Transcriptional Unit of the Murine Thy-1 Gene: Different Distribution of Transcription Initiation Sites in Brain

EUGENIA SPANOPOULOU, VINCENT GIGUERE,† AND FRANK GROSVELD*

Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom

Structural analysis of the mouse Thy-1.2 gene has shown that the major promoter of the gene is characterized by a tissue-specific DNase I-hypersensitive site and is located within a methylation-free island. The gene is regulated at the transcriptional level, and steady-state mRNA analysis reveals that the previously reported exon Ib contributes at most 5% of the total mRNA. The major promoter uses several transcription initiation sites within a region of 100 base pairs. The frequency of usage of these sites in brain is markedly different from that in other tissues.

The murine Thy-1 antigen is a cell-surface glycoprotein (12, 47) of 112 amino acids that is anchored to the cell membrane via a glycophospholipid tail (41, 55). Recently, a number of other eucaryotic proteins have been shown to share the same property (18); however, the functional significance of this type of anchorage remains to be studied. The protein exists in two allelic forms, Thy-1.1 (AKR mice) and Thy-1.2 (BALB/c mice), which differ by a single amino acid at position 89 (47). The mouse Thy-1 gene is tissue-specifically and developmentally regulated (24, 61). The antigen appears on the surface of cells in the central nervous system (Purkinje, cerebellum, spinal cord, sciatic nerve) during the second week of postnatal development as well as in the peripheral nervous system in almost every type of neuron and axon, with the exception of the olfactory nerve (45). In mice Thy-1 appears on thymocytes and mature T lymphocytes; however, it is absent from human and rat peripheral T cells (61). To a lesser extent, Thy-1 appears during different stages of development on several other cell types, e.g., connective tissue, epithelial cells, muscle, etc. (45). The Thy-1 protein belongs to the immunoglobulin superfamily (61), and its differentiation-specific pattern and structure suggest that it may play a role in intercellular communication during development. Recent evidence suggests that Thy-1 may be involved in the proliferation pathways of certain cell types, because anti-Thy-1 antibodies cause proliferation of murine T or transfected B lymphocytes by utilizing Ca^{2+} mobilization as a second message (25, 38). Moreover, transgenic mice were shown to develop either a lymphoid hyperplasia or a proliferative kidney disorder as a result of an augmented ectopic expression of the Thy-1 gene in either progenitor B cells (15) or the kidney tubuli, respectively (34). The structure of the Thy-1 gene is as interesting as its expression pattern. It has been reported that the gene contains two promoters separated by 210 base pairs (bp), with two first exons of 55 (24) and 158 (28, 29) bp, respectively, which are each spliced onto the same second exon (Fig. 1). Both promoters are G+C rich and lack a TATA box, which are characteristics of "housekeeping" genes.

In this paper we describe several structural and transcrip-

tional features of the mouse gene. First, we describe the extensive characterization of the G+C-rich methylation-free island and flanking regions and the identification of several DNase I-hypersensitive sites in the promoter and the gene. Second, the determination of the transcriptional rates of the gene in expressing and nonexpressing tissues by run-on assays and comparison with the steady-state levels of the corresponding mRNAs shows that this housekeeping-type promoter is transcriptionally regulated. Interestingly, this analysis revealed a different usage of transcription initiation sites from exon Ia in brain and thymus and, in contrast to an earlier report (28, 29), showed that exon Ib is only transcribed (as an exon) at very low levels in these tissues.

MATERIALS AND METHODS

DNase I sensitivity. Nuclei from C57/BL × CBA mouse tissues were prepared by the method of Linial et al. (40). The nuclear pellet was suspended in 10 mM Tris (pH 7.4)–10 mM NaCl–3 mM $MgCl_2$–0.1 mM $CaCl_2$. Digestions were performed at a DNA concentration of 1.0 mg/ml for 10 min at 37°C with the DNase I concentration ranging from 0 to 20 μg/ml. The reaction was terminated by the addition of sodium dodecyl sulfate-EDTA and incubation with proteinase K at 37°C for 2 h. The DNA was extracted with phenol-chloroform and ethanol precipitated. Purified DNA was digested with restriction enzymes, Southern blotted, and hybridized as described previously (52, 58). λ DNA was mixed into the digestion as an internal control to monitor for completion of the digests.

Transcriptional run-on assay. Run-on assays were performed by the method of Linial et al. (40). Nuclei were isolated and washed in 10 mM Tris (pH 7.4)–10 mM NaCl–5 mM $MgCl_2$ (RSB) and 0.1 mM phenylmethylsulfonyl fluoride and suspended at 3×10^7 nuclei per 100 μl in freezing buffer (50 mM Tris [pH 8.3], 40% [wt/vol] glycerol, 5 mM $MgCl_2$, 0.2 mM EDTA). Each transcription reaction consisted of 210 μl of nuclei in freezing buffer, 60 μl of runoff buffer (5× = 25 mM Tris [pH 8.0], 12 mM $MgCl_2$, 750 mM KCl, 1.25 mM ATP, 1.25 mM GTP, 1.25 mM CTP), and 40 μl of [α-^{32}P]UTP (3,200 Ci/mmol). The nuclear suspension was incubated at 30°C for 10 min; 15 μl of DNase I (200 μg/ml) in 10 mM $CaCl_2$ was added, and the mixture was incubated for 5 min at 37°C. After proteinase K treatment, phenol-chloroform extraction,

* Corresponding author.
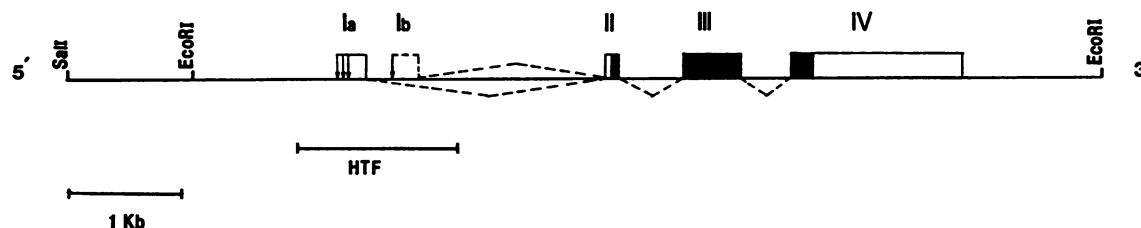† Present address: The Salk Institute, San Diego, CA 92138.

FIG. 1. Organization of the mouse Thy-1.2 gene. Exons are depicted as boxes, and introns and flanking regions are shown as lines. Filled boxes represent the coding part of the gene. Dotted lines indicate the splicing pattern. The length of the entire fragment is 9.5 kb.

and chromatography in G-50 columns to remove the unincorporated nucleotides, the purified RNA was hybridized to DNA probes slot blotted onto nitrocellulose filters for 36 h at 65°C. The filters were then washed in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate at 65°C for 2 h, incubated at 37°C in 2× SSC–10 µg of RNase A per ml for 30 min, and washed again in 2× SSC at 37°C for 1 h, followed by autoradiography with Kodak X-ray film.

