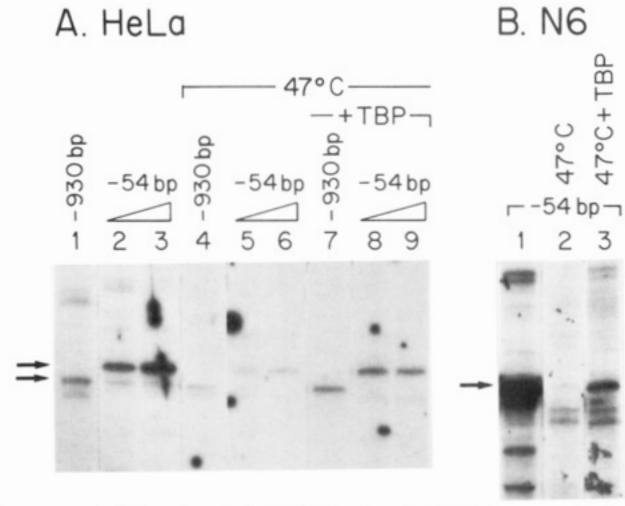

FIGURE 6: The role of the TFIID subunit, TBP, in transcription initiated from the human TdT promoter. (A) 15 μL (lanes 1, 2, 4, 5, 7, 8) or 40 μL (lanes 3, 6, and 9) of HeLa nuclear extract with –930 (lanes 1, 4, and 7) or –54 (lanes 2, 3, 5, 6, 8, and 9) bp templates. Extracts were preincubated for 15 min at 47 °C to inactivate TFIID (lanes 4–9), and 1 μL (20 ng) of purified TBP was added (lanes 7–9). (B) 15 μL of TdT$^+$, Nalm 6 lymphocyte nuclear extract with the –54 bp template. Lanes 2 and 3, heat-treated extract, and 1 μL of TBP was added in lane 3. Purified, recombinant human TBP was obtained from Promega (Madison, WI).

within –54 to –14 bp and the secondary element mapped between –14 and +21 bp.

Mutagenesis of Promoter Elements. Four constructs containing mutations within TdT basal promoter elements were cloned to identify sequences necessary for correct transcription initiation *in vitro*. These constructs were tested in the *in vitro* transcription assay using extracts from Nalm 6 and HeLa cells. Twenty base pairs at the 5′ end of the –54 construct were deleted to generate the –34 template. A similar level of promoter activity was observed with the –34 and –54 templates (Figure 5, compare lanes 1 with lanes 5), further narrowing the 5′ boundary of the primary element. Two adjacent substitution mutations were introduced within the primary element of the –54 template by using the PCR mutagenesis technique to test the specific roles of sequences within this region. In PM1, 8 bp from –32 to –25 were mutated, and in PM2, 5 bp from –24 to –20 bp were changed. Although PM1 had a small or no effect, PM2 resulted in an almost complete loss of transcription initiation (Figure 5, compare lanes 2 and 3). In addition, 5 bp were changed at the transcription start site (SSMT) to create a construct containing a secondary element mutation. The SSMT mutation completely abolished activity of the basal promoter (Figure 5, lanes 4). These data demonstrate that two separate 5 bp sequences, one in the primary element (ACCCT) and the other in the secondary element (GGGTG), were absolutely necessary for transcription initiation.

