Developmental Regulation of lck Gene Expression in T Lymphocytes

By Robert S. Wildin,*‡ Alex M. Garvin,§ Shashi Pawar,§
David B. Lewis,‡ Kristin M. Abraham,* Katherine A. Forbush,¶
Steven F. Ziegler,§ James M. Allen,§ and Roger M. Perlmutter*§‖¶

From the Departments of *Medicine (Medical Genetics), ‡Pediatrics, §Biochemistry, and
‖Immunology, University of Washington, Seattle, Washington 98195, and ¶Howard Hughes
Medical Institute, Seattle, Washington 98195

Summary

In the mouse and human, mRNA transcripts encoding the lymphocyte-specific protein tyrosine kinase p56^{lck} are derived from two separate promoters resulting in heterogeneity in the 5' untranslated region sequence. The proximal promoter lies just 5' to the coding region for the gene and is active only in thymocytes. In contrast, the distal promoter lies 34 kilobases (kb) 5' in the human, and is active both in thymocytes and mature peripheral T cells. As previously reported, transgenic mice bearing functional proximal promoter sequence juxtaposed with the SV40 large T antigen gene invariably develop lymphoid tumors confined to the thymus. In the current work, transgenic mice bearing a 2.6-kb fragment of the human distal promoter fused to the SV40 large T antigen gene express large T antigen in thymocytes and in peripheral lymphoid cells, and develop tumors of both the thymus and the peripheral lymphoid organs. The ability of the human distal promoter to function appropriately in transgenic mice is consistent with the strong similarity observed between the mouse and human distal promoter sequences. With the exception of a single short interval that serves as a target for binding of nuclear factors, significant sequence similarity is not seen when the distal and proximal promoter sequences are compared. Hence, developmentally regulated, lineage-specific transcription of the lck gene is mediated by distinct promoter sequences that appear to be capable of functioning independently.

The lck gene encodes a membrane-associated protein tyrosine kinase (p56^{lck}) that is expressed in cells of the lymphoid lineage, primarily T cells, lymphoid tumor cell lines, and a few nonlymphoid tumors (1-3). While the precise function of p56^{lck} in T cells remains undefined, it physically associates with the cytoplasmic tails of CD4 and CD8 coreceptor molecules (4-6). Hence it is believed that p56^{lck} may interact with the T cell antigen recognition complex and participate in an antigen-stimulated signal transduction pathway (7).

We and others have shown that lck transcripts derived from certain cultured cell lines are heterogeneous with respect to the nucleotide sequence of the 5' untranslated region (5' UTR)[1] (2, 6, 8-10). The two major transcript types seen in the cell lines, type I and type II, are derived from two alternative promoters. Transcription from a conserved promoter region contiguous with the lck coding sequence (the proximal promoter) results in type I transcripts. Type II transcripts are derived by splicing of a larger transcript originating at another promoter separated by many kb from the coding sequence (the distal promoter) (6, 8, 11). Since the splice ac-

ceptor site for the type II RNA precursor is just 5 bp 5' to the translation initiation codon, almost the entire 5' UTR sequence of type II lck transcripts differs from that of type I.

We do not know whether these results in cultured, often transformed cell lines reflect the true state of transcription in normal cells. Garvin et al. (8) have shown that transcripts containing type I and type II 5' UTRs are both present in normal human thymus RNA, suggesting that alternative promoter usage also occurs in normal human T cells. With this exception, the expression patterns of type I and type II lck mRNAs in normal lymphoid cells have not been examined.

To better define the regulation of lck gene expression, we have cloned genomic DNA fragments containing the human and mouse distal lck promoters, and have sequenced the regions flanking the transcription start sites. While these mouse and human sequences are very similar to one another, they share little similarity with the proximal promoter sequences from either species. Nonetheless, a short region with a high conservation between distal and proximal promoters was identified and found to be sufficient to bind nuclear proteins in a sequence- and tissue-specific manner. We demonstrate that type I and type II transcripts are differentially expressed in

---

[1] Abbreviation used in this paper: UTR, untranslated region.

a developmentally regulated fashion in normal lymphoid cells, and that the proximal and distal promoter elements can function independently in transgenic mice.

Materials and Methods

Genomic Libraries and Library Screening. Human clone SP2 (Fig. 1 *a*) was isolated from a genomic cosmid library (8) by hybridization screening with the human *lck* cDNA clone HK28 (2) using established procedures (8). A map was generated by analysis of restriction fragments and overlapping cosmids. Mouse clone 32.1 (Fig. 1 *b*) containing a 15.5-kb insert was isolated from a Charon 4A phage library of AluI and HaeIII partially digested BALB/c mouse sperm DNA (12) by hybridization screening with a 51 bp synthetic double-stranded oligonucleotide (5'AGGCTAGGAGCATCATG-TGAATAGGCCAGAAGGCTCCCAGGCTGGGCAGG-3') specific for the 5' UTR sequence of mouse type II transcripts (13). Plating of the phage, screening by hybridization, isolation, and large scale culture were performed as previously described (1).

DNA Sequencing and Sequence Analysis. Nucleotide sequence was determined by the Sanger dideoxy chain-termination method (14) using specific oligonucleotide primers applied to both strands of double stranded pUC18 or pUC19 plasmids containing the appropriate SP2 and 32.1 subclones.

Dot matrix comparisons, repeat analysis, and small-scale similarity searches were performed on a microcomputer running Genepro software (Riverside Scientific Enterprises, Seattle, WA). The parameters for dot matrix comparisons were chosen empirically so as to strictly limit the number of matches occurring randomly. Sequence alignments were performed piece-wise in the program Nucaln (D. J. Lipman, Bethesda, MD) using a K-tuple of 3, window size of 20, and gap penalty of 7, and then pieced together manually for best fit. The EMBL19 DNA sequence library was searched within PC-GENE (Intelligenetics, Inc., Mountain View, CA).

RNA Blotting and RNAase Protection. Human peripheral blood T cells were isolated as previously described (15, 16). Human thymocytes were subjected to complement lysis after incubation with either anti-CD4 (66.1), anti-CD8 (OKT8), or both monoclonal antibodies, as described (15). The efficiency of the separation was checked by flow cytometry of the surviving cells stained with conjugated antibodies specific for CD3ε, CD4, CD8, and γδ TCR. Total cellular RNA was prepared from the resulting cells by the guanidinium thiocyanate/cesium chloride method (17). 5 μg of RNA were analyzed in each lane by electrophoresis in 2.2 M formaldehyde/1% agarose gels and blotted as described (18). Blots were hybridized at 42°C in Southern blot buffer with 50% formamide (19) with a denatured synthetic double-stranded 64 bp [^{32}P]-labeled DNA probe specific for the 5' UTRs of type II human *lck* transcripts (bases +42 to +105). After autoradiography, the blot was stripped of probe by boiling and reanalyzed with a 64 bp probe specific for the human type I 5' UTR (bases +41 to +104). The probes were constructed and labeled as described below. After each hybridization, the blot was washed to a final stringency of 50°C in 1× SSC, 0.1% SDS (19). Densitometry was performed on a digitized video image using Visage software (Bio Image, Ann Arbor, MI).