RNA analysis. RNA was isolated with 3 M LiCl–6 M urea by the method of Auffray and Rougeon (2). A sample of the RNA was electrophoresed on a 1.5% agarose gel to check its quality. Poly(A)$^+$ RNA was prepared by oligo(dT) cellulose chromatography. Northern analysis was carried out as described previously (39). Single-stranded probes were generated by cloning DNA fragments into the vector pGEM-4 and carrying out Sp6 or T7 polymerase reactions. For S1 analysis, DNA probes were end labeled at one end with T4 polynucleotide kinase, and the hybridization and S1 protection reactions were carried out as described previously (6, 60). Primer extension analysis was performed with an 18-nucleotide synthetic oligonucleotide, 5' GCCTCCTCC TAGTGCGCC 3', from exon Ia. The primer was labeled with T4 polynucleotide kinase, gel purified, and dissolved in double-distilled H$_2$O. Then 10 ng of this primer was added to 7 µl of H$_2$O containing 0.5 µg of poly(A)$^+$ RNA from transgenic mice or 1 µg of poly(A)$^+$ RNA from wild-type animals. The mixture was denatured at 87°C for 10 min, 2 µl of 5× annealing buffer [50 mM piperazine-N,N'-bis(2-ethanesulfonic acid (pH 6.4), 2 M NaCl] was added; the nucleic acid was annealed at 52°C for 30 min and then at 42°C for 20 min, after which 90 µl of RT buffer (60 mM NaCl, 10 mM Tris [pH 8.3], 10 mM dithiothreitol, 1 mM dCTP, 1 mM dATP, 1 mM dGTP, 1 mM dTTP, 8 mM MgCl$_2$, 50 µg of actinomycin D per ml) was added to each reaction along with 100 U of avian myeloblastosis virus reverse transcriptase per ml. The reaction mixtures were incubated at 42°C for 1 h, phenol extracted twice, ethanol precipitated, and washed with 70% ethanol. The extension products were separated on a 12% polyacrylamide–urea gel and visualized by autoradiography.

RESULTS

Structural analysis of the mouse Thy-1 gene. The sequence of the mouse Thy-1.2 gene (14, 24, 29) shows that a region of approximately 1 kilobase (kb) in length, flanking exons Ia and Ib (Fig. 1) between position −300 from the principal transcription initiation site of the first promoter and position +700 into the first intron, contains a high frequency of CpG dinucleotides. In the remainder of the gene the dinucleotide CpG is four- to fivefold underrepresented, as in bulk vertebrate DNA (7, 23). Regions of DNA with high CpG (meth-

ylation-free islands [HTF]) have been found in many housekeeping genes (7) as well as in some tissue-specific genes (23). These observations led us to investigate the pattern of DNA methylation of the Thy-1 gene. The DNA from six tissues was digested with various restriction enzymes in combination with the methylation-sensitive enzymes HpaII and HhaI, electrophoresed through agarose gel, and transferred onto nitrocellulose filters, which were hybridized to six probes spanning the 5'-flanking region and the structural region of the gene (Fig. 2E). The results of four of these probes are shown in Fig. 2. Probes A and B covered a CpG-depleted region upstream of the gene, whereas probe C covered the CpG-rich area. The remainder of the probes covered the CpG-depleted area downstream (only probe D is shown). Probe A hybridized to a 2.6-kb EcoRI-BglII fragment that was largely resistant to further cleavage by methylation-sensitive enzymes (HpaII and HhaI), with one exception. This site was sensitive to cleavage in thymus DNA, resulting in a 0.6-kb band (Fig. 2A, lane HpaII 2). The same site might also be cleaved in nerve cells, but since these form only a small percentage of the brain it is not clearly visible (lane HpaII 1). The control MspI digest cleaved both HpaII sites (lanes MspI 1 and 2). Probe B hybridized to a region that was completely resistant to the enzymes HpaII and HhaI (Fig. 2B). Probe C hybridized to a 1.6-kb BglII-SacI fragment that was completely sensitive to further digestion. Both HpaII (indistinguishable from the MspI control) and HhaI cleaved this fragment at many sites, indicating that this area was completely unmethylated in all tissues (Fig. 2C). The downstream region was completely methylated and resistant to further cleavage by the methylation-sensitive enzymes (probe D [Fig. 2D] and others not shown) with one exception, a site close to the ApaI restriction site (Fig. 2E) that was partially methylated in all tissues. Additional restriction digests and blots showed that the methylation-free area almost exactly coincided with the probe C fragment. We conclude from these results that the Thy-1 gene contains a "classical" HTF covering the promoter and first exon(s) which is flanked by highly methylated CpG-depleted sequences. Interestingly, the most upstream region of this contains a sequence which is unmethylated in expressing tissues, as found for many (non-HTF) genes (11, 66). We therefore have the exceptional situation of an HTF flanked by an upstream region that shows the usual inverse correlation of hypermethylation and gene expression.

Chromatin structure of the Thy-1 gene. As the next step in the structural characterization of the gene, we localized all of the DNase I-hypersensitive sites in the Thy-1 gene. These sites have been determined for many genes and are believed to result from the binding of nonhistone proteins to chromatin. Many of these have been shown to be associated with regulatory sequences that are required to form transcriptionally active chromatin structures (27, 53, 63). Nuclei were
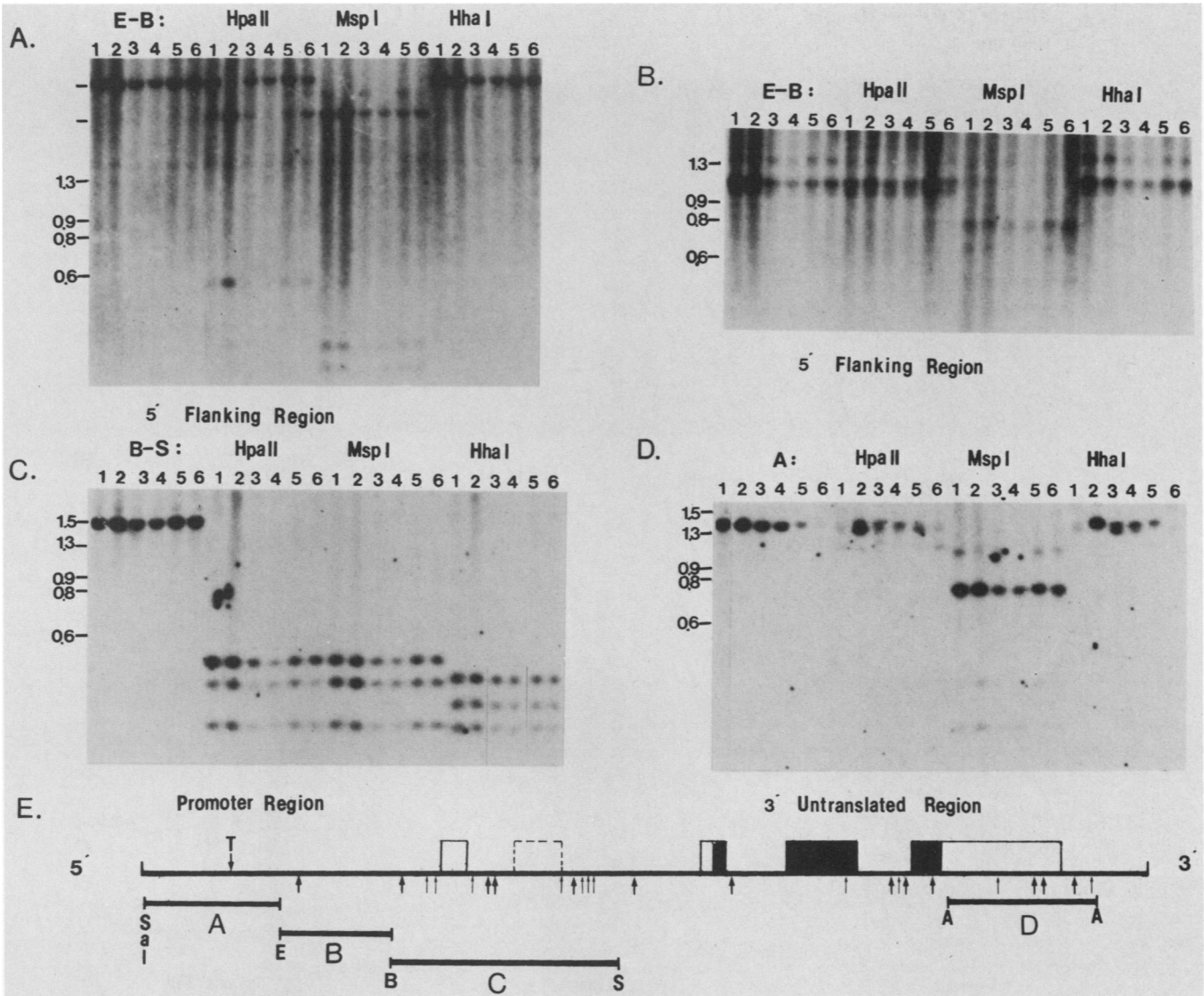
FIG. 2. Methylation pattern of the mouse Thy-1.2 gene. DNA from BALB/c mouse brain (lanes 1), thymus (lanes 2), sperm (lanes 3), liver (lanes 4), kidney (lanes 5) and muscle (lanes 6) was digested with EcoRI (E), BglII (B), BglII and SacI (S), or ApaI (A) and HpaII, MspI, or HhaI. After electrophoresis on a 1% agarose gel, DNA was transferred to nitrocellulose filters and hybridized with the 5'-flanking region probes EcoRI-SalI (A) (1,350 bp) or EcoRI-BglII (B) (1,140 bp). A 0.6-kb fragment detected in thymus after HpaII digestion indicates the presence of an unmethylated site 600 bp upstream of the EcoRI site in this tissue. Both probes A and B contain repetitive sequences which result in a background signal despite the presence of competitor DNA. Probe C, a BglII-SacI 1,450-bp fragment, spans the promoter area (panel C), whereas the ApaI-ApaI 1,410-bp fragment (probe D) hybridizes to the 3'-untranslated region (panel D). The probes and the endonuclease restriction sites relevant to this study are shown in panel E. Bold and thin arrows represent HpaII and HhaI sites, respectively.