Purified TBP Substitutes for the TFIID Complex during Transcription Initiation from the Human TdT Promoter. The 34 bp of 5′ sequence necessary for basal human TdT promoter activity contained no candidate sites for binding by the TATA binding protein subunit (TBP) of the general transcription factor TFIID or for binding by the transcription factor Sp1. After purified TBP was incubated with human TdT promoter elements, no binding was observed by gel mobility shift or DNase I footprinting assays (data not shown). We were interested in determining, then, the role for TFIID and its DNA binding subunit, TBP, in transcription from this TATA-less promoter. TFIID activity can be partially inactivated by mild heat treatment of transcriptionally active nuclear extracts (Nakajima et al., 1988; Pugh & Tjian, 1992; Zenzie-Gregory et al., 1992). We followed this protocol to determine whether inactivation of TFIID had an effect on human TdT promoter activity. Two nuclear extracts were used (HeLa and Nalm 6). The extracts were subjected to mild heat treatment and were added to the transcription reaction either with or without TBP. Heat treament of HeLa nuclear extract resulted in loss of transcription from the –930 and –54 bp templates (Figure 6A, compare lanes 1–3 with 4–6). Addition of TBP clearly restored activity to the TFIID-inactivated extracts (Figure 6A, compare lanes 4–6 with 7–9). Apparently, purified TBP, by itself, replaced the multisubunit TFIID complex in directing transcription from the human TdT promoter. In Figure 6B, the results from the TFIID inactivation experiment using lymphocyte extracts from TdT$^+$, Nalm 6 cells are illustrated. Partial inactivation of TFIID by heat treatment completely abolished transcription from the –54 bp template (Figure 6B, compare lanes 1 and 2). After adding TBP to the heat-treated extract, transcriptional activity was restored (Figure 6B, compare lanes 2 and 3).

DNA–Protein Interactions at the Human TdT Basal Promoter. Since two separate 5 bp regions were critical for functional promoter activity, we performed DNase I footprinting and gel mobility shift experiments to visualize protein binding at these two elements. Sequences from –54 to the ATG translation start (117 bp) were end-labeled on the plus strand and incubated with nuclear extract before limited digestion with DNase I. Two regions were protected from DNase I digestion (brackets in Figure 7A). Partial DNase I digestion extended from –27 to +4 bp (large bracket in Figure 7A). Within this large DNase I footprint, more complete protection from DNase I was observed from –27 to –17 bp (small bracket in Figure 7A). The nucleotides at the 5′ and 3′ boundaries of both of these regions exhibited hypersensitivity to DNase I. Another hypersensitive site was observed at the transcription start (see arrow in Figure 7A). Interestingly, this site was included in the 5 bp start site mutation (SSMT) which abolished transcription initiation. The 5 bp sequence from –24 to –20 bp, which was shown to be necessary for TdT promoter activity (see PM2, Figure 5), was in the middle of the smaller DNase I footprint.