Whole cellular RNA samples from normal mouse thymocytes and LN cells were analyzed by RNAase protection as described (9). The antisense *lck* RNA probe was transcribed from a subcloned SfaN1-EcoRI mouse genomic fragment spanning the 5' UTR of type I *lck* transcripts, the first exon, and the flanking promoter (5') and intron (3') sequences (Fig. 2 *a*). The CD4 RNA probe

was made in a similar fashion from a 145 bp XhoI-PstI fragment from the 5' end of the mouse CD4 cDNA (20).

Gel Mobility Shift Assay. Extracts of nuclear proteins were prepared from the organs of freshly sacrificed normal mice as described (21), except that protease inhibitors were not added. Labeled probe was run on a 0.4 mm thick, 16% polyacrylamide gel and eluted from the gel slice into water. Extracts were incubated with DNA essentially as described (21). Briefly, 2–5 μl of extract was incubated with 1.5 μg of poly dIdC (Sigma Chemical Co., St. Louis, MO) for 15 min at 4°C in a total volume of 8 μl. 10 ng (1 μl) of unlabeled competitor DNA, or water, was then added, and after 10 min at room temperature, 10^4 cpm (~60 pg) of labeled probe was added and the mixture incubated for an additional 25 min. Samples were run on a 0.8 × 140 × 170-mm, 6% polyacrylamide gel at 175 V for 2 h at 4°C with recirculation of the buffer (50 mM Tris-Cl, 1 mM EDTA, pH 8.5), the gel was dried, and autoradiographed.

Oligonucleotide Probes and Competitors. Short, double-stranded DNA fragments were prepared for use as probes or unlabeled competitors by annealing pairs of synthetic DNA oligonucleotides that are complementary at their 3' ends in 50 mM NaCl, 10 mM Tris-Cl, pH 7.5, 6 mM MgCl$_2$, 3 mM dithiothreitol for 15 min at 37°C, then adding the four deoxynucleotide triphosphates to 30 μM and filling in the recessed ends with Klenow fragment at room temperature for 5–10 min. For radiolabeled probes, [^{32}P]dATP was substituted for dATP in the Klenow step, and chased with unlabeled dATP for an additional 5 min. Competitors prepared for use in gel mobility shifts were phenol/chloroform extracted, precipitated with ethanol, and redissolved in water before use.

Transgenic Animals. The human genomic clone SP12 (Fig. 1 *a*) which includes 2.5 kb of the region 5' to the human type II transcription start site (i.e., the distal *lck* promoter) was juxtaposed with a StuI-BamHI fragment of the SV40 genome containing the gene for large and small T antigen (22) to construct a clone called SP12-T. Mice transgenic for this construct were generated as previously described (22). Animals were sacrificed at various ages or when they appeared ill. Tumors were visualized upon gross dissection and confirmed in some cases by histology. Cells from tumors and normal appearing lymphoid organs were suspended in solution, stained with fluorescence-conjugated antibodies to mouse CD4, CD8, and CD3ε, and analyzed for the presence of those molecules and for size (forward scatter) by flow cytometry as previously described (22). Northern blot of tissue RNA from a transgenic mouse was prepared as described above and probed with the StuI-BamHI SV40 T antigen gene fragment labeled by random priming (19).

Results

The Human Distal lck Promoter Resides 34 kb 5' of the Proximal Promoter. Though the presumed mechanism by which type II *lck* transcripts are generated is via splicing of a precursor mRNA derived from the distal promoter, physical linkage of these sequences has not been shown. By screening a human genomic cosmid library, we isolated a single cosmid clone (SP2) containing both type I and type II 5' UTR sequences. Mapping of the entire region by subcloning, restriction enzyme analysis, and sequencing demonstrated that the type II transcription start site is positioned 34 kb 5' to that of the type I *lck* transcripts in human DNA (Fig. 1 *a*). The fidelity of the cosmid was confirmed by genomic blotting analyses (data not shown). The long interval separating these tran-

a

b
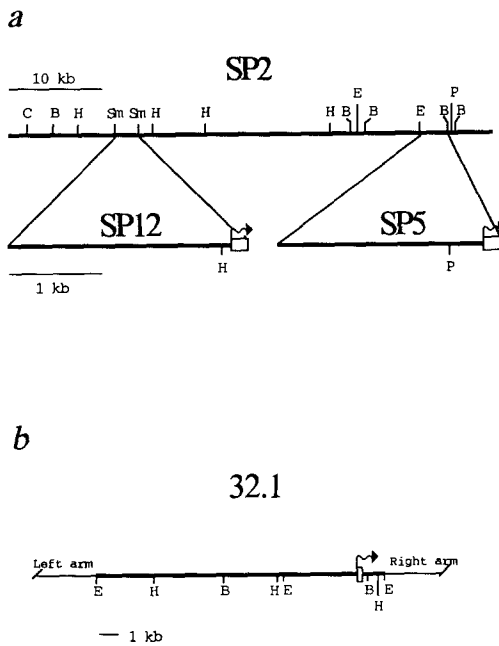
Figure 1. Structure of the human and murine *lck* genomic clones. (*a*) Map of the human genomic cosmid clone SP2 containing the proximal and distal *lck* promoter regions. SP5 and SP12 are 2.1- and 2.6-kb subclones of SP2 containing sequences flanking the proximal and distal promoter transcription start sites, respectively. The cosmid vector arms are not shown. (*b*) Restriction map of the mouse genomic phage clone 32.1 containing the distal *lck* promoter region. Left arm and right arm refer to the phage arms. Open bars represent 5' UTRs. Restriction sites (map in *a* may not include all sites for every enzyme): C, ClaI; B, BamHI; H, HindIII; Sm, SmaI; E, EcoRI; P, PstI.

scriptional units is surprising and suggests that regulatory sequences for each promoter might act independently. Indeed, we have previously demonstrated that a fragment spanning −3280 to +37, with respect to the type I start site, functions as a lymphocyte specific transcriptional regulatory element in transgenic mice (22). To better understand the regulation of *lck* transcription, we obtained a mouse genomic clone (32.1) containing the type II transcription start site (Fig. 1 *b*). Adler et al. (13) previously reported that the murine type II 5' UTR sequence was not present in a clone extending 9 kb 5' of the type I transcription start site, indicating that, in the mouse, the distal promoter lies at least 9 kb 5' to the proximal promoter. Hence the organization of the mouse and human *lck* transcriptional units is grossly conserved. We therefore asked whether type I and type II transcripts accumulate in different lymphoid tissues.