prepared from brain, thymus, liver, and spleen and subjected to limited DNase I digestion. The DNA was isolated, digested with various enzymes, subjected to Southern blotting, and hybridized to several probes spanning the entire Thy-1 gene (Fig. 3). The region upstream of the EcoRI site which contained the tissue-specific methylation site did not show a DNase I-hypersensitive site (data not shown). The methylation-free region contained two DNase I-hypersensitive sites. Probe B (a 1-kb PstI fragment containing some repetitive sequences) hybridized to a 3.3-kb HindIII main band, clearly visible above the repetitive hybridization background, and two sub-bands with increasing DNase I digestion. One of these (arrow 1) was 1,310 bp and was only visible in brain and thymus and is therefore tissue specific.

The second band (790 bp, arrow 2), was very weak but it was present in all tissues (Fig. 3A). To confirm the position of these sites, EcoRI and PstI digests were also hybridized to probe A (a 520-bp PstI fragment), which placed them respectively 80 bp upstream of the major cap site and 215 bp downstream from the BstEII site in the first intron (Fig. 3B and E). The upstream DNase I-hypersenitive site has been previously reported but was mapped within exon Ia (51). This apparent discrepancy with our data is due to the fact that the previous mapping does not take into account a small insertion (24), which shifts the coordinates between the sequenced gene (Thy-1.2, BALB/c mice) and the mapped DNA (C57/BL mice). Rehybridization of the HindIII digests to probe C (a 1,410-bp ApaI fragment; Fig. 3E) revealed the
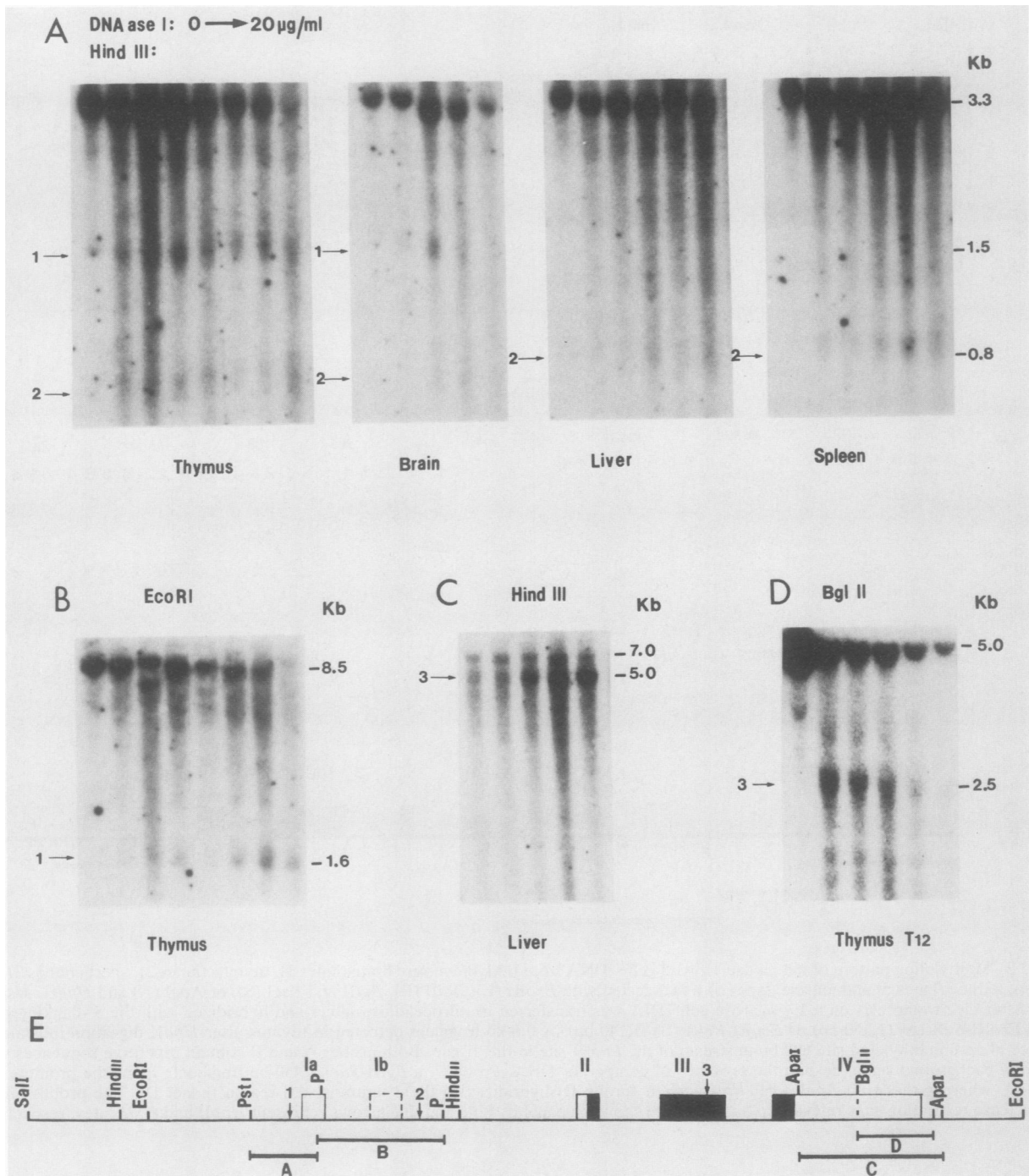
FIG. 3. Mapping of the hypersensitive sites in the Thy-1 gene. Nuclei were prepared as described elsewhere (40) from CBA × C57/BL mouse brain, thymus, liver, and spleen and digested with increasing amounts of DNase I (0, 0.4, 1.0, 1.5, 4.0, 8.0, 12.0, and 20.0 μg/ml from left to right in lanes) in a buffer containing 10 mM Tris (pH 7.4), 10 mM NaCl, 3 mM MgCl$_2$, 0.1 mM CaCl$_2$, and 1 mg of DNA per ml. Digestions were carried out for 6 to 15 min (depending on the tissue) at 37°C. The mixture was then digested with proteinase K for 2 h, extracted with phenol-chloroform three times, and ethanol precipitated. (A) After purification the DNA was redigested with HindIII, electrophoresed on a 1.2% agarose gel, transferred to nitrocellulose filters, and then hybridized with probe B (1,060-bp PstI-PstI fragment), which revealed a main band of 3.3 kb and two sub-bands of 1,310 bp (arrow 1) and 800 bp (arrow 2). (B) The purified DNA was digested with EcoRI and hybridized to a PstI-PstI 520-bp fragment covering the promoter region of the gene. A sub-band of 1.6 kb located the previously detected hypersensitive site (arrow 1) 80 bp upstream from the cap site. (C) The HindIII filters were rehybridized with a 3′ probe (probe C) (ApaI-ApaI 1,410-bp fragment), which revealed the presence of a third hypersensitive site (arrow 3) mapping inside exon 3. (D) Hypersensitive site mapped in transgenic mice carrying a mouse-human hybrid. Nuclei were subjected to the same DNase I conditions, digested with BglII, and probed with a human BglII-NcoI fragment of 950 bp (probe D). A 5.0-kb main band and a 2.5-kb sub-band localize the hypersensitive site (arrow 3) in the border between exon 3 and intron 3. (E) Probes used in this study are depicted as solid bars, and the hypersensitive sites are indicated by arrows. The BglII site that belongs to the human gene is indicated with a dotted line. (P, restriction site of PstI).
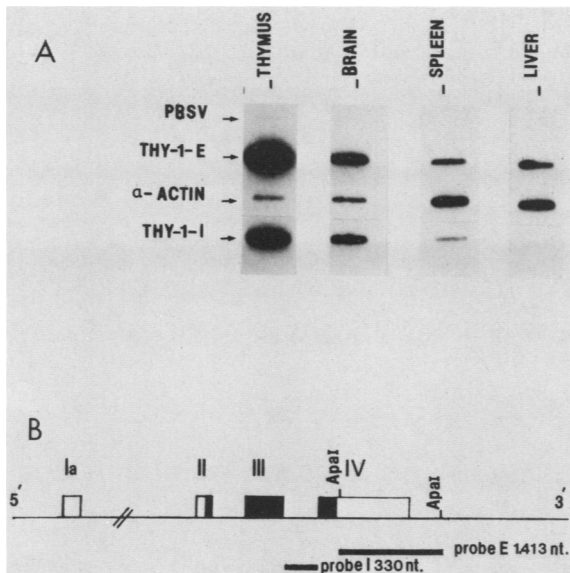
FIG. 4. Estimation of the transcriptional rates of the Thy-1 genes. (A) Nuclei prepared were from mouse brain, thymus, spleen, and liver as described in Materials and Methods. Each hybridization reaction was carried out under identical conditions with 6×10^5 cpm of ^{32}P-labeled nascent RNA in a 2-ml hybridization volume. Results were corrected for probe length and composition after various exposures of the autoradiography were scanned. (B) An ApaI-ApaI 1,410-bp fragment (Thy-1-E) covering the 3'-untranslated region of the gene and a SacI-PvuII 340-bp fragment (Thy-1-I) spanning the third intron were used as Thy-1 probes. An EcoRI-BamHI 8.2-kb fragment of the PBSV plasmid was used as a control for background signal, whereas the PstI-PstI cDNA fragment of mouse α-actin was used as a positive control in the assay. Each slot contained 10 μg of linearized, denatured probe.

