MOLECULAR AND CELLULAR BIOLOGY, Apr. 1991, p. 2216–2228
0270-7306/91/042216-13$02.00/0

Vol. 11, No. 4

The Functional Domains of the Murine *Thy-1* Gene Promoter

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

Received 18 September 1990/Accepted 21 December 1990

The *Thy-1* gene promoter resembles a "housekeeping" promoter in that it is located within a methylation-free island, lacks a canonical TATA box, and displays heterogeneity in the 5′-end termini of the mRNA. Using transgenic mice, we show that this promoter does not confer any tissue specificity and is active only in a position-dependent manner. It can only be activated in a tissue-specific manner by elements that lie downstream of the initiation site. We have analyzed the functional domains of the minimal *Thy-1* promoter and show that the dominant promoter elements consist of multiple binding sites for the transcription factor Sp1, an inverted CCAAT box, and sequences proximal to the transcription start site. DNase I and gel mobility shift assays show the binding of a number of nuclear factors to these elements, including Sp1 and CP1. Our results show that the structure of this promoter only permits productive interactions of the two transcription factors Sp1 and CP1 with the basal transcription machinery in the presence of enhancer sequences.

The murine Thy-1 antigen, a member of the immunoglobulin superfamily, is a cell surface glycoprotein anchored on the membrane through a phospholipid tail (47, 66, 70). In the mouse, Thy-1 is expressed at high levels in the neuronal and lymphoid systems (55, 57, 70), being found on the surface of all neurons, in which it represents a terminal differentiation marker (54). In the lymphoid system, Thy-1 is present at high levels on most cortical thymocytes and mature T lymphocytes. This expression pattern is the result of both tissue-specific and developmental regulation exerted at the transcriptional level (44, 62).

The promoter is located within a methylation-free, CpG-rich island (HTF island) (4) and lacks a canonical TATA box (23), features which are found often in "housekeeping" gene promoters (18) but less frequently in tissue-specific genes (19). Initiation of transcription occurs from multiple sites the distribution of which differs in murine brain and thymus (62). In apparent contrast to the absence of tissue specificity and its housekeeping characteristics, the *Thy-1* gene promoter is nevertheless characterized by the presence of a tissue-specific DNase I-hypersensitive site, located 80 bp upstream of the start site (62).

By using transgenic mice, a number of *Thy-1* transcriptional elements have been identified. Expression of the gene in early thymocytes is controlled by enhancer sequences located within the murine third intron (67). However, this thymocyte-specific enhancer is not sufficient to maintain expression of the gene in mature T cells (44), which requires the presence of distinct regulatory sequences (67). Transcriptional activation of the *Thy-1* gene in the majority of neuronal cells is dictated by *cis*-acting sequences present at the 3′ end of the first intron (7, 67). Thus, tissue-specific and developmental transcriptional activation of the *Thy-1* promoter is achieved through separate elements which function independently of each other and are located downstream of the principal site of initiation. Experiments in transgenic mice have also established that a 300-bp region of the *Thy-1* promoter (−270 to +36) contains all of the elements required to cooperate with the downstream tissue-specific elements (67). Moreover, hybrid transgenes in which the *Thy-1* promoter was replaced by different heterologous promoters exhibited a tissue-specific and developmental profile similar to that of the endogenous *Thy-1* gene. These data suggest that the presence of the tissue-specific hypersensitive site in the *Thy-1* promoter is presumably the result of productive interactions between the promoter and distant regulatory factors, with the promoter elements acting solely as the mediators rather than the initiators of activation. This scenario constitutes a common motif underlying the regulation of a number of different genes (56). Alternatively, the promoter may have some tissue-specific components to allow expression.

In this study, we define the specificity of the *Thy-1* promoter by using transgenic mice, and we describe the identification of the functional elements required for activity of the *Thy-1* promoter by mutational analysis and the characterization of the nuclear factors that bind to these elements.

MATERIALS AND METHODS

Preparation of the hybrid construct and analysis of transgenic mice. The hybrid construct, microinjected into fertilized mouse oocytes, was prepared by replacing the promoter of the human γ-globin gene with a 310-bp *Sma*I-*Pst*I fragment of the *Thy-1* promoter corresponding to sequences −270 to +36 of the promoter. *Bgl*II and *Hin*dIII linkers were inserted in the *Sma*I and *Pst*I sites, respectively, and the *Bgl*II-*Hin*dIII fragment of the *Thy-1* promoter was fused to the *Nco*I site present in the first exon of the human γ-globin gene by means of a *Hin*dIII linker inserted in this site. DNA fragments were isolated free of vector sequences from a low-melting-point agarose gel (Sea-Plaque) after phenol extraction. Fertilized oocytes from (CBA × C57BL/10)F_1 mice were isolated, and pronuclei were injected as described previously (44). Transgenic mice were identified by Southern blot analysis of 10 μg of genomic DNA. Total mRNA from mouse tissues was isolated and analyzed by S1 nuclease protection as described previously (62).

Construction of promoter mutants. The linker-scanning (LS) and internal deletion (ID) mutations were engineered by

* Corresponding author.
† Present address: Division of Endocrinology, Hospital for Sick Children, Toronto, Canada N5G 1X8.

joining 5′ and 3′ deletions covering the promoter region (49). Construction of the plasmid (pGSE628) containing 520 bp of the mouse *Thy-1* 5′ region fused to the chloramphenicol acetyltransferase (CAT) gene (by means of a *Hin*dIII linker inserted in the *Pst*I site of the *Thy-1* first exon) has been described elsewhere (23). Plasmid pGSE668 is a derivative of pGSE628 in which part of the *Thy-1* promoter region was deleted by excising a restriction fragment delineated by the *Cla*I site in the vector and the *Sma*I site in the *Thy-1* promoter (−270 bp relative to the cap site). This plasmid was recircularized after the addition of a *Bgl*II linker. Plasmids pGSE628 and pGSE668 were linearized by digestion with *Sma*I and *Hin*dIII, respectively, and then digested with exonuclease BAl 31. DNA samples with deletions extending to the desired positions in the promoter were then introduced into plasmids by standard procedures. The recombinants were screened for the presence of the *Cla*I linker, and the selected clones were sequenced by the Maxam-Gilbert method (48). Matching 5′ and 3′ deletions were made by exchange of the corresponding *Cla*I and *Hin*dIII restriction fragments between the two sets of plasmid. All of the resulting LS or ID mutants contain 277 bp of the *Thy-1* promoter with an introduced *Cla*I linker between positions −120 and +8 relative to the cap site. To express the mutants transiently in HeLa cells and analyze them by S1 nuclease protection, the CAT gene was replaced by an *Nco*I-*Hin*dIII γ-globin fragment (see Fig. 3b) which contains the globin poly(A) site, while pBR322 was replaced a pBSV vector (pBR328 *Bam*HI-*Eco*RI fragment linked to a simian virus 40 [SV40] *Bam*HI-*Eco*RI fragment).

DNA-mediated transfections, CAT assays, and S1 nuclease protection analysis. Transfections into L cells and CAT assays were performed as described previously (23). Transfections into HeLa cells were performed by calcium phosphate precipitation. Total mRNA was isolated and analyzed by S1 nuclease protection as described elsewhere (62). The 5′-end S1 probes were a *Bgl*II-*Bgl*I 360-bp or an *Xba*I-*Bgl*I 145-bp fragment labeled at the *Bgl*I site. Both probes protect 81- to 86-bp fragments from the hybrid transcripts. In all cases, values were normalized for transfection efficiency by comparison with a cotransfected human β-globin plasmid. A 540-bp *Acc*I-*Acc*I 5′-end probe was used to detect the human β-globin transcripts.

Preparation of nuclear extracts and DNase I footprinting. All cells were grown in suspension to a density 0.6×10^6 to 1.0×10^6 cells per ml. Conditions for the preparation of nuclear extracts were identical to those described by Dignam et al. (16) for HeLa cells and by Gorski et al. (24) for whole mouse brain, EL-4, and K562 extracts. For DNase I footprinting assays, a *Sma*I-*Pst*I fragment representing the *Thy-1* promoter region −270 to +36 (relative to the cap site) was cloned into pGEM4 vector. The plasmid was linearized with either *Eco*RI or *Hin*dIII (polylinker sites), labeled with T4 polynucleotide kinase, recut with a second enzyme, and isolated from a preparative gel by phenol extraction. Binding reactions and analysis of the purified DNA were performed as described elsewhere (15).

Gel mobility shift assays. A chemically synthesized single-stranded oligonucleotide was ^{32}P labeled with T4 kinase; after separation of the unincorporated nucleotides, the labeled strand was annealed to a fourfold excess of the nonlabeled complementary strand. Binding reactions and analysis on a native acrylamide gel were performed as described previously (15). In competition assays, a 100- to 200-fold excess of cold probe was added.

Sequences of oligonucleotides used as probes in gel retardation assays. Sequences of probes were as follows: THY-A, 5′-CGATTACCAACCCCACCCCCTGGAT-3′; THY-B, 5′-CGATGGCCCCCCTCCATCCGTCAT-3′; THY-C, 5′-CGATGTCCATCCTTTTCCCTCGGAT-3′; THY-D, 5′-CGATGTCTCAGCCTCCGATTCGAT-3′; THY-E, 5′-CGATGCTGACCGATTGGCTGCACAT-3′; THY-F, 5′-CGATCGAGTCCCTCCCTGCTGATAT-3′; THY-G, 5′-CGATTGCTCCCCCCTCTCTCCCCACCCTAT-3′; THY-H, 5′-CGATGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACAT-3′; THY-K, 5′-CGATGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCCAT-3′; THY-M, 5′-CGATAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGAT-3′; Sp1 GC III,IV, 5′-CGATGGGCGGAGTTAGGGGCGGGACTAT-3′Sp1 GC III, 5′-GATCCGAACTGGGCGGAGTTAGGGGCA-3′; Sp1 mut., 5′-GATCCGAACTGGACTGAGTTAGGGGCA-3′; α-globin CCAAT; 5′-CTCCGCGCCAGCCAATGAGCGCCGCCC-3′; and adenovirus NF-I, 5′-TATACCTTATTTTGGATTGAAGCCAATATGATTGC-3′.