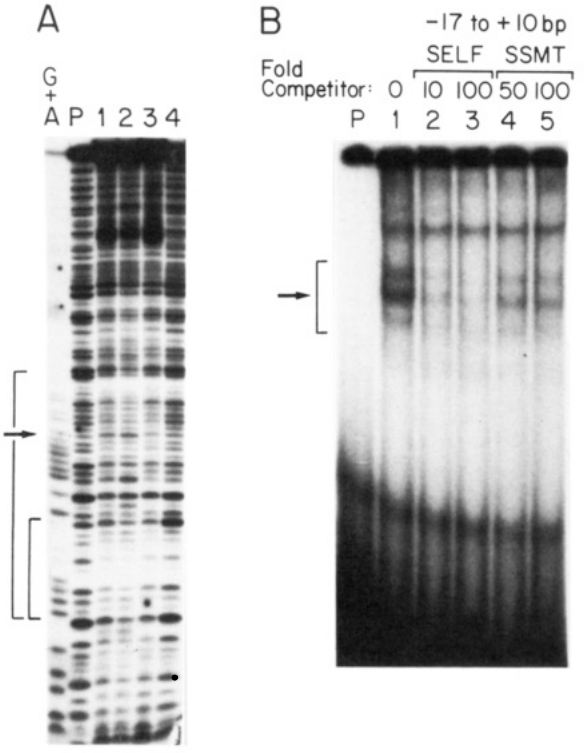

FIGURE 7: DNA–protein interactions at functionally significant sequences in the human TdT basal promoter. (A) DNase I footprinting. 30 μg (lanes 1, 3, and 4) or 60 μg (lane 2) of nuclear extract from Nalm 6 (lanes 1 and 2), KE37 (lane 3), or HeLa (lane 4) was incubated with ^{32}P-labeled 117 bp probe in a DNA binding reaction before addition of DNase I. The protected region is indicated by brackets. The arrow points to the DNase I hypersensitive nucleotide. G + A is Maxam–Gilbert G and A sequencing (Williamson & Celander, 1990). (B) Specific binding to the start site region. An electrophoretic mobility shift assay with the ^{32}P-end-labeled probe (−17 to +10 bp) and 5 μg of Nalm 6 nuclear extract. The DNA–protein complex is indicated by brackets. In lane P no protein and in lane 1 no competitor was added. Lanes 2 and 3 contain 10- and 100-fold excess of competitor (self), and lanes 4 and 5 contain 50- and 100-fold excess of a mutated competitor in which 5 bp at the transcription start site were altered (SSMT).

DNase I hypersensitivity at the transcription start site suggested a boundary for protein–DNA interactions. We proceeded to analyze this region by testing sequences from −17 to +10 bp in a gel mobility shift assay (Figure 7B). Specific binding was demonstrated by self-competition with 10- and 100-fold excess of unlabeled oligonucleotide (Figure 7B, lanes 2 and 3). Furthermore, a minimal effect on binding was observed after incubation with 50- and 100-fold excess of competitor oligonucleotide containing the 5 bp mutation at the transcription start site (Figure 7B, lanes 4 and 5).

DISCUSSION

Basal promoter elements directing transcription of protein-encoding genes by RNA polymerase II can be divided into two classes: those containing a TATA box sequence (Breathnach & Chambon, 1981) and those gene promoters such as mouse TdT (Smale & Baltimore, 1989) which are TATA-less but contain Inr sequences (Roeder, 1991; Weis & Reinberg, 1992; Hernandez, 1993; Kollmar & Farnham, 1993). TATA-less promoters that contain an Inr element are dependent on the latter for initiation of transcription at a single site (Hernandez, 1993; Weis & Reinberg, 1993; Kollmar & Farnham, 1993). Exceptions to these two classes are promoters which lack both Inr and TATA elements and initiate transcription from many sites (Zenzie-Gregory et al., 1992; Geng & Thompson, 1993).

We describe in this report the development of an *in vitro* system to identify mechanistic features of transcription initiation at the human TdT promoter. This experimental system is of particular importance because the primary and secondary elements of the human TdT gene basal promoter are not homologous to known transcription factor binding sites (the sequence is illustrated in Figure 1) or to the corresponding promoter sequences for mouse TdT (Smale & Baltimore, 1989; Lo et al., 1991). Thus, the human TdT promoter appears to have novel characteristics.

Promoter elements for the human TdT gene were active in the *in vitro* transcription assay, and transcription was initiated at the same start site as that observed *in vivo* (Bhaumik et al., 1994). The primary element, which conveyed 80% of the activity of the basal promoter, resided within the 20 bp of sequence from −34 to −14 bp and specifically spanned the sequence ACCCT at −24 to −20 bp. This 5 bp element was also included in the region protected from DNase I digestion. We concluded that human TdT promoter activity was mediated by an interaction between a nuclear factor and the ACCCT sequence in the primary promoter element.

Promoter activity from a secondary element–CAT construct was observed *in vivo* in a transient transfection assay, but accurate initiation of transcription *in vivo* was not confirmed by primer extension of RNA synthesized from this CAT construct (Bhaumik et al., 1994). By using the sensitive *in vitro* transcription assay described herein, we were able to demonstrate conclusively accurate transcription initiation by the secondary promoter element. This secondary element was functionally important because promoter activity was not completely abolished unless sequences from −14 to +21 were deleted. The observation that this element directed accurate transcription beginning at a single site and had, by itself, some promoter activity was reminiscent of Inr sequences described by others (Weis & Reinberg, 1992; Kollmar & Farnham, 1993; Jahavery et al., 1994). The 5 bp sequence GGGTG (Figure 1) at the human TdT transcription start site (within the secondary element) was essential for accurate and efficient initiation of transcription. This finding suggests that the GGGTG sequence must be an integral part of the human TdT Inr. However, there was no homology between this sequence and the mouse TdT Inr element. Recently, loose consensus Inr sequences have been described (Kollmar & Farnham, 1993; Javahery et al., 1994), and these authors (Javahery et al., 1994) note that a wide range of sequences convey Inr activity. Although the sequence $GGG_{+1}TG$ within the human TdT Inr does not fit the Inr consensus sequence ($PyPyA_{+1}NPyPy$; Javahery et al., 1994), an adjacent sequence $CA_{+1}GTCT$ two bp downstream, and underlined in Figure 1, does have homology to the Inr consensus. We have no evidence that transcription initiates at the A in the CAGTCT downstream sequence. Perhaps the sequence GGGTG which is required for human TdT promoter activity cooperates with adjacent sequences to impart Inr function.