presence of another non-tissue-specific, hypersensitive site located in or close to exon III (Fig. 3C and E). More precise mapping of this site was obtained by analysis of BglII digests of the DNA isolated from transgenic mice carrying a mouse-human hybrid gene (35). In this gene the 3' ApaI-EcoRI fragment has been replaced with the corresponding human sequence. The human Thy-1 probe D detected a 5-kb main band and a 2.5-kb sub-band, which positioned the hypersensitive site just upstream of the exon-intron border of exon III (Fig. 3D and E). Several other probes which covered the remainder of the gene failed to detect any hypersensitive sites.

Transcriptional unit of the mouse Thy-1 gene. The presence of the HTF and the absence of a TATA box sequence in the promoter suggested that the Thy-1 gene might be transcribed constitutively like a typical housekeeping gene (21). However, the presence of a tissue-specific hypersensitive site in the promoter indicated that the gene is regulated at the transcriptional level. We therefore carried out a run-on assay on isolated nuclei to measure the transcription rate in different tissues. ^{32}P-labeled nascent RNA from thymus, brain, spleen, and liver nuclei was hybridized to filters containing double-stranded Thy-1 and positive (actin) and negative (PBSV plasmid) control probes (Fig. 4). Hybridization to the intron III (Thy-1-I) probe by the thymus, brain, spleen, and (at very low levels) liver RNA clearly indicated that the gene was regulated at the transcriptional level in nerve and T cells. This result was confirmed by hybridization to an exon IV probe. The low level of hybridization

detected with liver RNA is presently unexplained. It could be due to the fact that the liver at this stage of development (fourth postnatal week) is still partly a hematopoietic tissue, or that there is some expression in hepatic cells. This low level of transcription did not result in detectable levels of mRNA as measured by Northern blots (data not shown), S1 nuclease protection analysis (Fig. 5B, lane 5), or primer extension (Fig. 5A, lane 2).

Since the gene is clearly regulated at the transcriptional level in nerve and T cells, we decided to determine whether the same transcriptional unit is used in both tissues. It has previously been shown (24) that the Thy-1 gene uses multiple transcription initiation sites in T cells, giving rise to exon Ia (Fig. 1). It was subsequently shown that a second promoter was also used (28), giving rise to exon Ib. These data suggested the possibility that there were two similarly used promoters present in the gene, although no comparative studies were presented. In addition, it was not clear whether the same transcription initiation sites were used in nerve cells. We examined the steady-state levels of mRNA by Northern blot analysis with sense and antisense RNA probes of exons Ia and Ib (Fig. 6). The antisense probes gave no hybridization signal (data not shown), whereas the exon Ia and Ib probes (Fig. 6B) detected mRNA in normal spleen, thymus, and brain of (C57BL × CBA) mice. RNA of a similar size was detected in the brain and thymus of transgenic mice (35) and the transgenic kidney cell line GRK (34). Normalization of the values for the specific activities of probes Ia and Ib showed that exon Ib is present in less than 5% of the total mRNA. Clearly, exon Ia is the major exon I, both in normal (Thy-1.2) and transgenic mice (Thy-1.2 and Thy-1.1).