Methylation interference assays. The labeled probes were methylated as described by Maxam and Gilbert (48), and complexes were resolved on a native gel. Bands containing the appropriate complexes were excised from the gel and electroeluted. The nucleic acid was purified, and after NaOH treatment (48), the radioactive samples were separated on a 10% sequencing gel.

RESULTS

Analysis of promoter activity in transgenic mice. A complete *Thy-1* gene containing all of the downstream regulatory elements but only 310 bp of promoter sequences shows proper tissue- and developmental stage-specific expression in transgenic mice (67). To further analyze the specificity of these sequences, a fragment representing the promoter region between −270 and +36 was linked to the first exon of the γ-globin gene (Fig. 1B). The hybrid construct was microinjected into fertilized mouse oocytes, and six transgenic lines carrying different copy numbers of the transgene were generated (Fig. 1C). Total mRNA was isolated from a variety of adult transgenic tissues and assayed by S1 protection analysis for expression of the endogenous *Thy-1* gene and of the hybrid transgene (Fig. 1A). Transcripts of the endogenous gene were found at negligible levels in the kidney and liver, at low levels in muscle, spleen, and lung, and at high levels in brain and thymus. In contrast, transcripts from the transgene were present at very low levels (about 0.75% in relation to the endogenous levels) in the brain of three of six lines, and these did not show the distribution of transcription initiation typical for brain (62). The transgene was silent in the rest of the tissues with the exception of line Tr.4 (containing the highest number of copies of the transgene), which expressed the hybrid gene at very low levels in all tissues analyzed (Fig. 1A). These results confirm that the *Thy-1* promoter is an extremely weak or inactive unit which has no tissue specificity and functions only in the presence of enhancer sequences as previously suggested (23, 33, 67).

Mapping of the functional elements of the *Thy-1* promoter. To define the sequences required for promoter function, a series of deletion mutants with 5′-terminal endpoints located between −489 and −80 were constructed and linked onto the CAT gene in the presence of an SV40 enhancer (see Materials and Methods). The mutants were cotransfected with the β-globin control plasmid into L cells, and transcriptional activity was determined 48 h later by CAT assays and S1 nuclease protection analysis. No decrease in the relative

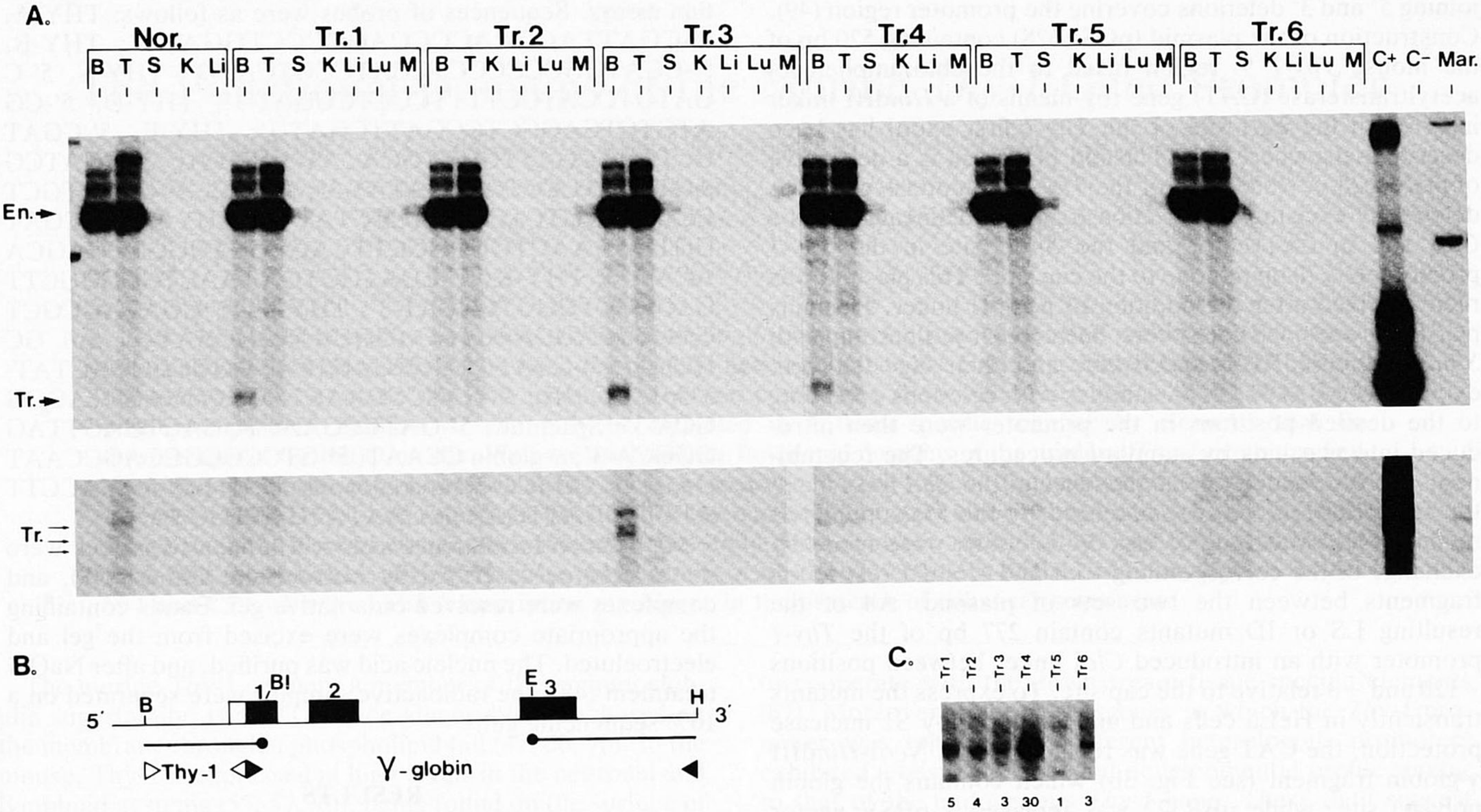

FIG. 1. Specificity of the *Thy-1* promoter as analyzed in transgenic mice. (A) A S*ma*I-*Pst*I 310-bp fragment of the *Thy-1* promoter (−270 to +36) was linked to the first exon of the γ-globin gene (see Materials and Methods). Six individual transgenic lines carrying various copy numbers of the transgene were generated. Total mRNA was isolated from a number of adult tissues and analyzed by S1 nuclease protection using three different probes (40 μg of total RNA per lane). Endogenous *Thy-1* transcripts were detected by a 680-bp S*ac*I-*Tth*111I probe that detects a 244-bp fragment (En.) corresponding to the 5′ border of exon IV. Transgenic transcripts were assayed by both a 5′-end *Bgl*II-*Bgl*I probe which protects 75- to 80- bp fragments representing multiple transcription initiation sites (triplet of thin arrows, Tr.; lower half of panel A) and a 3′-end *Eco*RI-*Hin*dIII 520-bp probe which detects a fragment of 170 bp corresponding to the polyadenylation site of the exogenous γ-globin gene (Tr.) and a minor polyadenylation site 180 bp further downstream (350-bp fragment). Note that the 5′ probe protects a nonspecific band in the normal brain lane. Lanes are designated as follows: brain (B), thymus (T), spleen (S), kidney (K), liver (Li), lung (Lu), and muscle (M). C+, positive control representing 40 μg of total mRNA isolated from a HeLa cell line, transiently transfected with the hybrid transgene linked to an SV40 enhancer; C−, negative control (40 μg of HeLa total mRNA); Mar., radiolabeled marker fragments. Tr.1 to Tr.6, Transgenic lines; Nor., a nontransgenic mouse control. (B) Schematic representation of the *Thy-1*/γ-globin gene hybrid construct, analyzed in transgenic mice. Probes used for S1 analysis are also indicated. Restriction sites: *Eco*RI (E), *Hin*dIII (H), *Bgl*II (B), *Bgl*I (BI). (C) Genomic DNA obtained from the transgenic lines was digested with *Sac*I, subjected to Southern blot analysis, and hybridized to a human γ-globin probe to estimate the copy number of the integrated transgenes. Transgenic lines are indicated above the lanes; estimated copy numbers of the hybrid construct are indicated below.

activity (RA; ratio of CAT activity to β-globin expression) was observed until deletions reached position −80 (data not shown and Fig. 2). Thus, it appears that the first 100 bp upstream of the major cap site are essential and sufficient for efficient initiation of transcription by the mouse *Thy-1* promoter.

To define the regulatory domains of the minimal *Thy-1* promoter, a number of mutants were produced (between −120 and +36) in which wild-type *Thy-1* sequences were either substituted or deleted (see Materials and Methods). To increase the steady-state mRNA levels of the hybrid transcripts, the various promoter mutants were linked to the γ-globin gene. Constructs were transiently expressed either in L cells (*Thy-1*/CAT gene hybrids) or in HeLa cells (*Thy-1*/γ-globin gene hybrids) along with the cotransfected β-globin control plasmid and assayed by CAT assays (Fig. 2) or S1 analysis (Fig. 2 and 3), respectively. The results show that sequences between −91 and −77, −68 and −48, and −24 and −10 are essential for transcriptional activation, since mutations in these three regions almost abolish transcription (Fig. 2 and 3). Mutations in other regions (−41 to −31 and −1 to +8) have a moderate effect on promoter activity. An unexpected result is the high RA of ID −52/−31, since LS −41/−31 exhibits only 40% of wild-type activity. This could be explained by the presence of negative sequences within the region from −52 to −41 or by the fact that a positive regulatory element(s) has been brought close to the proximal promoter. Analysis of representative mutants by S1 nuclease protection (Fig. 3a) shows the heterogeneity at the 5′ end of the transcripts and establishes the functional domains of the *Thy-1* promoter. Comparison of lanes WT (270 bp of promoter sequences) and ID-120 (120 bp of promoter sequences) indicates that sequences upstream of −101 are dispensable for promoter activity in both HeLa and L cells (Fig. 2). The region between −101 and −41 is essential for transcription efficiency, but mutations in this region do not appear to change the site of initiation. However, three LS mutations in the region from −30 to +8 show decreased expression but also the appearance of novel start sites. The importance of this region is also accentuated by internal deletions ID −41/−19, −18/+8, and −1/+36.