The presence of a negative regulatory element between 54 and 356 bp upstream of the TdT transcription start site was observed *in vivo* by using the transient transfection assay (Bhaumik et al., 1994). However, inhibitory effects of this element were not observed in *in vitro* transcription reactions. The transfected DNA reporter plasmids used in the transfection assay associate with histone proteins while inside the cell and assemble into a nucleosomal structure (Reeves et al., 1985). In contrast, the human TdT promoter templates employed in the *in vitro* transcription reactions were not preincubated with histones. Since long-range effects of

transcriptional enhancers or repressors are not observed *in vitro* in the absence of added histones, *in vitro* evidence of the negative element upstream of the TdT promoter may require packaging of template DNA into a pseudo-chromatin structure before initiation of the transcription reaction (Laybourn & Kadanoga, 1992).

Interestingly, although there was extensive homology in the 5′ flanking sequences for mouse and human TdT (Ernst et al., 1993), the region identified as the human TdT basal promoter (−34 to +21 bp) had no homology to the corresponding sequence from mouse. However, there is limited conservation in promoter function between mouse and human TdT since both basal promoter elements include a functional Inr sequence. In contrast to the basal promoter for the mouse TdT gene which was contained in the Inr element (Smale & Baltimore, 1989), most of the activity for the human TdT basal promoter resided in sequences upstream of the transcription start site from −34 to −14 bp.

We have found that TFIID was also involved in the function of the human TdT basal promoter and that it functioned in the absence of Sp1. TFIID is composed of a TATA box binding polypeptide (TBP) and several tightly associated factors (TAFs) (Dynlacht et al., 1991; Tanese et al., 1991). TATA-containing promoters require only the TBP component of TFIID for initiation of transcription (Weis & Reinberg, 1992; Hernandez, 1993; Struhl, 1994). In contrast, transcription from TATA-less promoters has been shown to require the intact TFIID complex (Smale et al., 1990; Pugh & Tjian, 1991; Zhou et al., 1992). It has been postulated, as a result of heat inactivation and immunodepletion studies, that TBP was necessary but not sufficient for transcription from TATA-less, Inr-containing promoters and that a heat-labile tethering component of the TFIID complex was inactivated during heat treatment of the nuclear extract (Pugh & Tjian, 1991). The tethering factor may bridge Sp1 with the general transcription machinery through interactions with TFIID (Pugh & Tjian, 1991; Weis & Reinberg, 1992).

The Inr element in the human TdT gene was not analogous to other TATA-less, Inr promoters like that in the mouse TdT gene because the addition of purified TBP alone to TFIID-inactivated extracts was sufficient for restoration of transcription. This dramatic effect of TBP on transcription from the TdT promoter was observed in the absence of evidence of a stable protein–DNA interaction with TBP that can be detected by the gel mobility shift or DNase I footprinting assays. Activation of transcription by the addition of TBP to extracts in which TFIID was inactivated by heat treatment suggested that TBP was the only component of the TFIID complex necessary for basal human TdT promoter activity. However, transcription initiation in the heat-treated extract may have been restored by an interaction between the exogenous TBP and one of the subunits of TFIID that was resistant to heat treatment. Our results with the human TdT basal promoter were reminiscent of data from experiments employing the AdML promoter (Carcamo et al., 1991). These investigators found that TBP was able to substitute for the TFIID complex in directing transcription from a mutated AdML promoter in which TATA box sequences had been deleted. However, the human TdT promoter is the first such natural promoter described in which TBP alone is effective in supporting transcription.