Transcripts from the Proximal and Distal Promoters Are Regulated Differently In Vivo. Thymus, spleen, and PBLs all express *lck* transcripts (1–3, 8, 23) and protein (11, 24). In human thymus, *lck* transcripts containing both type I and type II 5' UTRs can be found (8).

Using a probe specific for the type I 5' UTR and the first exon of the coding region of *lck* (*lck* probe, Fig. 2 *a*) in an RNAase protection assay, mouse thymus RNA protects fragments of two sizes (Fig. 2 *b*, lane *1*). As diagrammed in Fig.

2 *a*, the 294-base fragment corresponds to type I mRNA. The 110-base fragment corresponds to type II transcripts, the sequence of which diverges from that of the probe 5' to the splice acceptor site. Lane *3* of Fig. 2 *b* shows that while type II *lck* transcripts are found in LN cells, type I RNA is conspicuously absent. The 145-base band in lanes *2* and *3* of Fig. 2 *b* results from a CD4-specific probe added to these lanes as an internal control. These observations suggest that mature peripheral T cells contain transcripts derived almost entirely from the distal *lck* promoter. Importantly, similar differences in *lck* expression are found between normal human thymocytes and PBLs: thymocyte RNA hybridizes with probes specific for the 5' UTRs of both type I and type II transcripts, but peripheral lymphocyte RNA hybridizes only to the probe for type II mRNAs (Fig. 3, lanes *1* and *6*).

We conclude that in both man and mouse, type II *lck* transcripts are present in both thymus and peripheral T cells, whereas type I transcripts disappear at some point before the exit of the mature thymocyte from the thymus. While this change in steady-state mRNA abundance could be due to a change in mRNA stability, it more likely reflects a cessation of transcription from the proximal promoter. This interpretation is supported by the fact that the mouse proximal promoter directs the expression of heterologous genes, whose transcript stability is presumably under separate control, to the thymus and not to the peripheral lymphoid organs in transgenic mice (reference 22 and our unpublished results).

Expression of Type I and Type II Transcripts In Human Thymocyte Subsets. Since the proximal promoter is active in the thymus and inactive in mature peripheral T cells, we wondered whether this developmental switch occurred at a particular stage of T cell maturation in the thymus. We examined the expression of type I or type II *lck* transcripts among normal human thymocyte subpopulations defined on the basis of their expression of the T cell surface coreceptor molecules CD4 and CD8 (Fig. 3). Using oligonucleotide probes specific for the 5' UTR of each type of transcript, both type I and type II transcripts were detected in each of the CD4/CD8 thymocyte subsets studied. Although *lck* mRNA levels did not differ greatly among the subsets, there was a trend toward higher expression of type II transcripts in the more mature, CD4$^+$/CD8$^-$ or CD4$^-$/CD8$^+$ (single-positive), thymocytes (Fig. 3, lanes *3* and *4*) than in the immature CD4$^-$/CD8$^-$ (double-negative) thymocytes (Fig. 3, lane *5*). Conversely, the levels of type I mRNA were lower in the more mature thymocyte subtypes, particularly the CD4$^-$ cells (Fig. 3, lane *4*). The technique used to isolate the subsets, complement lysis, did not allow isolation of intermediate maturity CD4$^+$/CD8$^+$ (double-positive) thymocytes from those missing one or both of the CD4 or CD8 molecules. However, double-positive cells constitute the majority of thymocytes and are thus represented in Fig. 3, lanes *1* and *2*, admixed with the other, less abundant subtypes. Furthermore, in other experiments, RNA from human CD4$^+$/CD8$^+$ thymocytes separated by FACS® (Becton Dickinson and Co., Mountain View, CA) or peanut agglutinin contained abundant *lck* mRNA (data not shown). As complement lysis could
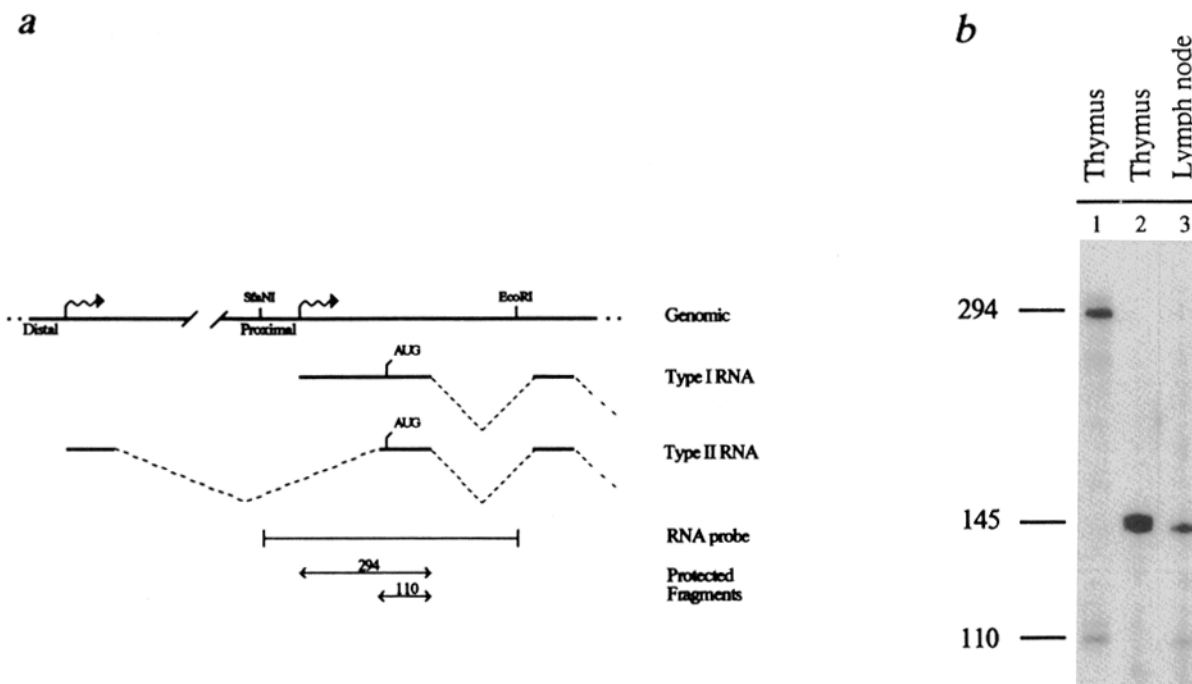
Figure 2. Type I *lck* transcripts accumulate specifically in thymocytes. (*a*) Drawing (not to scale) of the *lck* RNA probe used to perform RNAse protection analysis and the lengths protected by processed type I (294 bases) and type II (110 bases) *lck* mRNAs examined using this probe. The *lck* probe is a 602-base single-stranded uniformly labeled RNA fragment complementary to a mouse genomic fragment that spans the type I 5′ UTR, the first exon of coding sequence, and flanking 5′ (promoter) and 3′ (intron) sequences and delimited by SfaNI and EcoRI restriction sites. Distal and proximal refer to the two *lck* promoter regions. AUG, translation initiation codon. *Wavy arrows*, transcription start sites. (*b*) RNAse protection assay of mouse thymus and LN RNA. Lanes *1* and *3* contain *lck* probe; lanes *2* and *3* contain a CD4 probe (described in Materials and Methods) as a control, 145 bases of which are protected by normal CD4 mRNA.