Three main regions were identified as essential for efficient

	-118 -110 -100 -90 -80 -70 -60 -50 -40 -30 -20 -10 CAP +10 +20 +30	RA
wild type	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	1.00
CD-116	gatgCCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.95
CD-113	AtCgatgCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.90
CD-101	----------CatcgatgCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.25
CD-80	----------CaTCgatgTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.05
LS-120/-113	cAtCgatgCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	1.15
LS-118/-109	ACCAt-CgAtgCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.95
LS-110/-101	ACCACCCCACCatCgatg-CCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.70
LS-97/-89	ACCACCCCACCCCCTCCATCCTTcatCgatgGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	1.25
LS-91/-80	ACCACCCCACCCCCTCCATCCTTTTCCCTCAtCgatg---TGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.05
LS-83/-77	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCcATcGaCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.05
LS-71/-64	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTcGAtgCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	1.10
LS-68/-59	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAcA-TCgaTgCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.20
LS-59/-48	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCaT-Cgatg--CTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.05
LS-41/-31	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTC--CAtCgaTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.40
LS-24/-19	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGaGgGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.30
LS-18/-10	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGcaTcgAtgGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.20
LS-1/+8	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCcatcgaTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.45
ID-120/-101	catc----------gatgCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	1.00
ID-97/-77	ACCACCCCACCCCCTCCATCCTTcATc----------gatgCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.05
ID-83/-64	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCcATc----------GAtgCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.05
ID-67/-48	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGcaTC----------gatgCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.20
ID-52/-31	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCatC----------gaTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	1.00
ID-41/-19	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCatc----------gatGGGCTTCAGCGCTGGGTGCAGCAACTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.05
ID-18/+8	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGcaTc----------gaTGGAGGCGTTGGCGCACTAGGAGGAGGCTGCAG	0.15
ID-1/+36	ACCACCCCACCCCCTCCATCCTTTTCCCTCAGCCTCCGATTGGCTGAATCTAGAGTCCCTCCCTGCTCCCCCCTCTCTCCCCACCCCTGGTGAAAACTGCGGGCTTCAGCGCTGGGTGCcat----------agcttG	0.30

FIG. 2. Functional analysis of the mouse *Thy-1* promoter mutants. The nucleotide sequence of the wild-type mouse *Thy-1.2* gene promoter is shown on the top line (23). The CD mutants are designated by the position of the endpoint of the 5′ deletion. The LS and ID mutants are designated by the position of the nucleotide delimiting the 5′ and 3′ endpoint deletions. The sequence of the *Thy-1.2* promoter changed by the *Cla*I linker is underlined for all mutants. Nucleotides altered by the insertion are indicated by small letters. Deletions are represented by bars, and changes in the spacing of the insertions are shown above the sequences. RA values represent the expression levels of each individual mutant relative to wild-type activity after normalizing for the cotransfected β-globin levels. The mutants were tested for expression efficiency in both HeLa and L-cells. RA data are mean values of two different experiments in L cells and at least four different experiments in HeLa cells. Closed and open triangles indicate strong and weak transcription initiation sites, respectively.

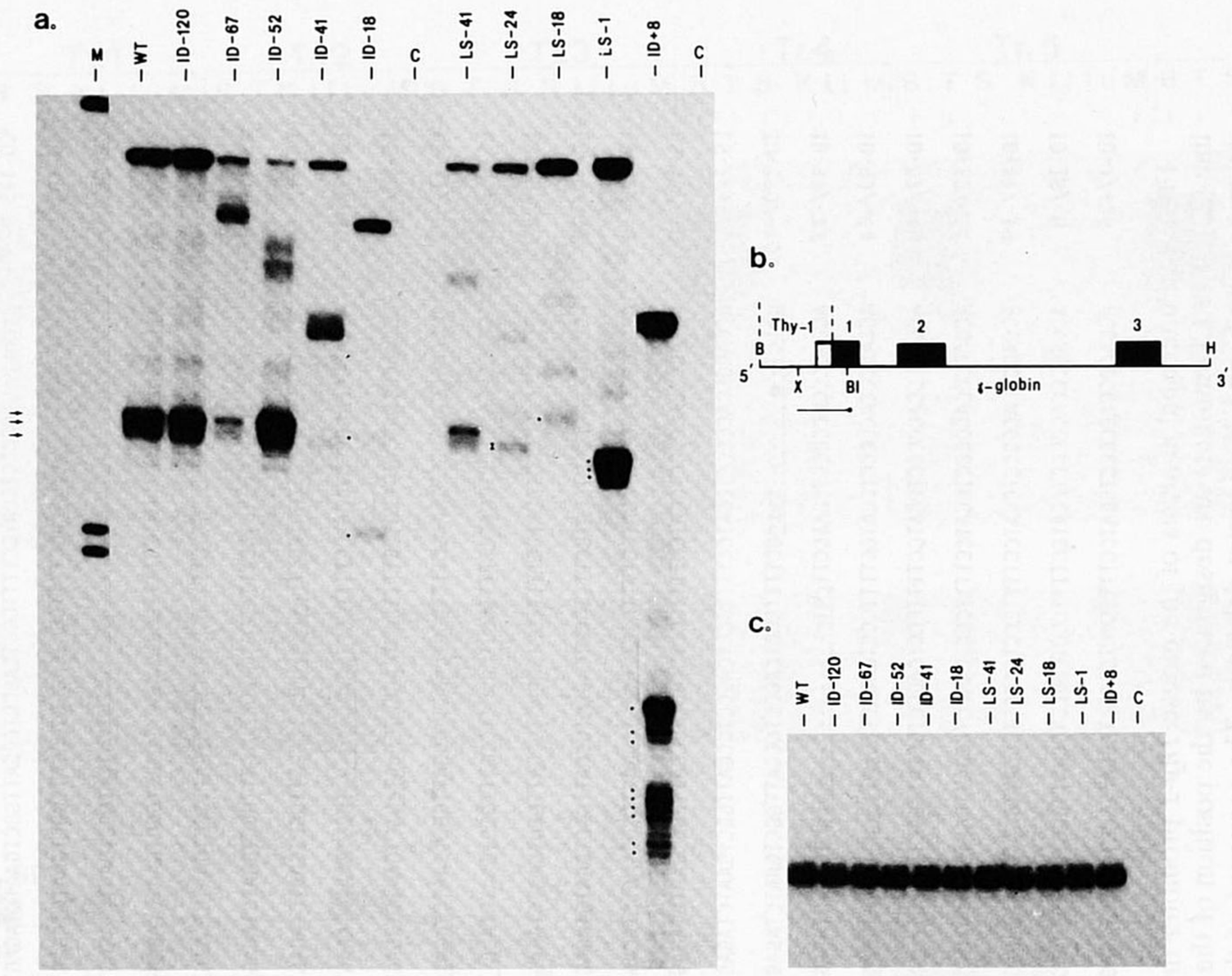

FIG. 3. Functional analysis of the *Thy-1* promoter by S1 nuclease protection. Mutants were cotransfected in HeLa cells with a human β-globin control plasmid and expressed transiently for 48 h. Total mRNA was analyzed by S1 protection. (a) Total RNA (30 μg) was hybridized to an *Xba*I-*Bgl*I 145-bp probe (labeled with T4 polynucleotide kinase) which detects the 5′-end termini of the fusion mRNAs and protects 84 to 90 bp of the transcripts. The position of wild-type 5′ ends is indicated by arrows; novel cap sites are indicated by dots. Of the larger protected bands, the uppermost band represents reannealed probe, while the lower bands represent upstream transcripts initiated in the vector. Lane M, marker; lane C, negative controls for wild-type (WT) and ID −18/+8 probes, representing hybridization to 30 μg of total RNA from nontransfected HeLa cells. (b) Schematic representation of the hybrid *Thy-1*/γ-globin constructs. The 5′-end S1 probe (*Xba*I to *Bgl*I fragment) is indicated. Restriction sites: B, *Bgl*II; S, *Sma*I; X, *Xba*I; BI, *Bgl*I; H, *Hin*dIII. (c) S1 analysis of human β-globin mRNA. A 5′-end probe (*Acc*I-*Acc*I 540-bp fragment which protects 160 bp) was used to detect the transcripts of the cotransfected human β-globin plasmid.

Thy-1 promoter function. The first (−91 to −77) contains an inverted CAAT box sequence homologous to the human α-globin CCAAT box motif (8). Regions from −68 to −48 and −41 to −19 include potential Sp1 binding sites. The third region is defined by sequences proximal to the transcription initiation site and is the subject of a separate report (61b). To further investigate the nature of the protein factors that bind to the Thy-1 promoter, DNase I footprinting and gel retardation analyses were used.

Binding of *trans*-acting factors to the minimal region of the *Thy-1* promoter. DNase I footprinting assays were performed by using a probe that contains the region between −270 and +36 of the *Thy-1* promoter in binding reactions with nuclear extracts derived from mouse brain, the murine T-cell line EL-4, HeLa cells, and the human erythroid line K562. Figure 4 shows the presence of four major footprints within the region from −100 to +8 which was found by expression analysis to be sufficient for promoter function. The four DNase I-protected areas (dashed lines in Fig. 5b) are readily detectable on both DNA strands and are present in all four extracts (Fig. 4A and B). Footprints III, IV, and V overlap with regions from −91 to −77, −68 to −48, and −41 to −19, respectively, which were found to be essential for promoter activity. However, incubation with different extracts produced a distinct pattern of hypersensitive sites (not related to species), and in some cases the borders of an individual protected region were also different between the various extracts. The region around −25 to −10 failed to show any obvious protection from DNase I, although incubation with HeLa extracts resulted in the formation of strong hypersensitive sites (Fig. 4A). Finally, one extensive, weak footprint was found in the sequences flanking the cap site, which we have arbitrarily subdivided into three regions on the basis of unprotected nucleotides (regions VI, VII, and VIII). Each of the footprints could be competed for by the corresponding oligonucleotides (Fig. 5b) in all four extracts (data not shown).