Since TdT expression occurs only during lymphocyte maturation, authentic transcriptional control must include mechanisms restricting expression to cells of lymphocytic lineage and specifically to pre-B- and pre-T-cells. Such enhancer or suppressor control elements may exist several kilobases distal to the basal promoter sequences described in the current study. Lymphocyte-specific or stage-specific expression may be detected *in vitro* but will most likely require reconstitution of chromatin by the association of histones with template DNA instead of using naked DNA templates alone (Laybourn & Kadonaga, 1992).

Our data suggest the following model for activity of the human TdT gene promoter. Basal transcription in lymphocytes and in HeLa cells is mediated by the interaction between components of the RNA polymerase II transcription complex and the −34 to −14 bp primary element. This primary element conveys 80% of promoter activity. We propose that TBP contacts DNA in combination with another, possibly novel protein as an initial event at the promoter. This protein contacts the sequence ACCCT in the primary promoter element. We do not believe TBP, by itself, is responsible for positioning because no obvious binding of TBP to the basal promoter was observed. After positioning of TFIID on the primary element, the TBP-associated factors TFIIA and -B bind to TFIID followed by binding of RNA polymerase II/TFIIF, TFIIE, and TFIIH to complete formation of the preinitiation complex (Roeder, 1992; Conaway & Conaway, 1991; Kollmar & Farnham, 1993). The sequence GGGTG within the secondary element of the basal promoter, an Inr element, then interacts directly with a component of the initiation complex or with an Inr binding factor (Seto et al., 1991; Javahery, 1994) and positions RNA polymerase II and the initiation complex at a single start site.

CONCLUSIONS

Thus, from the *in vitro* system we determined the contribution of distinct sequences within the primary and secondary promoter elements to initiation of transcription of the TdT gene. A new role for TBP in directing initiation of transcription at an apparently novel TATA-less promoter was also demonstrated. Current experiments are exploiting the *in vitro* assay to identify and isolate other transcription factors necessary for basal and tissue-specific TdT gene expression.

ACKNOWLEDGMENT

The authors acknowledge Jeanne Kahn and Jane Azizkhan for suggestions in the preparation of nuclear extracts and Judith Lesnaw (University of Kentucky), Beverly Mitchell, and Aziz Sancar for reviewing the manuscript.

REFERENCES

Alt, F. W., & Baltimore, D. (1982) *Proc. Natl. Acad. Sci U.S.A. 79*, 4118–4122.

Bhaumik, D., Yang, B., Trangas, T., Bartlett, J. S., Coleman, M. S., & Sorscher, D. H. (1994) *J. Biol. Chem. 269*, 15861–15867.

Blackwell, T. K., & Alt, F. W. (1988) in *Molecular Immunology* (Hames, B. D., & Glover, D. M., Eds.) pp 1–60, IRL Press, Washington, DC.

Bollum, F. J. (1974) in *The Enzymes* on (Boyer, P., Ed.) Vol. 10, pp 145–171, Academic Press, New York.

Breathnach, R., & Chambon, P. (1981) *Annu. Rev. Biochem. 50*, 349–383.

Carcamo, J., Maldonado, E., Ahn, M., Ha, I., Kasai, Y., Flint, J., & Reinberg, D. (1990) *Proc. Natl. Acad. Sci. U.S.A. 88*, 8052–8056.

Coleman, M. S., Yang, B., & Sorscher, D. (1992) *Crit. Rev. Eukaryotic Gene Expression 2*, 237–250.