not eliminate double-negative cells from the single-positives in Fig. 3, lanes *3* and *4*, we calculated the maximum proportion of hybridization signal potentially attributable to double-negative cells using flow cytometry analysis of the contributing pools (see Fig. 3, legend), and densitometry analysis of the blots. Fig. 3, *A* and *B*, lanes *3* and *4*, respectively, contain, at most, 2%, 31%, 0.5%, and 2%, respectively, signal from double-negative cells. Hence, single-positive cells in the thymus do express some type I, as well as type II, *lck* mRNA. We conclude that type I and type II *lck* mRNA is present in each of the human thymocyte subpopulations defined by CD4 and CD8 expression, although we detected only small amounts of type I mRNA in the CD4⁻/CD8⁺ compartment. These results are consistent with the recently reported findings of Reynolds et al. (25), who showed a similar distribution of differential promoter usage in flow cytometry-separated mouse thymocytes and peripheral T cells.

Virtually all peripheral T cells are derived from bone marrow progenitors that mature in the thymus (26). These immature thymocytes pass through a CD4⁺/CD8⁺ stage that gives rise to more differentiated single-positive cells that are the immediate precursors of peripheral T cells (27). Since single-positive thymocytes contain type I *lck* transcripts while peripheral T cells do not (Fig. 2 *b* and Fig. 3), the accumulation of these transcripts must cease in the short interval immediately preceding thymocyte emigration.

Sequence Elements of the Distal lck Promoter Are Highly Conserved. Since the pattern of expression of type I and type II *lck* transcripts is closely conserved in human and mouse lymphoid cells, one might expect to observe conservation in *cis*-linked sequences involved in distal and proximal promoter regulation. Indeed, we have previously reported that the human and mouse proximal promoter sequences contain regions of very high sequence identity (8). To examine the distal promoters for such conserved regions, we sequenced portions of the SP12 cosmid and the phage clone 32.1 containing the human and mouse distal promoters, respectively (Fig. 4).

Alignment of the mouse and human distal promoter sequence reveals a high degree of linear conservation within the first 200 bases 5′ to the transcription start site (Fig. 4). Further aligned similarities are seen 5′ to a 257 bp segment that appears in the mouse but not in the human. Notably, there is also a high degree of sequence conservation in the 5′ UTR for type II *lck* transcripts. Significant similarities between the mouse and human sequences disappear 3′ to the splice donor site (data not shown), indicating that at least the 5′ end of this intron is not conserved.

The Proximal and Distal lck Promoters Share Few Sequence Elements. While a dot matrix comparison of the mouse and human distal *lck* promoter sequences demonstrates their high degree of similarity (Fig. 5 *a*), very little similarity is noted when the mouse distal promoter is compared to the mouse

	Thymocytes					PTC
anti-CD4	−	−	−	+	+	−
anti-CD8	−	−	+	−	+	+
complement	−	+	+	+	+	+