Because of the surprising result with exon Ib and the apparent small difference in mobilities in the Northern blots, we confirmed these data with S1 nuclease protection assays. This would also allow detection of possible changes in the distribution of transcription initiation sites in different tissues that cannot be detected by Northern blot analysis. The 5' S1 nuclease probes from exons Ia and Ib (Fig. 5E) were hybridized to 20 μg of total RNA (the same as that used for the Northern blots) from either normal or transgenic tissues. The protection pattern of probe Ia by thymus RNA confirmed our previously reported data (24) that the major 5' ends of the mRNA form a triplet (around 30 bp; Fig. 5B, lanes 1 and 12) with a number of minor upstream starting sites. The same pattern (lanes 7 and 8) was found in transgenic thymus and transgenic kidney cells. Moreover, transfection of the Thy-1 promoter linked to the γ-globin gene and driven by the simian virus 40 enhancer into HeLa cells also showed prevalent usage of the same triplet of initiation sites (Fig. 5C). Surprisingly, the distribution of 5' ends was different in neuronal cells (either normal or transgenic) from that observed in thymus and transformed kidney cells. In addition to the characteristic triplet of initiation sites, the weak upstream starts detected in thymus represented major initiation sites in neuronal cells. Wild-type animals (Fig. 5B, lane 11) showed an increased usage of these upstream starts, whereas the mouse-human transgene showed an even more pronounced expression of these transcripts (Fig. 5B, lanes 6 and 14). To show that all S1 protected bands represent genuine 5' ends, we carried out a primer extension assay on the different tissues (Fig. 5A), using an 18-nucleotide primer (Fig. 5F); this confirmed the different usage of initiation sites in thymus (lanes 5, 6), transformed kidney cells (lane 9), and brain (lanes 7 and 8). Last, we carried out a 5'-end analysis with the exon Ib (Fig.
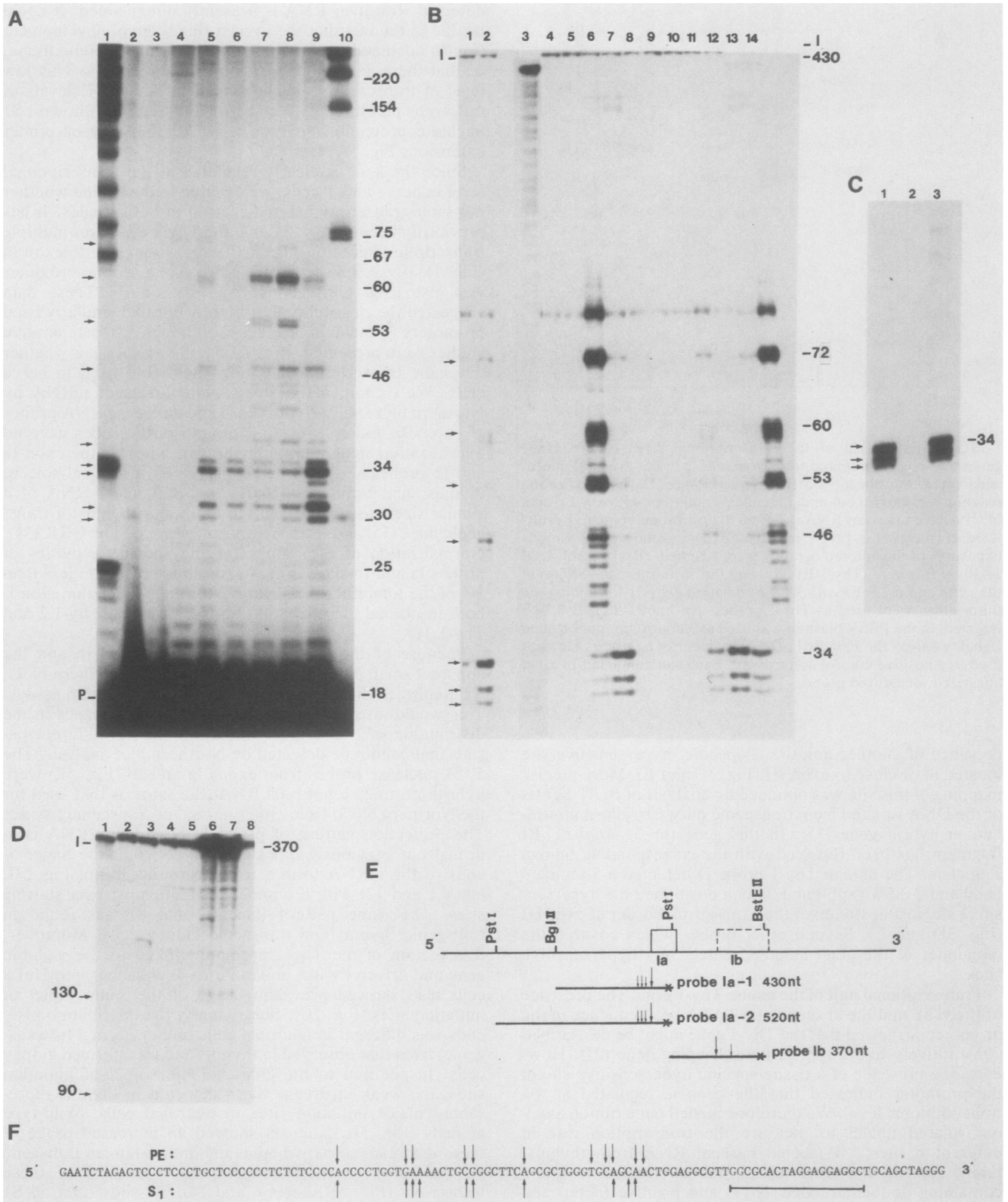
FIG. 5. 5'-end analysis of the Thy-1 transcripts. (A) Primer extension analysis. An 18-nucleotide primer spanning part of exon Ia (panel F) was annealed to 0.5 or 1 μg of poly(A)⁺ RNA from transgenic and wild-type mice, respectively. Extension products are indicated by arrows. Lanes: 1, marker; 2, liver; 3, kidney; 4, spleen; 5, transgenic thymus; 6, wild-type thymus; 7, wild-type brain; 8, transgenic brain; 9, GRK cells; 10, marker. The numbers represent the sizes of marker bands (220, 154, 75, 67, 34, and 25 kilobases) and elongation products (60, 53, 46, 34, and 30 kilobases). (B) Steady-state levels of Thy-1 mRNA measured by S1 nuclease protection. A 30-μg sample of total RNA was hybridized at 54°C with a BglII-PstI 430-bp probe (lanes 1 through 12) or a PstI-PstI 520-bp probe (lanes 13 and 14) (panel D). Both probes
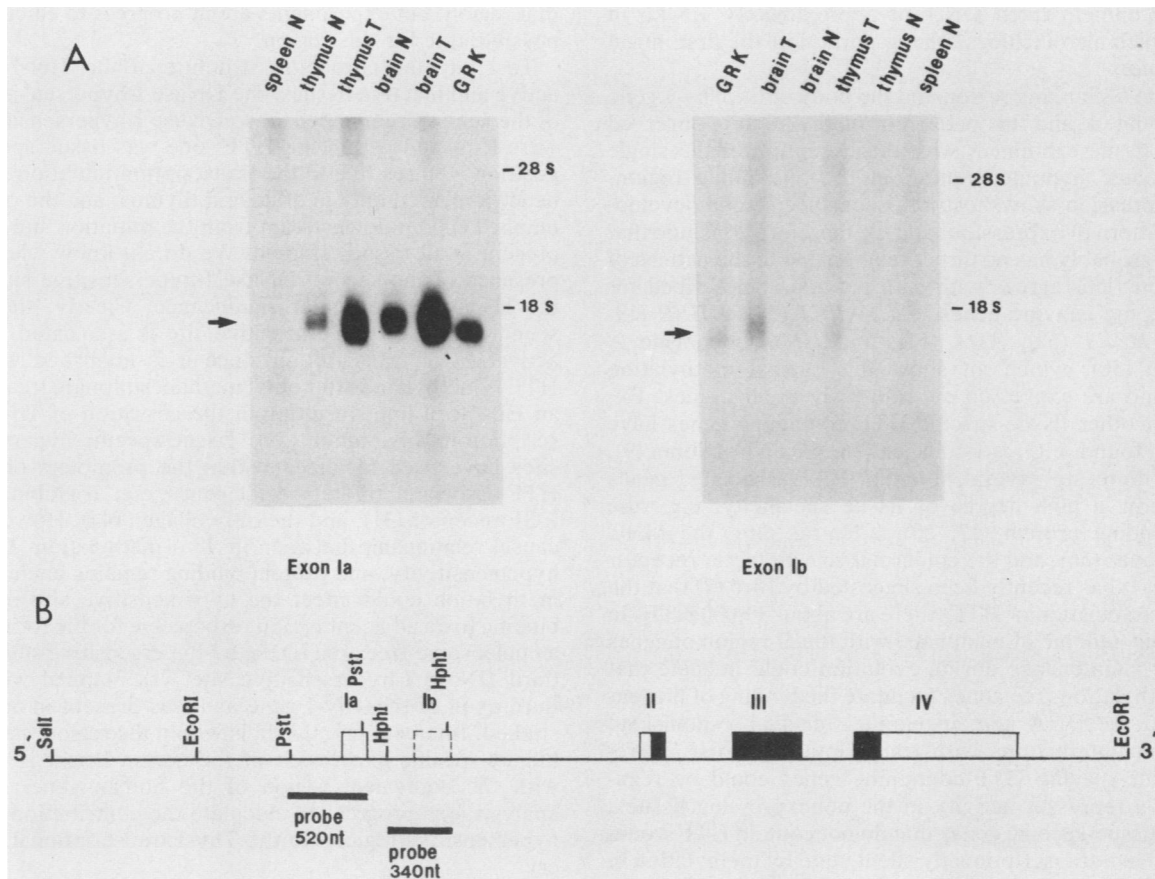
FIG. 6. Northern analysis. (A) RNA was extracted from wild type (N) and transgenic mice (T) tissues by the lithium chloride method (2), and poly(A)$^+$ RNA was prepared. A 1-µg sample of poly(A)$^+$ RNA from each tissue was electrophoresed on a 1.7% agarose–formaldehyde gel, transferred onto GeneScreen membrane, and hybridized with T7 polymerase-generated riboprobes at 52°C. Filters were washed in 3× SSC–0.1% sodium dodecyl sulfate for 1 h at 52°C and then in 0.3× SSC–1.0% sodium dodecyl sulfate at 65°C for 1 h and exposed to autoradiography. The mouse-human transgene hybrid has been previously shown to express five- to sixfold higher levels of Thy-1 mRNA, compared with the endogenous gene (35). GRK is a kidney cell line obtained from transgenic mice carrying a mouse-human-mouse Thy-1 hybrid (34), which codes for the human protein and results in proliferation of the tubular epithelia in kidney. Exon Ia indicates the Thy-1 mRNA detected by hybridization with a 520-nucleotide PstI-PstI probe that spans exon Ia. Exon Ib shows the same filter hybridized with an HphI-HphI 340-nucleotide fragment that detects exon Ib. (B) The probes used to detect both first exons are depicted. The specific activity of riboprobe Ib was 70% of that of probe Ia.