Conaway, J. W., & Conaway, R. C. (1991) *J. Biol. Chem. 266*, 17721–17724.
Desiderio, S., Yancopoulos, G., Paskind, M., Thomas, E., Boss, M., Landau, N., Alt, F., & Baltimore, D. (1984) *Nature 311*, 752–755.
Dignam, J. D., Lebovitz, R. M., & Roeder, R. G. (1983) *Nucleic Acids Res. 11*, 1475–1489.
Dynlacht, B. D., Hoey, T., & Tjian, R. (1991) *Cell 66*, 563–576.
Ernst, P., Hahn, K., & Smale, S. T. (1993) *Mol. Cell. Biol. 13*, 2982–2992.
Geng, Y., & Johnson, L. F. (1993) *Mol. Cell. Biol. 13*, 4894–4903.
Gilfillan, S., Dierich, A., Lemeur, M., Benoist, C., & Mathis, D. (1993) *Science 261*, 1175–1178.
Gu, H., Forster, I., & Rajewsky, K. (1990) *EMBO 9*, 2133–2140.
Hernandez, N. (1993) *Genes Dev. 7*, 1291–1308.
Higuchi, R., Krummel, B., & Saiki, R. K. (1988) *Nucleic Acids Res. 15*, 7351–7367.
Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., & Smale, S. T. (1994) *Mol. Cell. Biol. 14*, 116–127.
Kallenbach, S., Doyen, N., d'Anton, M. F., & Rougeon, F. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2799–2803.
Koiwai, O., & Morita, A. (1988) *Biochem. Biophys. Res. Commun. 154*, 91–100.
Kollmar, R., & Farnham, P. J. (1993) *Proc. Soc. Exp. Biol. Med. 203*, 127–139.
Komori, T., Okada, A., Stewart, V., & Alt, F. W. (1993) *Science 261*, 1171–1175.
Landau, N. R., St. John, T. P., Weissman, I. L., Wolf, S. C., Silverstone, A. E., & Baltimore, D. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5836–5840.
Landau, N. R., Schatz, D. G., Rosa, M., & Baltimore, D. (1987) *Mol. Cell. Biol. 7*, 3237–3243.
Laybourn, P. J., & Kadonaga, J. T. (1992) *Science 257*, 1682–1685.
Leiber, M. R., Hesse, J. E., Mizuuchi, K., & Gellert, M. (1987) *Genes Dev. 1*, 751–761.
Lo, K., Landau, N. R., & Smale, S. T. (1991) *Mol. Cell. Biol. 11*, 5229–5243.
Nakajima, N., Horikoshi, M., & Roeder, R. G. (1988) *Mol. Cell. Biol. 8*, 4028–4040.
Pugh, B. F., & Tjian, R. (1991) *Genes Dev. 5*, 1935–1945.
Reeves, R., Gorman, C. M., & Howard, B. (1985) *Nucleic Acids Res. 13*, 2279–2286.
Riley, L. K., Morrow, J. K., Danton, M. J., & Coleman, M. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2489–2493.
Roeder, R. G. (1991) *Trends Biochem. Sci. 16*, 402–407.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Smale, S. T., & Baltimore, D. (1989) *Cell 57*, 103–113.
Smale, S. T., Schmidt, M. C., Berk, A. J., & Baltimore, D. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4509–4513.
Srivastava, B. I. S., & Minowada, J. (1983) *Leuk. Res. 7*, 331–338.
Struhl, K. (1994) *Science 263*, 1103–1104.
Tanese, N., Pugh, B. F., & Tjian, R. (1991) *Genes Dev. 5*, 2212–2224.
Trangas, T., & Coleman, M. S. (1989) *Biochem. Biophys. Res. Commun. 164*, 750–757.
Weis, L., & Reinberg, D. (1992) *FASEB. J. 6*, 3300–3309.
Williamson, J. R., & Celander, D. W. (1990) *Nucleic Acids Res. 18*, 379.
Yancopoulos, G. D., Blackwell, K. T., Suh, L., Hood, L., & Alt, F. W. (1986) *Cell 44*, 251–259.
Zeleznik-Le, N. J., Azizkhan, J. C., & Ting, J. P.-Y. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1873–1877.
Zenzie-Gregory, B., O'Shea-Greenfield, A., & Smale, S. T. (1992) *J. Biol. Chem. 267*, 2823–2830.
Zhou, Q, Lieberman, P. M., Boyer, T. G., & Berk, A. J. (1992) *Genes Dev. 6*, 1964–1974.