A

B

1　2　3　4　5　6
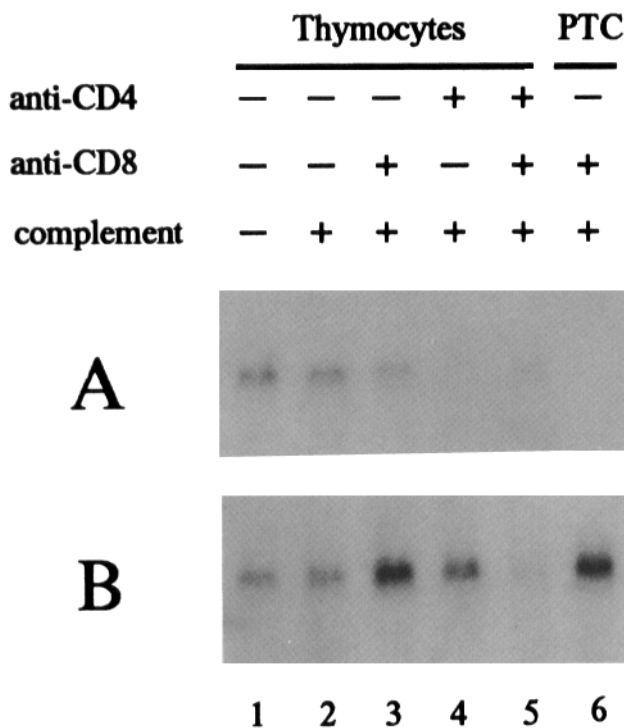
Figure 3. Steady state levels of human type I and type II *lck* mRNA in thymocyte subsets and peripheral T cells (PTC). RNA samples from whole thymus (lanes *1* and *2*), subpopulations of thymocytes isolated by complement lysis with the indicated reagents (lanes *3–5*), and PTC (lane *6*) were analyzed by Northern blotting and sequential probing for the unique 5' UTRs of human type I (*A*) and type II (*B*) *lck* transcripts. The composition of the lysis survivors was assessed by flow cytometry: lanes *1* and *2*, 74% CD4+, 75% CD8+, 76% CD3+; lane *3*, 95% CD4+, 8% CD8dull, 96% CD3+; lane *4*, 3% CD4dull, 91% CD8+, 99% CD3+; lane *5*, 10% CD4very dull, 19% CD8very dull, 51% CD3+, 37% γδ TCR+.
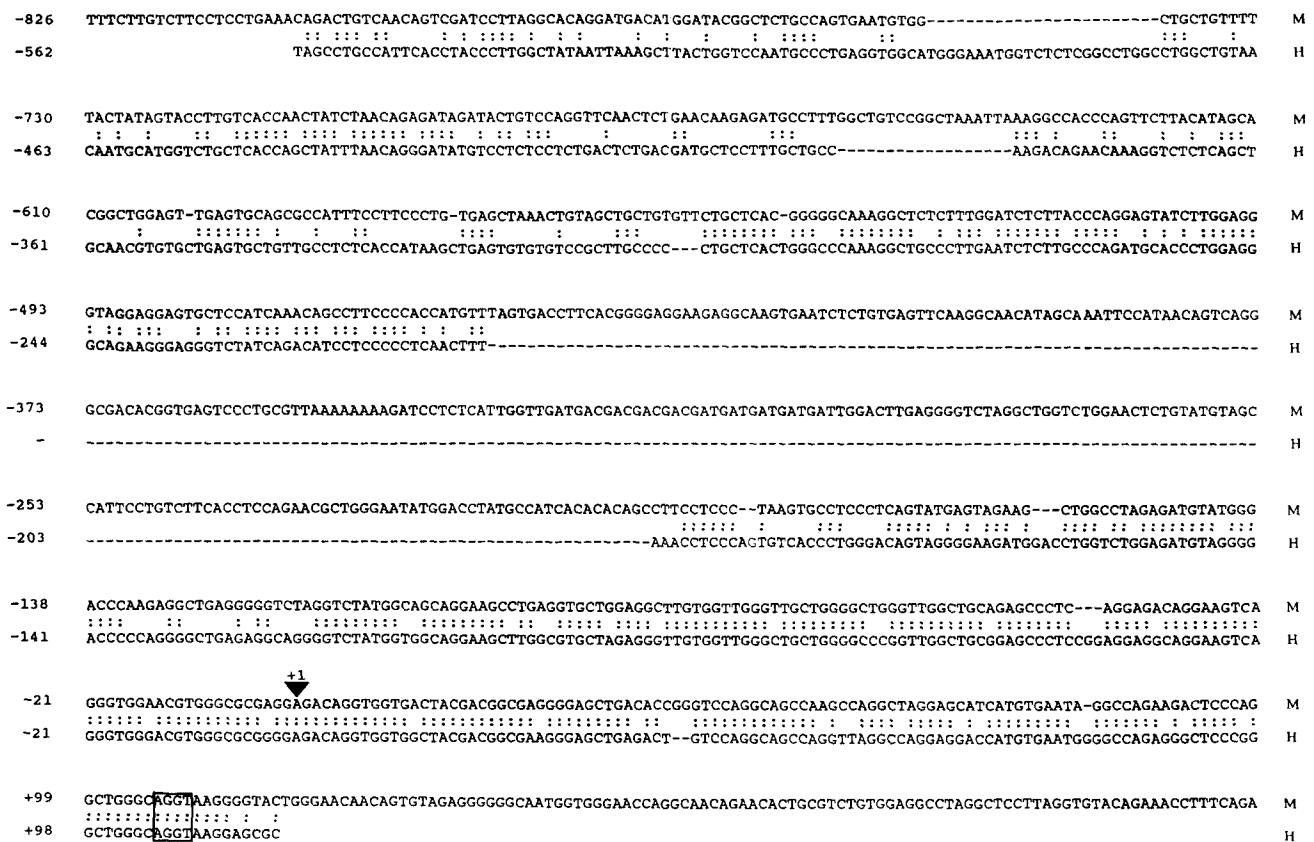
proximal promoter (Fig. 5 *b*). Nevertheless, careful comparison revealed a sequence element in the distal promoter regions of both the mouse and human that is also present in the proximal promoter sequences of both species (Fig. 6 *a*). This element lacks similarities to the core consensus sequences of any previously described *cis*-regulatory elements of which we are aware, but its "core" (5'-TGCTGGGG-3') does have some similarities to sequences within regions known to bind the murine MHC class I transcription factor H2TF-1 (includes 5'-GGCTGGGG-3') (28), as well as a chicken lysozyme silencer (includes 5'-AGCTGGGG-3') (29). A computer search for the "core" 8-mer of the element among lymphoid-specific promoters in the EMBL 19 DNA database failed to find an entry whose sequence conformed to all of the conserved regions of the entire element. We also noted that the sequence surrounding the element in the mouse distal promoter shows a curiously symmetric series of nested direct repeats (Fig. 6 *b*). Most of these repeats are also found in the same positions in the human distal promoter sequence. These observations are provocative and serve to focus attention on this region as a putative binding site for factors regulating expression of the *lck* promoter (see below). Several other short sequence

segments (8–11 bp in length) are shared between the proximal and distal promoters within a species, but unlike the element in Fig. 6 *a* are not shared by both species (data not shown).

We further examined the human and mouse proximal and distal *lck* promoter sequences for the presence of a variety of previously identified eukaryotic *cis*–regulatory sequences. We found that the human and mouse proximal promoters contain a consensus sequence (5'-GGAAGTG-3') for E4TF-1, a transcriptional regulator of adenovirus type 5 early region 4 (30, 31), located some 50 bases 5' to the IgH core enhancer sequence previously noted by both Takadera et al. (6) and Adler et al. (13). A target sequence (5'-GTCGTCA-3') for the cAMP response element binding protein was found in the inverse orientation in the mouse distal promoter (bases −318 to −324). The so-called PU-box (5'-GAGGAA-3'), an enhancer bound by the PU.1 protein (the protooncogene Spi-1, reference 32) in B lymphocytes and macrophages (33) is found three times in the mouse distal promoter (−189 to −194, −436 to −431, and −812 to −817) but not in the human distal promoter. CCAAT and TATAA elements are present in the human distal promoter sequence, but these are positioned far from the transcription start sites (bases −516 and −536, respectively), and thus probably represent chance occurrences.

The Conserved Sequence Element Is a Target for Nuclear Protein Binding. To examine the possibility that the conserved distal/proximal sequence element depicted in Fig. 6 *a* is an important functional element, we tested its ability to specifically bind components of nuclear protein extracts that presumably contain DNA-binding transcription factors, in a gel mobility shift assay. Fig. 6 *c* demonstrates that a protein (or protein complex) in lymphoid cell nuclei binds specifically to a 29 bp probe spanning this region of the mouse distal promoter (Fig. 6 *a*, MD). This binding (Fig. 6 *c*, lanes *1*, *8*, and *15*) is specifically competed by an excess of unlabeled probe (the MD competitor; Fig. 6 *c*, lanes *3*, *10*, and *17*), but not by the unrelated sequence OCT (Fig. 6 *c*, lanes *2*, *9*, and *16*) that includes Oct-1 binding motifs (34) and also binds nuclear proteins in this assay (data not shown). Binding is also not competed by a longer DNA fragment that contains sequences flanking the mouse distal promoter transcription start site (bases −38 to +44, data not shown). Hence, this nuclear extract–induced gel mobility shift is specific for the probe.