5D); this confirmed the Northern results, i.e., there was a small amount of transcript starting at the previously reported position in Ib. However, a number of other protected bands were also visible (lanes 3, 6, and 7), indicating that the small amount of Ib transcription is divided over a number of different start sites. All of these are less abundant than the minor start sites of Ia, and it is therefore extremely doubtful

whether the Ib mRNAs represent anything more than spurious transcripts (as observed in many other genes), which probably play no role in the biological function of Thy-1.

DISCUSSION

Structure of the mouse Thy-1 gene. The methylation analysis showed that the Thy-1 promoter is located within an

detect the transcription initiation sites for exon Ia. Lanes: 1, EL-4 cells (mouse T-cells); 2, GRK cells (34); 3, G+A track; 4, spleen; 5, liver; 6, brain; 7, thymus; 8, kidney (samples in lanes 1 through 8 were all from transgenic mice carrying a mouse-human hybrid (35)); 9, spleen; 10, kidney; 11, brain; 12, thymus (samples in lanes 9 through 12 were all from wild-type animals). Lanes 13 and 14 represent thymus and brain, respectively, from transgenic mice probed with the PstI-PstI 520-bp fragment, to probe for any other possible transcription initiation sites further upstream. Protected products are indicated by arrows. Numbers indicate the sizes of the major protection products. (C) S1 analysis in RNA from HeLa cells transfected with Thy-1-γ globin hybrid genes show that the same exon Ia triplet motif is found in HeLa cells as in thymus. Lanes: 1, wild-type Thy-1 in HeLa cells; 2, negative control; 3, deletion at position −270 (relative to the cap site) of the Thy-1 promoter. (D) S1 analysis of 30 µg of total RNA hybridized with a 370bp PstI-BstEII fragment detecting exon Ib. Lanes: 1, G+A track; 2, liver; 3, GRK cells; 4, kidney; 5, spleen; 6, brain; 7, thymus (samples in lanes 1 through 9 were all from transgenic mice carrying the mouse-human hybrid); 8, EL-4 cells. The protected input in lanes 6 and 7 is the result of precursor stability in these transgenic mice (as it was shown by S1 analysis; data not shown). Numbers indicate the sizes of the protected products in kilodaltons. (E) Probes used for S1 analysis. The probes were labeled at one end with T4 polynucleotide kinase. Arrows indicate the protected fragment. (F) Position of the transcription initiation sites of exon Ia as identified by primer extension and S1 analysis. The underlined sequence represents the region of exon Ia mRNA, which hybridizes to the primer.

island of unmethylated DNA of approximately 1.5 kb in length which also includes a large portion of the first intron of the gene.

Both the 5'-flanking region and the body of the Thy-1 gene are methylated, and this pattern of methylation is observed in every tissue examined, with the exception of the single tissue-specific hypomethylation in the 5'-flanking region. Thy-1 expression shows a strict tissue-specific and developmental pattern of expression, and we therefore conclude that the HTF probably has no direct relationship to the pattern of expression. This agrees with data previously described for housekeeping enzyme genes, e.g., APRT (54), DHFR (19, 50, 54), HPRT (62), ADA (57), and HMG coenzyme A reductase (50), which all show this typical methylation pattern and are expressed constitutively in all tissues. Examples of other tissue-specific HTF-containing genes have also been found, e.g., α_2I collagen (16, 43) and α-globin (9), whereas there are several potential HTF-associated genes which show a high degree of tissue specificity, e.g., the retinol-binding protein (17, 20), h-Ha-ras (30), the MHC class II genes (56), and the epidermal growth factor receptor gene (31). It has recently been suggested by Bird (7) that the frequent association of HTF (there are about 30,000 HTFs in the haploid genome of mammals) with the 5' region of genes and their maintenance during evolution could indicate that these methylation-free zones facilitate the binding of nuclear factors to the DNA and, in parallel, disqualify nonisland DNA from interactions with transcription factors. Therefore, tissue-specific HTF-containing genes could be regulated via a repressor activity in the nonexpressing tissues, whereas tissue-specific genes that do not contain HTF would be locked in a transcriptionally silent state by methylation in those tissues. A few examples have been reported that support this hypothesis. First, deletion analysis of the gene encoding retinol-binding protein indicated the presence of a repressor that was bound to the promoter region on negative tissues (17). Second, it has been shown that methylation of promoter regions results in the cessation of transcription in cell lines that do express the transfected nonmethylated genes (11, 46, 66). More recently, in vivo footprinting analysis of the TAT gene promoter allowed the identification of two binding sites containing HhaI sites. Footprints of methylated fragments containing these two sites showed that one of the two binding sites was completely abolished (3). A similar result was obtained for the E1A-inducible factor E2F, since methylation of its binding site inhibited binding of the factor, as was shown by dimethyl sulfate footprinting. Certainly, Thy-1 transcription is abolished, when at least part of the HTF is methylated with ethyl methanesulfonate in the T-cell line EL-4. Northern analysis proved the absence of any Thy-1 transcripts, but treatment with 5-azacytidine resulted in the expression of the Thy-1 gene (51); unfortunately, Sneller and Gunter did not present any data about the flanking regions, in particular the upstream tissue-specific hypomethylated site, and therefore this result is not yet unequivocal. Nevertheless, it indicates that the same rule applies to HTF genes as to other genes, i.e., hypomethylation is a necessary but insufficient condition to allow transcription. This is confirmed by the analysis of a nonmethylated Thy-1 construct introduced into transgenic mice, which showed complete restoration of HTF (37). It is, of course, not clear from our data whether the single unmethylated site in the otherwise methylated 5'-flanking region in T cells is significant. DNase I mapping of this region did not detect any hypersensitive sites as a measure of function in

that region, but experiments are in progress to elucidate any possible role for this region.

To study the chromatin structure of the Thy-1 gene in active and inactive tissues, the DNase I-hypersensitive sites of the gene were mapped. Two DNase I-hypersensitive sites were localized within the HTF; one was tissue specific (at position −80 relative to the transcription initiation site) and hence appeared only in brain and thymus, and the other one mapped 600 bp downstream from the initiation site and was present in all tissues studied. We do not know whether the presence of the second DNase I-hypersensitive site within the HTF is a particular significance. Clearly, the tissue-specific DNase I-hypersensitive site is associated with the activation of transcription since it is abolished when the HTF is methylated after ethyl methanesulfonate treatment of an EL-4 cell line, resulting in the cessation of Thy-1 transcription in this mutant (51). Tissue-specific hypersensitive sites have been localized within the promoters of several HTF-associated tissue-specific genes, e.g., α-globin (64), the EGF receptor (31), and the α2I collagen (44). However, the causal relationship between in vivo methylation, DNAse I hypersensitivity, and protein binding remains unclear, since methylation could affect the hypersensitive site per se or binding to an adjacent region responsible for the formation of a nucleosome-free area (DNase I-hypersensitive site) (32). A third DNase I-hypersensitive site was mapped within the third exon of the Thy-1 gene and was present in all tissues studied. Interestingly, this falls within the region that causes kidney-specific expression of the gene when it is replaced with the equivalent region of the human gene. Deletion analysis is in progress to elucidate the contribution of these hypersensitive regions to the Thy-1 transcriptional machinery.