Characterization of the protein factors by mobility shift assays. To identify the individual proteins interacting with each of the DNase I-protected regions, a number of oligonucleotides were used as probes or competitors in gel mobility shift assays using a HeLa nuclear extract (Fig. 5). Probe THY-A represents the regions from −115 to −105 and −41 to −33 and binds four distinct complexes (Fig. 5a, lane 1). The first two slowly migrating complexes were also detected with probes THY-F and THY-G and to a lesser extent with THY-D. Since the common motif present in all of these probes was reminiscent of the Sp1 binding site, some of the probes were tested in competition experiments with a genuine GC box (GC boxes III and IV from the SV40 early promoter; see probe Sp1 GC III, IV in Materials and Methods). THY-A and THY-F bind complexes which comigrate with those retarded by an Sp1 probe (Fig. 5a; compare lanes 21, 23, and 25). These complexes are competed for very efficiently by the Sp1 probe (lanes 22 and 24), and conversely, THY-A and F compete for the Sp1 binding

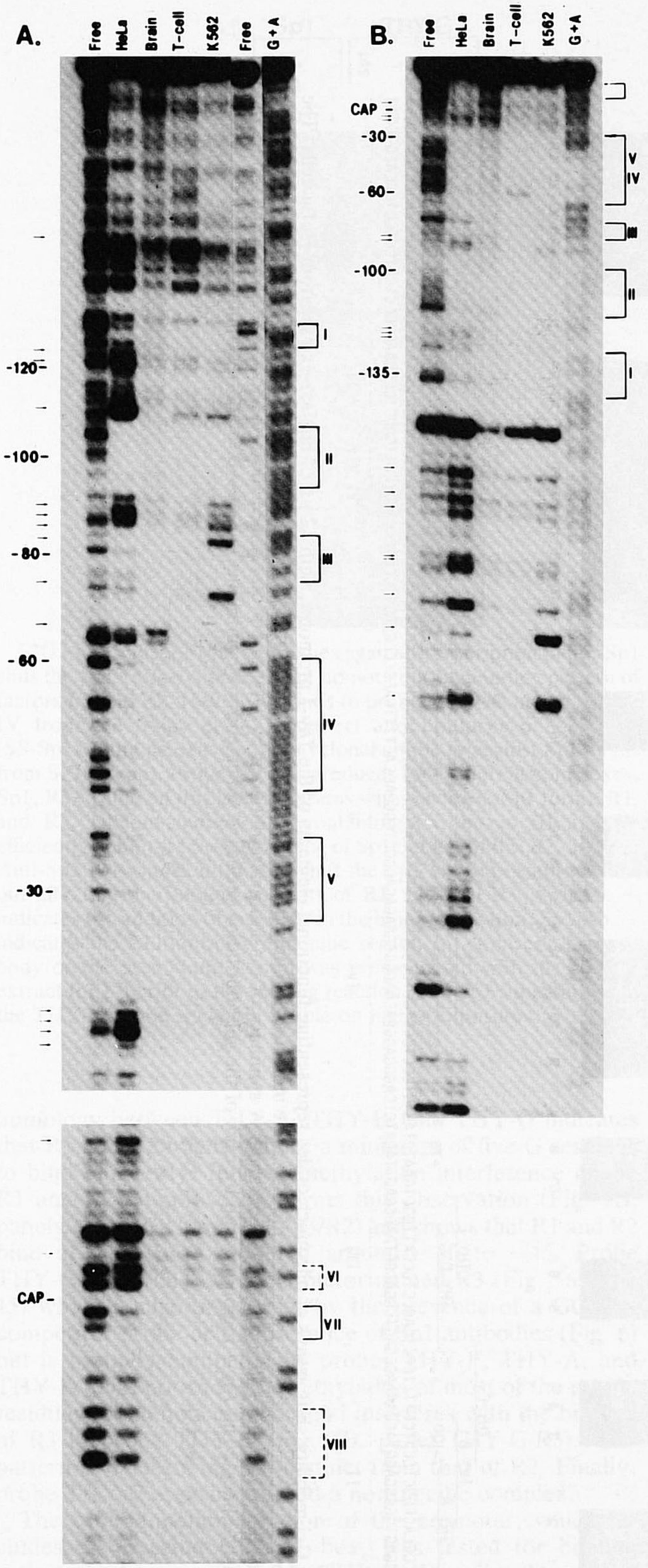

FIG. 4. Footprinting analysis of the *Thy-1* promoter. A 5′-end-labeled fragment containing the promoter region between −270 and +36 was used as a probe in binding reactions containing 50 μg of nuclear extracts from HeLa cells, mouse brain, T cells (cell line EL-4), and the erythroid line K562. Footprints are indicated on the right of each panel; hypersensitive sites are indicated by arrows on the left. (A [coding strand]) Two different lanes of DNA without extract (free); (B) noncoding strand.

pattern (lanes 27 and 28). The THY-A motif CCCCACCCCC represents a strong Sp1 binding site, as revealed by titration competition experiments (data not shown and Fig. 5a, lanes 25 and 27) and reported by Letovsky and Dynan (46). The THY-F CCCTCCC motif also binds Sp1 but with a lower affinity than the CCGCCC or the CCCCACCCCC motifs and is also present in the epidermal growth factor receptor promoter, where it was shown to bind purified Sp1 (42). Probe THY-G binds Sp1 by the CCCCACCC motif, whereas THY-D has a weak Sp1 site represented by the CCTCC sequences. Binding of Sp1 to these motifs was also tested by polyclonal antibodies raised against Sp1 (a gift from S. Jackson) which shift the upper complex formed by probes THY-A, THY-F, THY-G, and THY-D in gel mobility shift assays (Fig. 6 and data not shown). The upper complex resolves into two distinct bands which are consistent with the presence of two Sp1 forms (95 and 105 kDa) (5, 34, 39). The lower complex is designated ?Sp1 in Fig. 6 since it might represent a degradation or posttranslational product of Sp1 (14, 27, 46). In our experiments this complex shows binding affinity identical to that of Sp1 and an identical competition pattern, and it is present in all nuclear extracts tested. However, its binding remains unaffected by the presence of anti-Sp1 antibodies (Fig. 6) (which, however, do not recognize the carboxy terminus of the protein [33a]), and it has been suggested by others that it represents a factor distinct from Sp1 (71). Moreover, the factor may copurify with transcription factor LSF (30). It is therefore possible that ?Sp1 represents a factor other than Sp1.

Methylation interference experiments using probes THY-A, THY-F, and Sp1 GC III,IV show that methylation of most of the available G residues flanking the central T residue prevents the binding of Sp1 (Fig. 7A). This could explain the high binding affinity of the CCCCACCCCC motif. Partial proteolysis of complexes formed on probes THY-A, THY-F, and Sp1 GC III,IV also indicated that the upper complex is indeed Sp1 (data not shown). It should be mentioned that GC box IV of the SV40 promoter is a particularly weak Sp1 binding site, but its affinity is enhanced by mutations in GC box V (22). We therefore attribute the increased interactions between Sp1 and GC box IV (Fig. 7A) to the absence of GC box V.

Probe THY-B fails to bind Sp1 although it contains a candidate Sp1 binding site (also indicated by competition assays; data not shown). Instead, it binds a novel factor (Fig. 5a, lanes 3 and 4) which interacts weakly with most of the available purines present on the noncoding strand (Fig. 7B, panel THY-B). Probe THY-D binds a factor which comigrates with the complex observed on probe THY-B. However, competition assays between probes THY-B and THY-D strongly indicate that the two factors are different. This is verified by their distinct protection patterns obtained by methylation interference (Fig. 7B; compare panels THY-D and THY-B). We have not characterized these proteins further since they are not observed in T cells (see below) and deletion of these sequences only has a marginal effect on the RA (Fig. 2). In addition, THY-B binds two other complexes also obtained by probes THY-A and particularly THY-G. These two factors, which we call R1 and R2, bind with different affinities to most of the *Thy-1* probes that bind Sp1 (Fig. 5, lanes 1, 3, and 15), suggesting that they may be degradation products of Sp1. However, R1 and R2 do not bind to the GC box probe, nor is their binding to probes THY-A and THY-G prevented by low concentrations of GC box competitor (Fig. 6). Moreover, Sp1 polyclonal antibodies do not shift complexes R1 and R2 (Fig. 6), and the

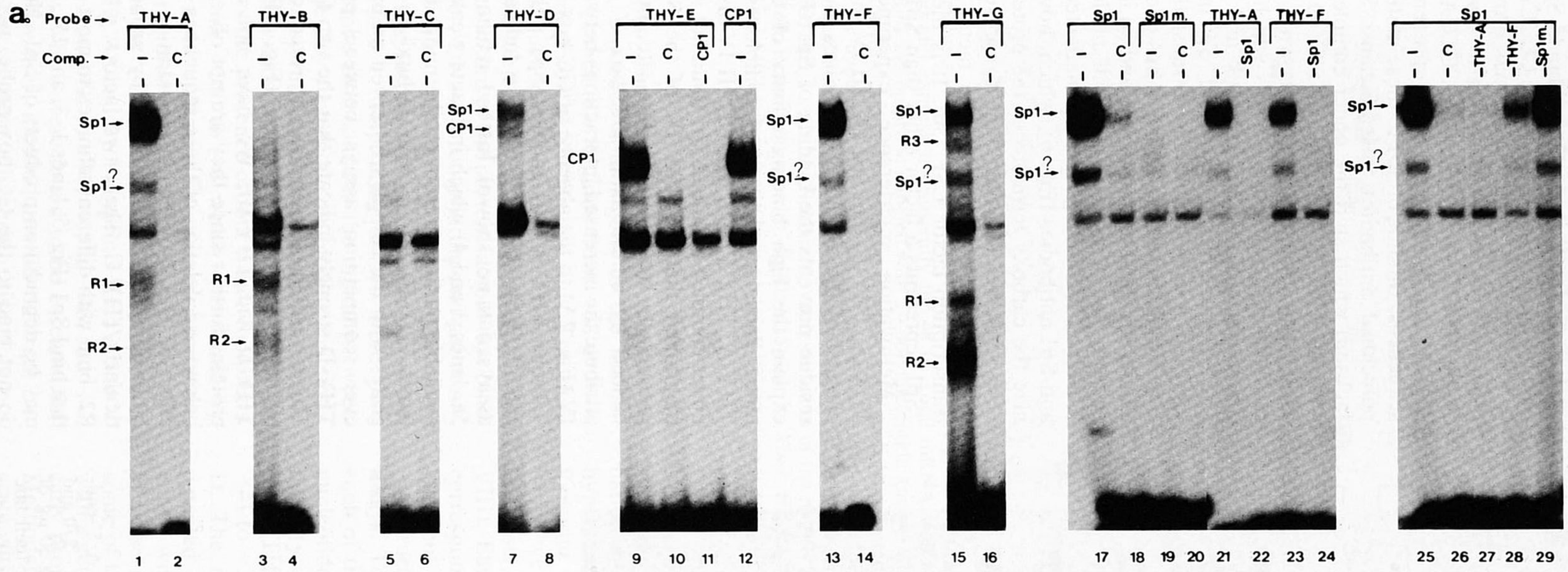

FIG. 5. (a) Gel mobility shift assays of oligonucleotides representing functional elements of the *Thy-1* promoter, with 12 μg of nuclear extract from HeLa cells. The upper line above each panel indicates the probe; the lower line indicates the competitor. Homologous competitors are denoted by a c above each lane. (b) *Thy-1* promoter sequences between −120 and +40 for both DNA strands. A summary of the footprints obtained by DNase I protection in the T-cell extract is presented below the sequences (dashed lines). The *Thy-1* region covered by each of the gel retardation probes is indicated by a continuous line above the corresponding sequences.