The sequence dependence of the binding was further examined by competing with an excess of cold competitors representing the similar sequences found in the human distal (HD, differing by four bases from the MD probe), the mouse proximal (MP, differing by eight bases), and two sequences from the human proximal (HPa and HPb, differing by 14 and 21 bases, respectively) *lck* promoters (Fig. 6 *a*). (HPb is quite divergent from MD but has more similarity to MP.) As seen in Fig. 6 *c* (lanes *4*, *11*, and *18*), the very similar sequence from the human distal promoter, HD, inhibited MD probe binding as effectively as unlabeled MD sequence. The MP competitor failed to compete significantly in the experi-

```
-826  TTTCTTGTCTTCCTCCTGAAACAGACTGTCAACAGTCGATCCTTAGGCACAGGATGACA1GGATACGGCTCTGCCAGTGAATGTGG---------------------CTGCTGTTTT  M
             :: ::: ::     : : :::: : : :     :   :   : ::::    ::                                                :::  :
-562                         TAGCCTGCCATTCACCTACCCTTGGCTATAATTAAAGCTTACTGGTCCAATGCCCTGAGGTGGCATGGGAAATGGTCTCTCGGCCTGGCCTGGCTGTAA  H


-730  TACTATAGTACCTTGTCACCAACTATCTAACAGAGATAGATACTGTCCAGGTTCAACTCTGAACAAGAGATGCCTTTGGCTGTCCGGCTAAATTAAAGGCCACCCAGTTCTTACATAGCA  M
        :  :    ::  :::::: ::::  ::::::: ::::  : :: :::     :    ::       :::                        :::: :     ::    :  :  :::
-463  CAATGCATGGTCTGCTCACCAGCTATTTAACAGGGATATGTCCTCTCCTCTGACTCTGACGATGCTCCTTTGCTGCC----------------AAGACAGAACAAAGGTCTCTCAGCT  H


-610  CGGCTGGAGT-TGAGTGCAGCGCCATTTCCTTCCCTG-TGAGCTAAACTGTAGCTGCTGTGTTCTGCTCAC-GGGGGCAAAGGCTCTCTTTGGATCTCTTACCCAGGAGTATCTTGGAGG  M
        :     :::::::: : : :::     :::: ::    ::: :::    :::::::: :::  :::::::: : ::: ::::::: :::::  : : ::::::::
-361  GCAACGTGTGCTGAGTGCTGTTGCCTCTCACCATAAGCTGAGTGTGTGTCCGCTTGCCCC---CTGCTCACTGGGCCCAAAGGCTGCCCTTGAATCTCTTGCCCAGATGCACCCTGGAGG  H


-493  GTAGGAGGAGTGCTCCATCAAACAGCCTTCCCCACCATGTTTAGTGACCTTCACGGGGAGGAAGAGGCAAGTGAATCTCTGTGAGTTCAAGGCAACATAGCAAATTCCATAACAGTCAGG  M
        : :: :::   : ::  ::::  :::  :: :::  :::: :  : ::
-244  GCAGAAGGGAGGGTCTATCAGACATCCTCCCCCTCAACTTT-----------------------------------------------------------------------------  H


-373  GCGACACGGTGAGTCCCTGCGTTAAAAAAAAAGATCCTCTCATTGGTTGATGACGACGACGACGATGATGATGATGATTGGACTTGAGGGGTCTAGGCTGGTCTGGAACTCTGTATGTAGC  M
-     ----------------------------------------------------------------------------------------------------------------------  H


-253  CATTCCTGTCTTCACCTCCAGAACGCTGGGAATATGGACCTATGCCATCACACACAGCCTTCCTCCC--TAAGTGCCTCCCTCAGTATGAGTAGAAG---CTGGCCTAGAGATGTATGGG  M
                                                                       :::::: :    ::: :::::: : : ::: :   :::: :: ::::::::: :::
-203  --------------------------------------------------------AAACCTCCCAGTGTCACCCTGGGACAGTAGGGGAAGATGGACCTGGTCTGGAGATGTAGGGG  H


-138  ACCCAAGAGGCTGAGGGGGTCTAGGTCTATGGCAGCAGGAAGCCTGAGGTGCTGGAGGCTTGTGGTTGGGTTGCTGGGGCTGGGTTGGCTGCAGAGCCCTC---AGGAGACAGGAAGTCA  M
        ::::   ::     : ::       ::::::::: ::::::::: ::  :::::  ::::   ::::::::::: ::::::::  ::::::::::: :::::::    :::::  :::::::::::
-141  ACCCCCAGGGGCTGAGAGGCAGGGGTCTATGGTGGCAGGAAGCTTGGCGTGCTAGAGGGTTGTGGTTGGGCTGCTGGGGCCCGGTTGGCTGCGGAGCCCTCCGGAGGAGGCAGGAAGTCA  H


            +1
             ▼
-21   GGGTGGAACGTGGGCGCGAGGAGACAGGTGGTGACTACGACGGCGAGGGGAGCTGACACCGGGTCCAGGCAGCCAAGCCAGGCTAGGAGCATCATGTGAATA-GGCCAGAAGACTCCCAG  M
        ::::::  ::::::::::::: :::::::::::::: :::::::::::::: :::::::::: ::       :::::::::::::: :   :::: ::::: : :::::::::  ::::::: : :::::  :
-21   GGGTGGGACGTGGGCGCGGGGAGACAGGTGGTGGCTACGACGGCGAAGGGAGCTGAGACT--GTCCAGGCAGCCAGGTTAGGCCAGGAGGACCATGTGAATGGGGCCAGAGGGCTCCCGG  H


+99   GCTGGGC[AGGT]AAGGGGTACTGGGAACAACAGTGTAGAGGGGGGCAATGGTGGGAACCAGGCAACAGAACACTGCGTCTGTGGAGGCCTAGGCTCCTTAGGTGTACAGAAACCTTTCAGA  M
      :::::::::::::::: : :
+98   GCTGGGC[AGGT]AAGGAGCGC  H
```

Figure 4.   Extended sequence similarity in the human and mouse distal *lck* promoter regions. The mouse sequence is at the top of each comparison. Bases that match are linked by a colon. Sequences are 5' to 3'. Numbers are with respect to the predominant (IIc) transcription start site (*filled triangle*) reported by Takadera et al. (6). The splice donor site is boxed. These sequence data have been assigned GenBank accession numbers M36823 (mouse) and M36824 (human).

ment in Fig. 6 *c* (lanes 5 and 12) but did compete successfully at much higher molar ratio (data not shown) in a sequence-specific manner. This observation suggests that the DNA-binding protein complex has some specificity for the MP sequence, but binds with a much lower affinity. The two sequences from the human proximal promoter were unable to compete effectively (Fig. 6 *c*, lanes 5, 6, 13, and 14).

The gel mobility shifts in thymus have a slower mobility than the otherwise similar shifts seen in spleen or liver, though their sequence-specificity is the same (Fig. 6 *c*). Hence, the nature of the complex found in thymus is different than that in total spleen, or in liver. By varying the conditions of the gel shift assay, the mobility differences noted between thymus and spleen or liver extracts are much more pronounced, indicating that this is a real and consistent difference (data not shown).