The measurements of the transcriptional rates of the Thy-1 gene show that the gene is not transcribed in nonexpressing tissues and constitute the first evidence for a tissue-specific regulation of a housekeeping promoter that lack a TATA box at the level of transcription. Although many tissue-specific genes are included in the list of housekeeping promoters, there are no data at present to distinguish between a transcriptional regulation or a tissue-specific stabilization of the gene products.

The absence of a TATA box in the Thy-1 promoter produces the same effect as that observed in housekeeping genes, i.e., heterogeneity at the 5' ends of the mRNAs (42). In most cases, this has been correlated to the presence of a number of GC boxes (e.g., Sp1-binding sites) in the corresponding promoters. Tissue-specifically expressed housekeeping promoter genes (encoding epidermal growth factor receptor, h-Ha-ras, etc.) also follow the typical pattern of 5' heterogeneity and the presence of GC boxes (30, 31, 48). The only exception to date is the insulin receptor promoter, which is highly GC rich (methylated or not), contains several GC boxes, and lacks TATA and CAAT boxes but utilizes a single transcription initiation site (1). Although at present we do not understand what determines the site of initiation of transcription in the absence of a TATA box (10), the observation that T and nerve cells use a different distribution strongly suggests that it is a specific and regulated process. Clearly, this process is faithfully followed when the gene is introduced into mice. Interestingly, a fragment of the Thy-1 promoter at −270 bp (relative to the initiation site) initiates transcription from the same sites seen on thymus and kidney when introduced into HeLa cells. Although a number of genes are known to use either two different promoters with a distinguished tissue specificity (e.g., α-amylase [49] and

the yeast invertase gene [13]) or alternative first exon (e.g., c-*abl* [5]), there is only one example in which a developmental regulation of transcription initiation sites has been proposed (5C actin gene in *Drosophila* [59]).

To our knowledge, no other examples of tissue-specific utilization of transcription initiation sites which belong to the same exon have been reported for mammals. This diversity of 5' ends in the Thy-1 mRNA could merely be the result of a different initiation mechanism between the two tissues, without any further significance. On the other hand, in several cases it has been reported that the 5' untranslated region might be involved in translational efficiency (26) or mRNA instability (4).

Last, based on the S1 nuclease experiments and the Northern blots, we conclude that exon Ib could only play a minor (if any) role in the transcription of the Thy-1. In conclusion, the Thy-1 gene is a single strong promoter gene which offers an excellent example for the study of tissue-specific regulation of the position of transcription initiation rather than a gene that uses alternate promoter regions.

LITERATURE CITED

1. Araki, E., F. Shimada, H. Uzawa, M. Mori, and V. Ebina. 1987. Characterization of the promoter region of the human insulin receptor gene. J. Biol. Chem. 262:16186–16191.
2. Auffray, C., and O. Rougeon. 1979. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur. J. Biochem. 107:303–307.
3. Becker, P. B., S. Ruppert, and G. Schütz. 1987. Genomic footprinting reveals cell type specific DNA binding of ubiquitous factors. Cell 51:435–443.
4. Belasco, J. G., G. Nilsson, A. von Gabain, and S. N. Cohen. 1986. The stability of *E. coli* gene transcripts is dependent on determinants localized to specific mRNA segments. Cell 46:245–251.
5. Ben-Neriah, Y., A. Bernards, M. Paskind, G. Daley, and D. Baltimore. 1986. Alternative 5' exons in c-abl mRNA. Cell 44:577–586.
6. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 enconuclease digested hybrids. Cell 12:721–732.
7. Bird, A. P. 1986. CpG rich-islands and the function of DNA methylation. Nature (London) 321:209–213.
8. Bird, A. P. 1987. CpG rich islands as gene markers in the vertebrate nucleus. Trends Genet. 3:342–347.
9. Bird, A. P., M. H. Taggart, R. D. Nicholls, and D. R. Higgs. 1987. Non-methylated CpG-rich islands at the human α-globin locus: implications for evolution of the α-globin pseudogene. EMBO J. 6:999–1004.
10. Breathnach, R., and P. Chambon. 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. 50:349–383.
11. Busslinger, M., J. Hurst, and R. A. Flavell. 1983. Methylation and the regulation of globin gene expression. Cell 34:197–206.
12. Campbell, D. G., J. Gagnon, K. B. Reid, and A. F. Williams. 1981. Rat brain Thy-1 glycoprotein. The amino acid sequence, disulphide bonds and unusual hydrophobic region. Biochem. J. 195:15–30.
13. Carlson, M., and D. Botstein. 1982. Two differentially regulated mRNAs with different 5' ends encode secreted and intra cellular forms of yeast invertase. Cell 28:145–154.
14. Chang, H.-C., T. Seki, T. Moriuchi, and J. Silver. 1985. Isolation and characterization of mouse Thy-1 genomic clones. Proc.

Natl. Acad. Sci. USA 82:3819–3823.
15. Chen, S., F. Botteri, H. van der Putten, C. R. Landel, and G. Evans. 1987. A lymphoproliferative abnormality associated with inappropriate expression of the Thy-1 antigen in transgenic mice. Cell 51:7–19.
16. Chu, M.-L., W. de Wet, M. Bernard, and F. Ramirez. 1985. Fine structural analysis of the human Pro-α1(I) collagen gene. J. Biol. Chem. 260:2315–2320.
17. Colantuoni, V., A. Pirozzi, C. Blance, and R. Cortese. 1987. Negative control of liver-specific gene expression: cloned human retinol-binding protein gene is repressed in HeLa cells. EMBO J. 6:631–636.
18. Cross. G. 1987. Eukaryotic protein modification and membrane attachment via phosphatidylinositol. Cell 48:179–181.
19. Crouse, G. F., C. C. Simonsen, R. N. McEwan, and R. T. Schimke. 1982. Structure of amplified normal and variant dihydrofolate reductase genes in mouse sarcoma S180 cells. J. Biol. Chem. 257:7887–7897.
20. d'Onofrio, C., V. Colantuoni, and R. Cortese. 1985. Structure and cell-specific expression of a cloned human retinol binding protein gene: the 5' flanking region contains hepatoma specific transcriptional signals. EMBO J. 4:1981–1989.
21. Dynan, S. W. 1986. Promoters for housekeeping genes. Trends Genet. 6:196–197.
22. Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6–13.
23. Gardiner-Garden, M., and M. Frommer. 1987. CpG islands in vertebrate genomes. J. Mol. Biol. 196:261–282.
24. Giguere, V., K.-I. Isobe and F. G. Grosveld. 1985. Structure of the murine Thy-1 gene. EMBO J. 4:2017–2024.
25. Gunter, K. C., R. N. Germain, R. A. Kroczek, T. Saito, W. M. Yokoyama, C. Chan, A. Weiss, and E. M. Shevach. 1987. Thy-1 mediated T-cell activation requires co-expression of CD3/Ti complex. Nature (London) 326:505–507.
26. Hultmark, D., R. Klemenz, and W. J. Gehring. 1986. Translational and transcriptional control elements in the untranslated leader of the heat-shock gene hsp22. Cell 44:429–438.
27. Hutchison, N., and H. Weintraub. 1985. Localization of DNaseI-sensitive sequences to specific regions of interphase nuclei. Cell 43:471–482.
28. Ingraham, H. A., and G. A. Evans. 1986. Characterization of two atypical promoters and alternate mRNA processing in the mouse Thy-1.2 glycoprotein gene. Mol. Cel. Biol. 6:2923–2931.
29. Ingraham, H. A., G. M. Lawless, and G. A. Evans. 1986. The mouse Thy-1.2 glycoprotein gene: complete sequence and identification of an unusual promoter. J. Immunol. 136:1482–1489.
30. Ishii, S., G. T. Merlino, and I. Pastan. 1985. Promoter region of the human Harvey ras proto-oncogene: similarity to the EGF receptor proto-oncogene promoter. Science 230:1378–1381.
31. Ishii, S., Y.-H. Xu, P. H. Stratton, B. A. Roe, G. T. Merlino, and I. Pastan. 1985. Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. Proc. Natl. Acad. Sci. USA 82:4920–4924.
32. Keshet, I., J. Lieman-Hurwitz, and H. Cedar. 1986. DNA methylation affects the formation of active chromatin. Cell 44:535–543.
33. Kioussis, D., E. Vanin, T. deLange, R. A. Flavell, and F. G. Grosveld. 1983. β-Globin gene inactivation by DNA translocation in γ-β thalassaemia. Nature (London) 306:662–666.
34. Kollias, G., D. J. Evans, M. Ritter, J. Beech, R. Morris, and F. Grosveld. 1987. Ectopic expression of Thy-1 in the kidneys of transgenic mice induces functional and proliferative abnormalities. Cell 51:21–31.
35. Kollias, G., E. Spanopoulou, F. Grosveld, M. Ritter, J. Beech, and R. Morris. 1987. Differential regulation of a Thy-1 gene in transgenic mice. Proc. Natl. Acad. Sci. USA 84:1492–1496.
36. Kollias, G., N. Wrighton, J. Hurst, and F. Grosveld. 1986. Regulated expression of human ^Aγ-β- and hybrid γβ-globin genes in transgenic mice: manipulation of the developmental expression patterns. Cell 46:89–94.
37. Kolsto, A.-B., G. Kollias, V. Giguere, K.-I. Isobe, H. Prydz, and F. Grosveld. 1986. The maintenance of methylation-free islands