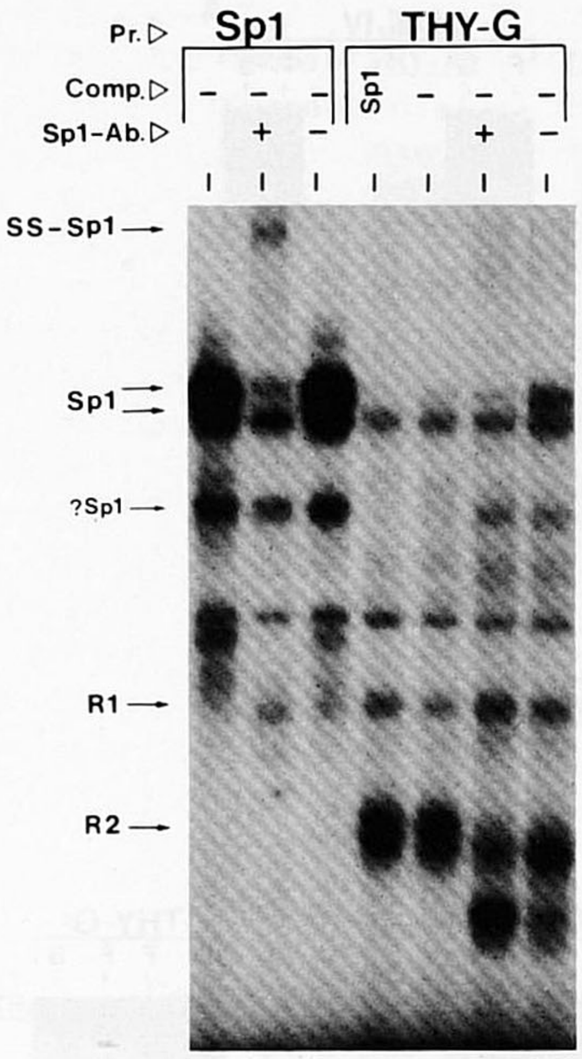

FIG. 6. Evidence that antibodies against transcription factor Sp1 shift the Sp1-formed complex but do not alter the binding pattern of factors R1 and R2. HeLa Sp1 binds to probe Sp1 (GC boxes III and IV from the SV40 early promoter) and produces a supershift (SS-Sp1) in the presence of a polyclonal antibody against Sp1 (a gift from S. Jackson). Probe THY-G produces four different complexes, Sp1, R3 (which on this gel comigrates with the lower Sp1 form), R1, and R2. Oligonucleotide Sp1 containing GC boxes III and IV efficiently competes for the binding of Sp1 but not of R1, R2, and R3. Anti-Sp1 polyclonal antibodies shift the Sp1-formed complex without affecting the binding patterns of R1, R2, and R3. Sp1-Ab. + indicates the addition of antibody in the binding reaction; Sp1-Ab. − indicates the addition of preimmune serum. In all cases, the antibody or the preimmune serum was preincubated with the nuclear extract for 1 h prior to the binding reaction. The Sp1 supershift with the THY-G probe is clearly visible on higher exposures.

homology between THY-A, THY-B, and THY-G indicates that R1 and R2 might require a minimum of five G residues to bind efficiently. Indeed, methylation interference of the R1 and R2 complexes confirms this observation (Fig. 7B, panels THY-B/R1 and THY-G/R2) and shows that R1 and R2 bind at a position centered around −50 to −45. Probe THY-G binds an additional factor named R3 (Fig. 5a, lane 15) which is also unaffected by the presence of a GC box competitor probe or the presence of Sp1 antibodies (Fig. 6) but is partially competed by probes THY-F, THY-A, and THY-D (data not shown). Methylation of most of the purine residues in the noncoding strand interferes with the binding of R3 to probe THY-G (Fig. 7B, panel THY-G/R3). This pattern is overlapping but distinct from that of R2. Finally, probe THY-C appears to bind a nonspecific complex.

The other functional region of the promoter, which includes an inverted CCAAT box, was tested for binding activity with oligonucleotide THY-E (Fig. 5b). The probe binds the CCAAT-box-specific transcription factor CP1 (Fig. 5a, lanes 9 to 12; Fig. 8B) which has been extensively described (8), and its binding pattern is not competed for by an oligonucleotide containing the NF-I binding site (Fig. 8B). Figure 8B also shows the binding pattern of the human α-globin CCAAT box (CP1 activity) and that of a probe containing the adenovirus major late origin of replication (NF-I activity) in all four nuclear extracts tested. Surprisingly, probe THY-D was also able to bind CP1 although with a lower affinity since it contains part of the binding site but carries a mutation (G to C) in the G doublet which has been shown to interact with the protein (8). Thus, the results obtained from the binding assays and the expression analyses indicate that within the region from −100 to −35 of the *Thy-1* promoter the two dominant transcription factors are Sp1 and CP1, which bind to sequences centered around −80 (CP1) and −60 (Sp1).

Complex formation in the presence of brain and T-cell nuclear extracts. The binding pattern of Thy-1 probes was also tested in the presence of brain and T-cell nuclear extracts. With the exception of THY-C, which is from a nonessential part of the promoter, all oligonucleotides exhibit a binding pattern resembling the one obtained with HeLa extracts, but the presence of factors R1 and R2 is more evident in the T-cell extract (Fig. 8A, lanes 1, 3, and 11). Probe THY-F shows enhanced binding of a factor that migrates below the Sp1 complex and is more prominent in T-cell and brain nuclear extracts, although it is also present at low levels in the HeLa extract. Probe THY-G selectively binds a large amount of an R2-like complex even though Sp1 is present in the brain extract (lanes 12 and 14).

The *Thy-1* CCAAT box probe (THY-E) exhibits a binding pattern similar to that of the α-globin CCAAT box in brain and thymus extracts, and this pattern is not affected by the NF-I competitor (Fig. 8B). However, other patterns can be obtained with other extracts (e.g., K562; Fig. 8B).

DISCUSSION

Specificity of the *Thy-1* promoter. We have previously used transgenic mice as an in vivo expression system to identify a number of tissue-specific enhancer elements which are located downstream from the *Thy-1* promoter. These elements also function when the *Thy-1* promoter is replaced by heterologous (non-tissue-specific) promoters (67). However, these results did not exclude that the *Thy-1* promoter itself is not tissue specific. The data obtained in this study by using the *Thy-1*/γ-globin transgene clearly show that the *Thy-1* promoter fails to activate initiation of transcription in the absence of enhancer sequences. With the exception of transgenic line Tr.4, which expresses very low levels of the transgene in all of the examined tissues, the hybrid construct remained transcriptionally silent in all six transgenic lines. Some transgenic transcripts were present in the brain of three lines at levels of less than 1% compared with the endogenous *Thy-1* mRNA. Taking into account the high copy number of the transgenes, the expression level per transgene becomes even lower. The brain-specific transcripts might stem from the diminished expression of the hybrid gene in the different neuronal subtypes expressing *Thy-1* or the specific expression of the transgene in a particular subtype of neuronal cells. Alternatively, they might represent aberrant expression of the transgene in nonneuronal cells because they do not show the pattern of initiation typical for the brain (62). The inability of the *Thy-1* promoter to initiate transcription in the absence of enhancer sequences agrees with the results obtained from a number of cell lines (23, 33). Clearly no specificity is conferred by the promoter, suggesting that *Thy-1* is one of the few genes (3, 9, 20) with tissue-specific control elements only located downstream of promoter elements. In contrast, a number of tissue-specific genes contain promoters with housekeeping characteristics (for a review, see reference 19), but unlike