We conclude that both lymphoid and nonlymphoid organs contained nuclear proteins or complexes that specifically bind the sequence element highlighted by the sequence comparisons, that the nature of the complexes varies by organ type, and that the binding affinity in the gel mobility shift assay is sensitive to increasing variations from the MD target sequence.

The Human Distal Promoter Functions Independently in Transgenic Mice.   Sequence analysis showed that *cis* promoter elements of the *lck* distal promoter have been conserved in evolution and suggested that the corresponding *trans*-acting factors responsible for tissue-specific and developmental regulation of the distal promoter may also have been conserved. We tested this hypothesis by assessing the ability of the human distal promoter to function in a heterologous background, i.e., by using it to drive a reporter gene in transgenic mice. Animals bearing a transgene (SP12-T) consisting of a human genomic segment containing 2.5 kb of sequence 5' to the type II *lck* mRNA transcription start site (subclone SP12, Fig. 1) driving the SV40 large T antigen gene expressed SV40 T mRNA in thymus and spleen, and faintly in heart, but not in liver or kidney (Fig. 7). The very small amount of T antigen transcript detected in heart may represent that derived from the blood lymphocytes contained within the heart at the time of dissection rather than from the myocardium. Starting at 21 wk of age, mice bearing this transgene developed hemorrhagic tumors of the LN and spleen and gross enlargement of the thymus. Tumors were not noted in the heart or in any other non-lymphoid organs. Thus, unlike previously described transgenic mice bearing a similar T antigen trans-

a

b
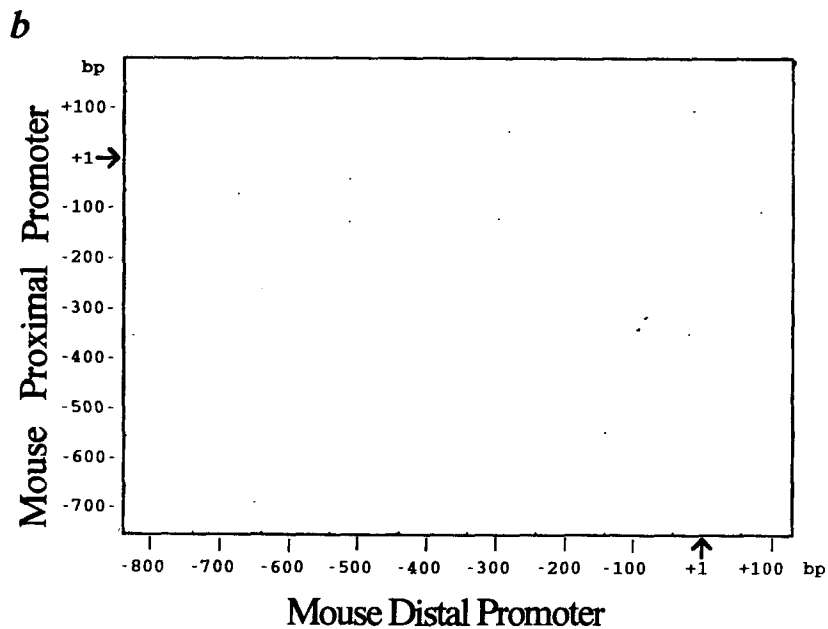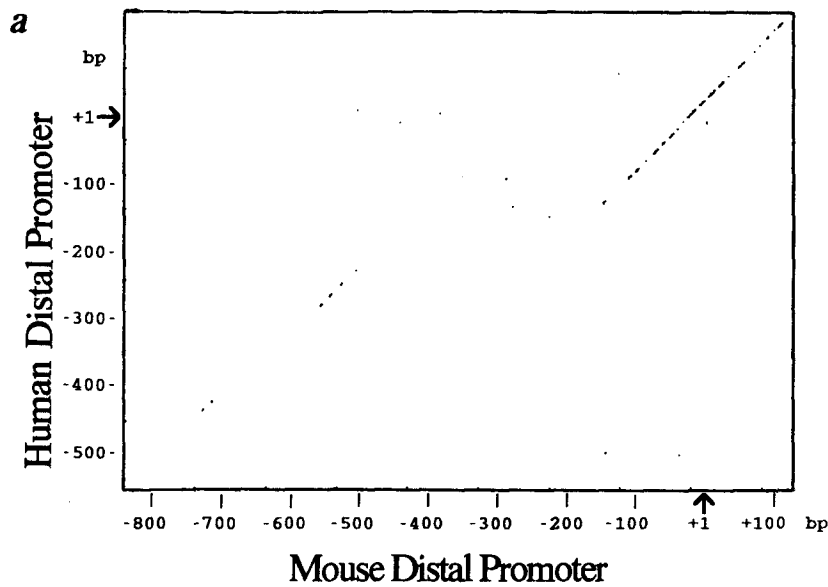
Figure 5. The proximal and distal *lck* promoters share little sequence similarity. (*a*) Dot matrix comparison of the mouse and human distal *lck* promoter region sequences. Each dot represents a match of 12 out of 16 bases. The 5′ end of each sequence begins at the lower left and the numbers are bases with respect to the transcription start sites (*arrows*). (*b*) Dot matrix comparison of the mouse proximal and mouse distal *lck* promoter sequences. The former sequence is taken from Garvin et al. (8). Parameters are as in *a*.

gene driven by the mouse proximal *lck* promoter, which develop only thymic tumors (22), the human distal promoter directed T antigen mRNA expression and T antigen-induced transformation to the peripheral lymphoid organs as well as to the thymus. The behavior of the distal and proximal promoters in transgenic mice thus conforms to the expression patterns of type I and type II *lck* transcripts demonstrated in Figs. *2* and *3*.

Discussion

We have demonstrated that the *lck* gene can be transcribed via two separate promoters that behave in a developmentally

regulated fashion: while both promoters are active in thymocytes, only the distal promoter is active in mature T cells. One may conceive of three explanations for the presence of dual *lck* promoters. First, the simple requirement for p56^{lck} in both thymocytes and peripheral T cells may have driven conservation of a "constitutive" promoter (the distal promoter), while the presence and activity of the proximal promoter is genetically accidental and superfluous. The high degree of sequence conservation noted between the human and mouse, the persistence of a functional proximal promoter in both species, and the lack of major sequence similarities between the distal and proximal promoter sequences argue against this hypothesis.