in transgenic mice. Nucleic Acids Res. **14:**9667–9678.

38. **Kroczek, R. A., K. C. Gunter, R. C. Germain, and E. M. Shevach.** 1986. Thy-1 functions as a signal transduction molecule in T lymphocytes and transfected B lymphocytes. Nature (London) **322:**181–184.

39. **Krumlauf, R., P. W. H. Holland, J. McVey, and B. L. M. Hogan.** 1987. Developmental and spatial patterns of expression of the mouse homeobox gene, Hox 2.1. Development **99:**603–617.

40. **Linial, M., N. Gunderson, and M. Groudine.** 1985. Enhanced transcription of c-myc in bursal lymphoma cells requires continuous protein synthesis. Science **230:**1126–1132.

41. **Low, M. G., and P. W. Kincade.** 1985. Phosphatidyl inositol in the membrane-anchoring domain of the Thy-1 glycoprotein. Nature (London) **318:**62–64.

42. **McGrogan, M., C. C. Simonsen, D. T. Smouse, P. J. Farnham, and R. T. Schimke.** 1985. Heterogeneity at the 5' termini of mouse dihydrofolate reductase mRNAs. J. Biol. Chem. **260:**2307–2314.

43. **McKeon, C., H. Ohkubo, I. Pastan, and B. deCrombrugghe.** 1982. Unusual methylation pattern of the α₂I collagen gene. Cell **29:**203–210.

44. **McKeon, C., I. Pastan, and B. de Crombrugghe.** 1984. DNaseI sensitivity of the α₂(I) collagen gene: correlation with its expression but not with its methylation pattern. Nucleic Acids Res. **12:**3491–3502.

45. **Morris, R.** 1985. Thy-1 in developing nerve tissue. Dev. Neurosci. **7:**133–160.

46. **Murray, E., and F. Grosveld.** 1987. Site specific demethylation in the promoter of human γ-globin gene does not alleviate methylation mediated suppression. EMBO J. **6:**2329–2335.

47. **Reif, A., and J. Allen.** 1964. The AKR thymic antigen and its distribution in leukemias and nervous tissues. J. Exp. Med. **120:**413.

48. **Reynolds, G. A., S. K. Basu, T. F. Osborne, D. J. Clein, G. Gel, M. S. Brown, J. L. Goldstein, and K. L. Luskey.** 1984. HMG-CoA reductase: a negatively regulated gene with unusual promoter and 5' untranslated region. Cell **38:**275–285.

49. **Saluz, H. P., J. Jiricny, and J. P. Jost.** 1986. Genomic sequencing reveals a positive correlation between the kinetics of strand-specific DNA demethylation of the overlapping estradiol/glucorticoid-receptor binding sites and the rate of the avian vitellogenin mRNA synthesis. Proc. Natl. Acad. Sci. USA **83:**7167–7171.

50. **Schibler, U., O. Hagenbüchle, P. K. Wellauer, and A. C. Pittet.** 1983. Two promoters of different strengths control the transcription of the mouse alpha-amylase gene Amy-1ᵃ in the parotid gland and the liver. Cell **33:**501–508.

51. **Sneller, M., and K. Gunter.** 1987. DNA methylation alters chromatin structure and regulates Thy-1 expression in EL-4 T cells. J. Immunol. **138:**3505–3512.

52. **Southern, E. M.** 1975. Detection of specific sequences among

53. **Stalder, J., A. Larsen, J. D. Engel, M. Dolan, M. Groudine, and H. Weintraub.** 1980. Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNaseI. Cell **20:**451–460.

54. **Stein, R., N. Sciaky-Gallili, A. Razin, and H. Cedar.** 1983. Pattern of methylation of two genes coding for housekeeping functions. Proc. Natl. Acad. Sci. USA **80:**2422–2426.

55. **Tse, A. G. C., A. N. Barclay, A. Watts, and A. F. Williams.** 1985. A glycophospholipid tail at the carboxyl terminus of the Thy-1 glycoprotein of neurons and thymocytes. Science **230:**1003–1008.

56. **Tykocinski, L. M., and E. E. Max.** 1984. CG dinucleotide clusters in MHC genes and in 5' demethylated genes. Nucleic Acids Res. **12:**4385–4396.

57. **Valerio, D., M. G. C. Duyvesteyn, B. M. M. Decker, G. Weeda, Th. M. Berkvens, L. van der Voorn, H. van Ormond, and A. J. van der Eb.** 1985. Adenosine deaminase: characterization and expression of a gene with a remarkable promoter. EMBO J. **4:**437–443.

58. **Van der Ploeg, L. H. T., and R. A. Flavell.** 1980. DNA methylation in the human γδβ-globin locus in erythroid and non-erythroid tissues. Cell **19:**947–958.

59. **Vigoreaux, J. O., and S. L. Tobin.** 1987. Stage-specific selection of alternative transcriptional initiation sites from the 5C actin gene of Drosophila melanogaster. Genes Dev. **1:**1161–1171.

60. **Weaver, R. F., and C. Weissman.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure. The 5' termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. **6:**1175–1192.

61. **Williams, A., and J. Gagnon.** 1982. Neuronal cell Thy-1 glycoprotein: homology with immunoglobulin. Science **216:**696–703.

62. **Wolf, S. F., D. J. Jolly, K. D. Lunmen, T. Freidmand, and B. R. Migeon.** 1984. Methylation of the hypoxanthine phosphoribosyl-transferase locus on the human X-chromosome: implications for X-chromosome inactivation. Proc. Natl. Acad. Sci. USA **81:**2806–2810.

63. **Wu, C.** 1987. The 5' ends of Drosophila heat shock gene in chromatin are hypersensitive to DNaseI. Nature (London) **286:**854.

64. **Yagi, M., R. Gelinas, J. T. Elder, M. Peretz, T. Papayannopoulou, G. Stamatoyannopouplos, and M. Groudine.** 1986. Chromatin structure and developmental expression of the human α-globin cluster. Mol. Cell. Biol. **6:**1108–1116.

65. **Yee, A. S., R. Reichel, I. Kovesdi, and J. R. Nevins.** 1987. Promoter interaction of the E1A-inducible factor E2F and its potential role in the formation of a multi-component complex. EMBO J. **6:**2061–2068.

66. **Yisraeli, J., R. S. Adelstein, D. Melloul, V. Nudel, D. Yaffe, and H. Cedar.** 1986. Muscle-specific activation of a methylated chimeric actin gene. Cell **46:**409–416.

DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.