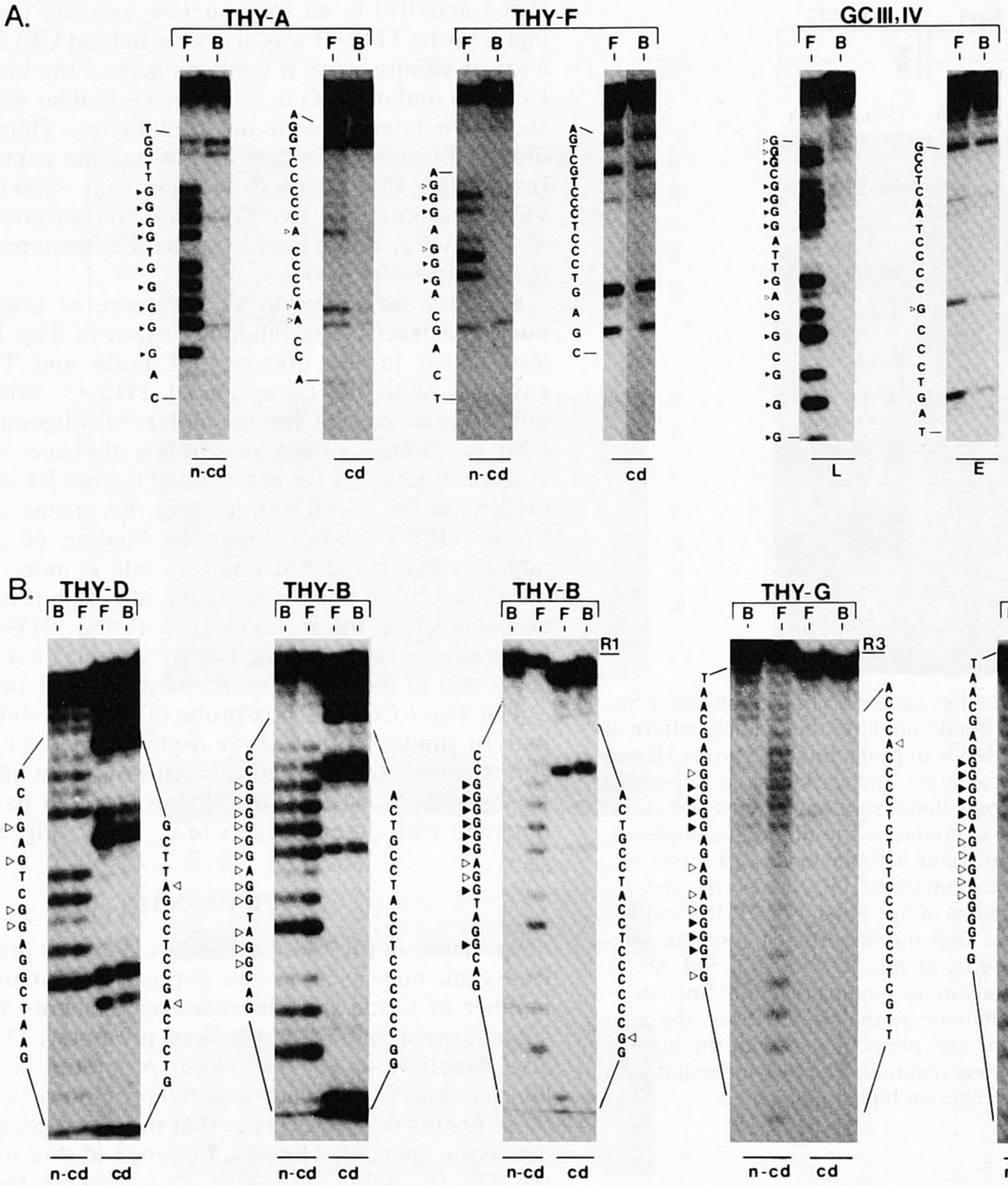

FIG. 7. (A) Methylation interference patterns of transcription factor Sp1. The coding (cd) and noncoding (n-cd) strands of oligonucleotides THY-A, THY-F, and the early (E) and late (L) strands of GC boxes III and IV of the SV40 promoter region, were partially methylated (48) and incubated with nuclear extracts derived from HeLa cells (gel retardation binding reactions were scaled up 10-fold). DNA was purified from the excised bands by electroelution and subjected to sodium hydroxide-sodium phosphate treatment to cleave the methylated purine residues. Free (F) and complexed (B) DNA is indicated above each lane. Open and filled triangles indicate weak and strong interference, respectively. (B) Methylation interference patterns of the abundant complexes bound to probes THY-D and THY-B and of factors R1, R2, and R3. Factor R1 is complexed with probe THY-B (since it exhibits the highest affinity), whereas factors R3 and R2 are detected by the same probe (THY-G).

Thy-1, they exhibit transcriptional promiscuity in a diverse number of heterologous cells after gene transfer (60, 69).

Distal domains required for promoter function. (i) The CCAAT box region. The first element essential for promoter function maps in the region containing an inverted CCAAT box. The corresponding *Thy-1* CCAAT probe reveals a certain heterogeneity when analyzed by gel retardation assays. The probe mainly detects CP1 activity in all four extracts tested. However, it also binds to a faster-migrating complex that is competed for by the NF-I probe. An additional minor activity unique to the *Thy-1* probe is observed in the T-cell extracts. It is now well established that distinct CCAAT-binding proteins with variable binding affinities exist within a cell (8, 17, 25, 35, 36, 59). In the case of the *Thy-1* promoter, the major CCAAT-binding activity is contributed by CP1, which therefore is the primary candidate for participating in the transcriptional activation of the promoter.

(ii) The region from −68 to −31 and binding of several proteins to GC-rich motifs. The functional importance of this region is demonstrated by a number of promoter mutations. Substitution of the CCCTCCC motif (−63 to −57) by linker sequences has a detrimental effect on promoter efficiency. This motif binds transcription factor Sp1, as suggested by the competition assays and shown by antibody experiments. Sp1 also binds to the CCCCACCCC motif (−41 to −33) which constitutes the other important element for promoter function. These data suggest that Sp1 binding activity is essential for the in vivo activation of the *Thy-1* promoter, whereas the data from transgenic mice indicate that Sp1

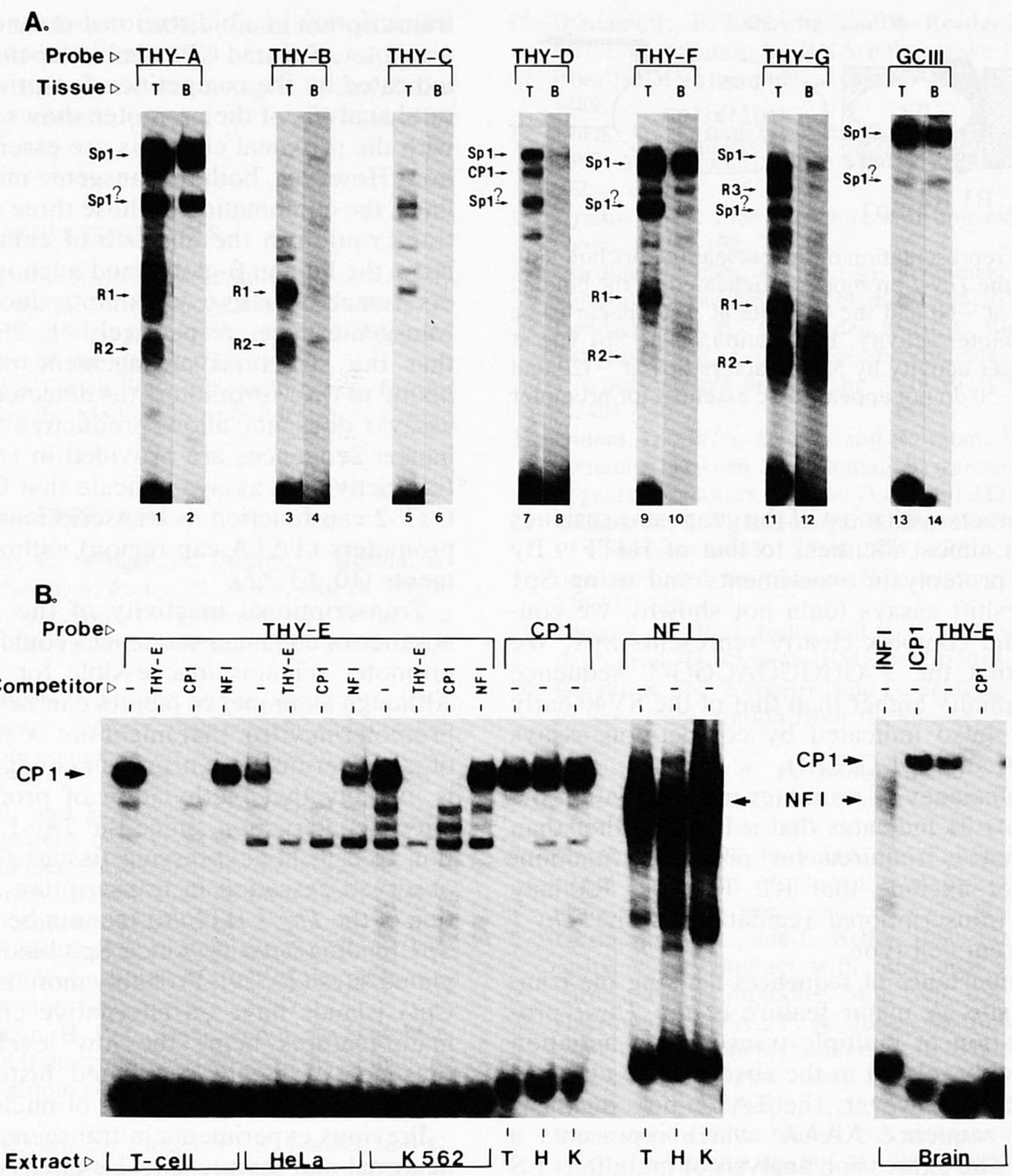

FIG. 8. Complex formation on oligonucleotides THY-A to THY-G and the Sp1 GC III, IV probe in the presence of brain or T-cell nuclear extracts. (A) Oligonucleotides shown in Fig. 5 were tested for binding activity in the presence of 12 μg of whole mouse brain (B) or T-cell (EL-4 line) (T) nuclear extracts. (B) The binding pattern of the *Thy-1* probe, THY-E, containing the CAAT box motif is shown in reactions containing 12 μg of nuclear extracts from cell lines HeLa, EL-4 (T cell), and K562 as well as from mouse brain. The binding pattern is compared with that of the α-globin CAAT box motif (CP1) and the NF-I binding site derived from the adenovirus origin of replication.

(and/or CP1) alone is not sufficient for expression. The binding assays revealed the abundant presence of Sp1 not only in HeLa cells but also in K562, murine brain, and T-cell nuclear extracts. Transcription factor Sp1 is found in two forms, both of which are glycosylated (34). The DNA binding and transcription activation domains of the protein have been extensively characterized (11, 39, 40); Sp1 interacts with (possibly) only one of the two DNA strands (21) and activates transcription in a bidirectional manner (22). Activation can occur from either distal or proximal elements and in a synergistic way (10). These properties of the protein are particularly interesting in the context of potential interactions with other transcription factors. However, it remains to be determined whether Sp1 mediates its effect directly or through an as yet unidentified factor(s).

The region from −57 to −33 binds three proteins, termed R1, R2, and R3, which appear to be distinct from Sp1 by a number of criteria. They bind with high affinity to probe THY-G but not to the Sp1 probe (Fig. 5 to 7). Probe THY-B preferentially binds R1, whereas probe THY-A binds R1 and R2 with low affinities. R1 and R2 activities are more abundant in the T-cell extract than in HeLa cells. The homology between the probes predicts that factors R1 and R2 require the presence of five G residues to bind efficiently. This view is supported by the methylation interference data which show that R1 and R2 bind to five adjacent G residues, whereas factor R3 interacts weakly with all purines available (Fig. 7B). Lastly, Sp1 antibodies only interfere with Sp1 binding. An increasing number of nuclear proteins, exhibiting activator or repressor function, have been found to interact with GC-rich motifs (12, 13, 32, 42, 43, 52, 53, 68). The most striking example of these is factor LSF, which binds to two adjacent Sp1 binding sites positioned 10 bp apart (30). It remains to be determined what the nature of factors R1, R2, and R3 is, but on the basis of the methylation interference data we cannot correlate them to any other reported factors. Surprisingly, we did not detect an activity equivalent to the HeLa factor H4TF1, which binds to the sequence 5′-GGGGGAGGG-3′ (12, 13). Our methylation interference data on probe THY-A clearly indicate that the

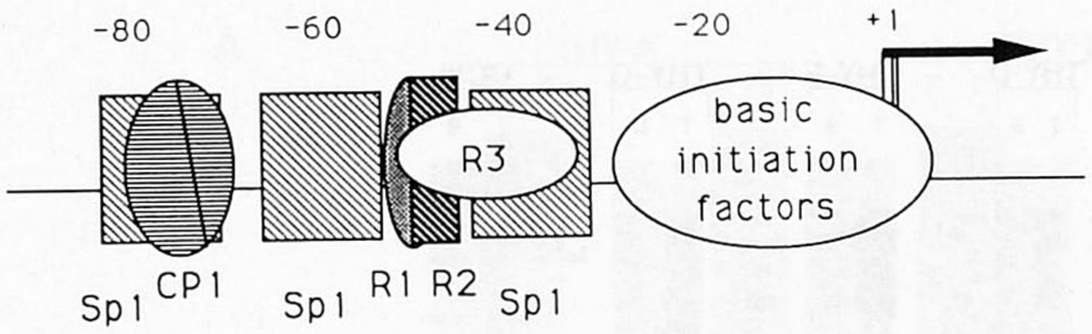

FIG. 9. Schematic representation of the nuclear factors bound to the minimal region of the *Thy-1* promoter as indicated by the binding studies. The Sp1 site at −60 and the CP1 site at −80 appear to be indispensable for promoter activity. Elimination of the Sp1 site at −40 decreases promoter activity by 55%. Factors Sp1 at −120 and factors R1 and R2 at −50 do not appear to be essential for promoter activity in HeLa cells.

bound complex interacts with most of the guanosine residues available, a pattern almost identical to that of H4TF1. By performing partial proteolysis experiments and using Sp1 antibodies in supershift assays (data not shown), we conclude that the bound complex clearly represents Sp1. We therefore believe that the 5′-GGGGGAGGG-3′ sequence binds Sp1 with an affinity higher than that of the SV40 early promoter fragment (also indicated by competition assays with probes THY-A and Sp1 GC III).

The expression efficiency of promoter mutants within the region from −68 to −31 indicates that it is Sp1 rather than R1, R2, or R3 that is required for promoter function, although we cannot exclude that R1, R2, and R3 may participate in the transcriptional regulation of the *Thy-1* promoter in a different cell type.

The functional importance of sequences flanking the transcription initiation site. A major feature of the *Thy-1* promoter is the utilization of multiple transcription initiation sites, which is possibly related to the absence of a classical TATA motif (23, 62). However, the TATA box region is substituted by the sequence AAAA, which represents a weak TATA motif. The expression analysis of mutations LS −24/−19 and −18/−10 and ID −18/+8 clearly shows that sequences flanking the cap site are essential for promoter function, whereas mutations in the cap site region (LS −1/+8) result in a 55% decrease in transcription efficiency. In addition, S1 analysis of the exogenous transcripts revealed that all mutations located around the principal start site (LS −24/19, LS −18/−10, and LS −1/+8) produce mRNAs with novel 5′ ends. Hence, one or more basal transcriptional elements are situated in the region from −30 to +8, regulating the level and site of transcription initiation. Preliminary studies have identified a number of binding activities within this region (61a).

Possible mechanisms involved in *Thy-1* promoter activation. The individual functional domains of the *Thy-1* promoter are summarized in Fig. 9. The dominant distal elements are the inverted CCAAT box at −85 to −75 and transcription factor CP1 that binds to it, the Sp1 binding site at −63 to −56, and, to a lesser extent, the Sp1 site at −41 to −33. The presence of a functional CCAAT box in a promoter without a TATA box is a rather uncommon motif in this class of promoters. Tissue-specific promoters without a TATA motif constitute 10 to 20% of the published promoter sequences. Those containing an HTF island and a CAAT box in addition constitute less than 1% (61a). The functional significance of coordinated interactions of Sp1 and CCAAT-binding proteins with additional promoter elements was initially shown in the context of the herpes simplex virus thymidine kinase promoter (38, 51). Both elements are capable of activating transcription in a bidirectional manner (22, 50). In the *Thy-1* promoter, Sp1 and CP1 bind independently of each other, as indicated by the competition footprinting assays. The structural analysis of the promoter shows that CP1 and Sp1 along with the proximal elements are essential for promoter function. However, both in transgenic mice and in tissue culture lines, the combination of those three elements fails to initiate transcription in the absence of enhancer sequences. Similarly, the human β-globin and adenosine deaminase promoters remain inactive when introduced into cell lines and transgenic mice, respectively (1, 26). This finding implies that the structural arrangement of transcription factors bound to these promoters (as detected by the in vitro binding assays) does not allow productive interactions unless enhancer sequences are provided in *cis*, even though in vivo transactivation assays indicate that factors such as Sp1 and OTF-2 can function as transcriptional activators of minimal promoters (TATA-cap region) without any other distal elements (10, 63, 65).

Transcriptional inactivity of the *Thy-1* promoter in the absence of enhancer sequences could also be explained if the promoter remains inaccessible for Sp1 and CP1 to bind. Although a number of factors can be complexed to the *Thy-1* promoter in vitro, that might not be the case in vivo because of a nonpermissive chromatin configuration. At this stage it is unlikely that methylation of promoter sequences might affect its activation, since the *Thy-1* HTF remains methylation free in nonexpressing tissues (45, 62). Moreover, the observed cessation in transcription after artificial methylation of the *Thy-1* HTF (61) cannot be attributed to the lack of Sp1 binding activity, since Sp1 binds equally well to methylated sites (28, 29). Probably more important is the fact that CpG islands have an alternative chromatin structure, the main features being the low levels of histone H1, the presence of highly acetylated histones H3 and H4, and finally an overrepresentation of nucleosome-free areas (64).

Previous experiments in transgenic mice have shown that neuronal and thymic Thy-1 expression is absolutely dependent on the presence of downstream enhancer sequences (67). This is confirmed by the analysis of the *Thy-1*/γ-globin hybrid gene, which shows the inability of the promoter to function in the absence of enhancer sequences or to confer any specificity to the foreign gene. The alternative explanation for this result is a dominant suppression of the promoter in the nonexpressing tissues. A repressor activity was implicated in cell fusion experiments between *Thy-1*-positive and -negative cell lines (31, 58). Elimination of different *Thy-1* sequences (upstream of −270 or downstream of +36) did not result in a reproducible gain of expression in the negative tissues (67). This suggests that a potential repressor function could be present only in the promoter. Interestingly, the region between −50 and −41 of the *Thy-1* promoter shows homology in eight of nine nucleotides with a silencer sequence found in the human ϵ-globin gene (6) and the chicken lysozyme box 1 (2). In addition, deletion of this sequence in the mutant ID −52/−31 leads to fourfold stimulation of transcription (compared with LS −41/−31) when assayed in the presence of enhancer sequences.

In conclusion, the results indicate that the *Thy-1* promoter functions through the coordinated interaction of the two transcription factors CP1 and Sp1 and the proximal promoter elements. However, the structure of the promoter does not permit productive interactions between the different factors as a result of the absence of *cis*-acting enhancer sequences or the presence of a repressor activity. Tissue specificity is conferred by the downstream elements, and therefore the

promoter acts as the mediator rather than the primary activator of transcription initiation. The extensive analysis of promoter mutants in transgenic mice might uncover the mechanisms underlying transcriptional activation in the *Thy-1* gene.

ACKNOWLEDGMENTS

We are grateful to N. Yannoutsos for generating the transgenic mice lines, S. Jackson for the gift of Sp1 polyclonal antibodies, P. Rigby and D. Kioussis for helpful suggestions, and Cora O'Carroll for assistance in preparation of the manuscript.

E.S. was supported by the Greek State Scholarship Foundation. This work was supported by the Medical Research Council, United Kingdom.

REFERENCES

1. **Aronow, B., D. Lattier, R. Silbiger, N. Dusing, J. Hutton, G. Jones, J. Stock, J. McNeish, S. Potter, D. Witte, and D. Wiginton.** 1989. Evidence for a complex regulatory array in the first intron of the human adenosine deaminase gene. Genes Dev. **3:**1384–1400.
2. **Baniahmad, A., M. Muller, C. Steiner, and R. Renkawitz.** 1987. Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements. EMBO J. **6:**2297–2303.
3. **Basler, K., P. Siegrist, and E. Hafen.** 1989. The spatial and temporal expression pattern of sevenless is exclusively controlled by gene-internal elements. EMBO J. **8:**2381–2386.
4. **Bird, A.** 1986. CpG-rich islands and the function of DNA methylation. Nature (London) **321:**209–213.
5. **Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science **234:**47–52.
6. **Cao, S. X., P. D. Gutman, H. P. G. Dave, and A. N. Schechter.** 1989. Identification of a transcriptional silencer in the 5′-flanking region of the human ϵ-globin gene. Proc. Natl. Acad. Sci. USA **86:**5306–5309.
7. **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.
8. **Chodosh, C., A. Baldwin, R. Carthew, and P. Sharp.** 1988. Human CCAAT-binding proteins have heterologous subunits. Cell **53:**11–24.
9. **Clevers, H., N. Lonberg, S. Dunlap, E. Lacy, and C. Terhorst.** 1989. An enhancer located in a CpG-island 3′ to the TCR/CD3-ϵ gene confers T lymphocyte-specificity to its promoter. EMBO J. **8:**2527–2535.
10. **Courey, A. J., D. A. Holtzman, S. P. Jackson, and R. Tjian.** 1989. Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell **59:**827–836.
11. **Courey, A. J., and R. Tjian.** 1988. Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell **55:**887–898.
12. **Dailey, L., S. M. Hanly, R. G. Roeder, and N. Heintz.** 1986. Distinct transcription factors bind specifically to two regions of the human histone H4 promoter. Proc. Natl. Acad. Sci. USA **83:**7241–7245.
13. **Dailey, L., S. B. Roberts, and N. Heintz.** 1988. Purification of the human histone H4 gene-specific transcription factors H4TF-1 and H4TF-2. Genes Dev. **2:**1700–1712.
14. **Dawson, P. A., S. L. Hofmann, D. R. van der Westhuyzen, T. C. Sudhof, M. S. Brown, and J. L. Goldstein.** 1988. Sterol-dependent repression of low density lipoprotein receptor promoter mediated by 16-base pair sequence adjacent to binding site for transcription factor Sp1. J. Biol. Chem. **263:**3372–3379.
15. **de Boer, E., M. Antoniou, V. Mignotte, L. Wall, and F. Grosveld.** 1988. The human β-globin promoter; nuclear protein factors and erythroid specific induction of transcription. EMBO J. **7:**4203–4212.
16. **Dignam, J., R. Lebowitz, and R. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.
17. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CAAT-box binding proteins. Cell **50:**863–872.
18. **Dynan, S. W.** 1986. Promoters of housekeeping genes. Trends Genet. **6:**196–197.
19. **Gardiner-Garden, M., and M. Frommer.** 1987. CpG islands in vertebrate genomes. J. Mol. Biol. **196:**261–282.
20. **Georgopoulos, K., P. van den Elsen, E. Bier, A. Maxam, and C. Terhorst.** 1988. A T-cell specific enhancer is located in a DNase I hypersensitive area at the 3′ of the CD3-δ gene. EMBO J. **7:**2401–2407.
21. **Gidoni, D., W. S. Dynan, and R. Tjian.** 1984. Multiple specific contacts between a mammalian transcription factor and its cognate promoters. Nature (London) **312:**409–413.
22. **Gidoni, D., J. T. Kadonaga, H. Barrera-Saldana, K. Takahashi, P. Chambon, and R. Tjian.** 1985. Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. Science **230:**511–517.
23. **Giguere, V., K.-I. Isobe, and F. G. Grosveld.** 1985. Structure of the murine Thy-1 gene. EMBO J. **4:**2017–2024.
24. **Gorski, K., M. Carneiro, and U. Schibler.** 1986. Tissue-specific *in vitro* transcription from the mouse albumin promoter. Cell **47:**767–776.
25. **Graves, B. J., P. F. Johnson, and S. L. McKnight.** 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell **44:**565–576.
26. **Green, M. R., R. Treisman, and T. Maniatis.** 1983. Transcriptional activation of cloned human β-globin genes by viral immediate-early gene products. Cell. **35:**137–148.
27. **Gustafson, T. A., and L. Kedes.** 1989. Identification of multiple proteins that interact with functional regions of the human cardiac α-actin promoter. Mol. Cell. Biol. **9:**3269–3283.
28. **Harrington, M. A., P. A. Jones, M. Imagawa, and M. Karin.** 1988. Cytosine methylation does not affect binding of transcription factor Sp1. Proc. Natl. Acad. Sci. USA **85:**2066–2070.
29. **Holler, M., G. Westin, J. Jiricny, and W. Schaffner.** 1988. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. Genes Dev. **2:**1127–1135.
30. **Huang, H.-C., R. Sundseth, and U. Hansen.** 1990. Transcription factor LSF binds two variant bipartite sites within the SV40 late promoter. Genes Dev. **4:**287–298.
31. **Hyman, R., and V. Stallings.** 1978. Evidence for a gene extinguishing cell-surface expression of the Thy-1 antigen. Immunogenetics **6:**447–458.
32. **Imagawa, M., R. Chiu, and M. Karin.** 1987. Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell **51:**251–260.
33. **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. Cell. Biol. **6:**2923–2931.

33a. **Jackson, S.** Personal communication.

34. **Jackson, S. P., and R. Tjian.** 1988. Unexpected O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. Cell **55:**125–133.
35. **Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. L. McKnight.** 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes Dev. **1:**133–146.
36. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58:**799–839.
37. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48:**79–89.
38. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell **42:**559–572.
39. **Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian.** 1987. Isolation of cDNA encoding transcription factor Sp1 and

functional analysis of the DNA binding domain. Cell **51**:1079–1090.
40. **Kadonaga, J. T., A. J. Courey, J. Ladika, and R. Tjian.** 1988. Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science **242**:1566–1570.
41. **Kadonaga, J. T., K. A. Jones, and R. Tjian.** 1986. Promoter-specific activation of RNA polymeraseII transcription by Sp1. Trends Biochem. **11**:20–23.
42. **Kageyama, R., G. T. Merlino, and I. Pastan.** 1988. A transcription factor active on the epidermal growth factor receptor gene. Proc. Natl. Acad. Sci. USA **85**:5016–5020.
43. **Kageyama, R., and I. Pastan.** 1989. Molecular cloning and characterization of a human DNA binding factor that represses transcription. Cell **59**:815–825.
44. **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.
45. **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**:9667–9778.
46. **Letovsky, J., and W. S. Dynan.** 1989. Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence. Nucleic Acids Res. **17**:2639–2653.
47. **Low, M. G., and P. W. Kincade.** 1988. Phosphatidyl inositol in the membrane anchoring domain of the Thy-1 glycoprotein. Nature (London) **318**:62–65.
48. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavage. Methods Enzymol. **65**: 499–560.
49. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-encoding gene. Science **217**:316–324.
50. **McKnight, S. L., R. C. Kingsbury, A. Spence, and M. Smith.** 1984. The distal transcription signals of the herpes virus tk gene share a common hexanucleotide control sequence. Cell. **37**:253–262.
51. **McKnight, S. L., and R. Tjian.** 1986. Transcriptional selectivity of viral genes in mammalian cells. Cell **46**:795–805.
52. **Mermod, N., T. J. Williams, and R. Tjian.** 1989. Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcripts *in vitro*. Nature (London) **332**:557–560.
53. **Mitchell, P. J., C. Wang, and R. Tjian.** 1987. Positive and negative regulation of transcription *in vitro*: enhancer-binding protein AP-2 is inhibited by SV40 T-antigen. Cell **50**:847–861.
54. **Morris, R. J.** 1985. Thy-1 in developing nervous tissue. Dev. Neurosci. **7**:133–160.
55. **Morris, R. J., and F. Grosveld.** 1989. Expression of Thy-1 in the nervous system of the rat and mouse, p. 121–148. *In* A. E. Reif and M. Schlesinger (ed.), Cell surface antigen Thy-1. Marcel Dekker, Inc., New York.
56. **Ptashne, M.** 1986. Gene regulation by proteins acting nearby and at a distance. Nature (London) **322**:697–701.
57. **Reif, A., and J. Allen.** 1964. The AKR thymic antigen and its distribution in leukemias and nervous tissues. J. Exp. Med. **120**:413–433.
58. **Saleh, M., and P. F. Bartlett.** 1989. Evidence from neuronal heterokaryons for a trans-acting factor suppressing Thy-1 expression during neuronal development. J. Neurol. Sci. **23**: 406–415.
59. **Santoro, C., N. Mermod, P. C. Andrews, and R. Tjian.** 1988. A family of CAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature (London) **334**:218–224.
60. **Sehgal, A., N. Patil, and M. Chao.** 1988. A constitutive promoter directs expression of the nerve growth factor receptor gene. Mol. Cell. Biol. **8**:352–362.
61. **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.
61a. **Spanopoulou, E.** Unpublished data.
61b. **Spanopoulou, E., et al.** Submitted for publication.
62. **Spanopoulou, E., V. Giguere, and F. G. Grosveld.** 1988. Transcriptional unit of the murine Thy-1 gene: different distribution of transcription initiation sites in brain. Mol. Cell. Biol. **8**:3847–3856.
63. **Tanaka, M., V. Grossniklaus, W. Herr, and N. Hernandez.** 1988. Activation of the U2 sn RNA promoter by the octamer motif defines a new class of RNA polymerase II enhancer elements. Genes Dev. **2**:1764–1778.
64. **Tazi, J., and A. Bird.** 1990. Alternative chromatin structure at CpG islands. Cell **60**:909–920.
65. **Thali, M., M. Muller, M. DeLorenzi, P. Matthias, and M. Bienz.** 1988. Drosophila homeotic genes encode transcriptional activators similar to mammalian OTF-2. Nature (London) **336**:598–601.
66. **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 **330**:1003–1009.
67. **Vidal, M., R. Morris, F. G. Grosveld, and E. Spanopoulou.** 1990. Tissue specific control elements of the Thy-1 gene. EMBO J. **9**:833–840.
68. **Westin, G., and W. Schaffner.** 1988. A zinc-responsive factor interacts with a metal-regulated enhancer element (MRE) of the mouse metallothionein-I gene. EMBO J. **7**:3763–3770.
69. **Whitelaw, E., P. Hogben, O. Hanscombe, and N. J. Proudfoot.** 1989. Transcriptional promiscuity of the human α-globin gene. Mol. Cell. Biol. **9**:241–251.
70. **Williams, A., and J. Gagnon.** 1982. Neuronal cell Thy-1 glycoprotein: homology with immunoglobulin. Science **216**:696–703.
71. **Xiao, J.-H., I. Davidson, M. Macchi, R. Rosales, M. Vigneron, A. Staub, and P. Chambon.** 1987. *In vitro* binding of several cell-specific and ubiquitous nuclear proteins to the GT-I motif of the SV-40 enhancer. Genes Dev. **1**:794–807.