a

```
-83    AGGCTTGTGGTTGGGTTGCTGGGGCTGGG       Mouse Distal    (MD)
-352   GT-G-G-G---A---G--------T----       Mouse Proximal  (MP)
-86    ---G-----------C---------CC--       Human Distal    (HD)
-420   TCA--G-A--G-TTTC--------AGAA-       Human Proximal a (HPa)
-347   GT-AGG-GAACA-C-GG-GCTT-ATG-T-       Human Proximal b (HPb)
```

b

-103
|
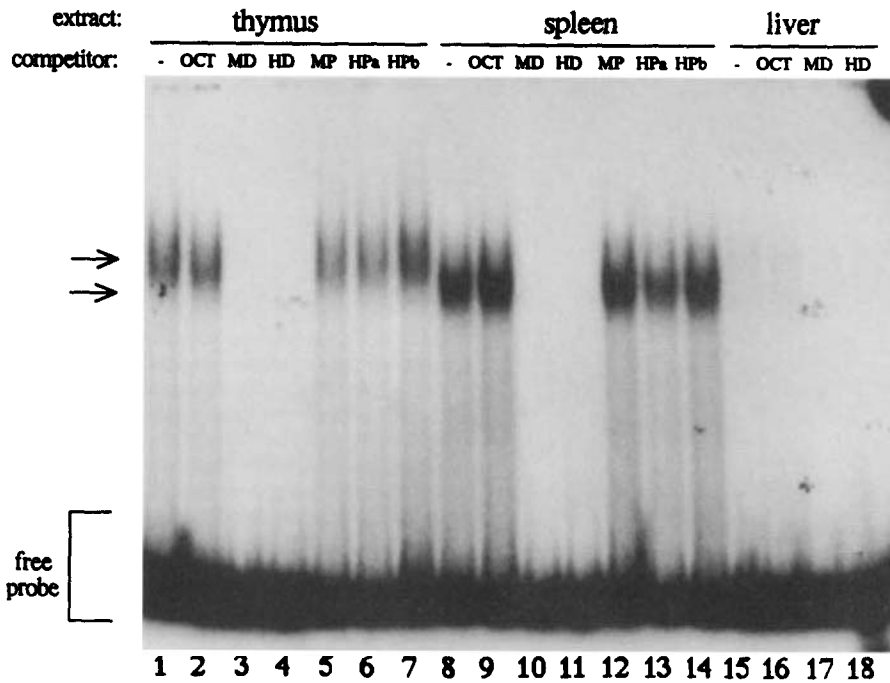CAGGAAGCCTGAGGTGCTGGAGGCTTGTGGTTGGGTTGCTGGGGCTGGGTTGGCTGCAGAGCCCTCAGGAGACAGGAA

-24
|

c

Figure 6. An element conserved in the proximal and distal *lck* promoters binds nuclear factors. (*a*) Alignment of a promoter element found in four *lck* promoter sequences (Fig.4, and reference 8). (–) Bases that match the mouse distal promoter sequence (*top line*). MD, MP, HD, HPa, and HPb: sequence of double-stranded DNA segments used as probe and competitors in *c*. (*b*) Sequence of the mouse distal *lck* promoter surrounding the conserved element in *a*. Nested direct repeats surrounding the core sequence (*bold type*) of the element are underlined and connected by a dotted line. Nucleotide position with respect to the transcription start site is noted. (*c*) Gel mobility shift assay. Nuclear protein extracts were incubated with a double-stranded 29-bp radiolabeled DNA probe (MD, see *a*), with or without the prior addition of unlabeled competitors. The resulting protein:DNA complexes were separated from the faster moving free DNA probe ("free probe") on a native polyacrylamide gel. Binding of the probe by proteins with nonspecific affinity for DNA was reduced by the prior addition of 1.5 μg of poly dIdC per reaction. Double-stranded DNA competitors: -, none; OCT, (5'-GCAGAAGTATGCAAAGCA-TGCATCTCAATTAGTCAG-3'), from the SV40 enhancer B element, target for the octamer binding protein, OBP-100 (34); See *a* for competitor designations. Upper and lower arrows indicate bands of different mobility seen in thymus, spleen, or liver, respectively.
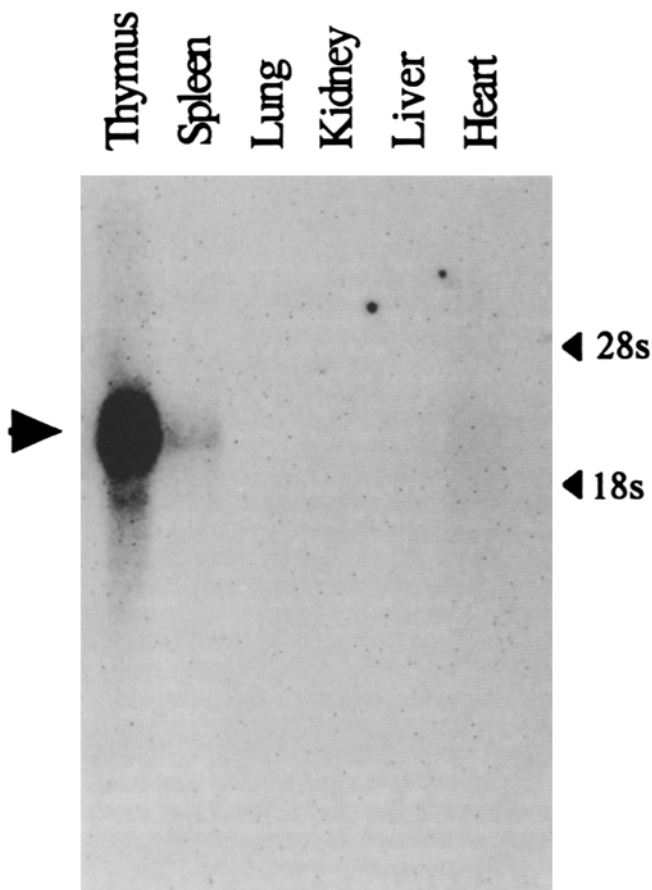
The dual promoter system might also reflect the need for continued expression of *lck* in the context of a changing transcriptional milieu that could not support activation from a single set of regulatory elements under all conditions. In this model, the changing milieu would be driven not by the need for modified *lck* expression but by the regulation of other genes with related transcriptional control elements that must be modulated in the process of thymocyte maturation. Contradicting this hypothesis, we have shown that thymocytes of all developmental stages as well as peripheral lymphocytes transcribe effectively from the distal promoter.

It seems much more likely that the dual *lck* promoter system evolved as a mechanism for allowing differential regulation under different developmental conditions: i.e., there is an advantage to be gained from controlling *lck* expression in peripheral lymphocytes in a manner different from that in thymocytes undergoing selection and maturation.

While the nature of this advantage is obscure, it is impor-

Thymus Spleen Lung Kidney Liver Heart

◀ 28s

◀ 18s

Figure 7. Expression of the SP12-T transgene in lymphoid tissues. Shown is an RNA blot performed using samples derived from the indicated tissues of a 6.5-wk-old SP12-T transgene mouse. The blot was hybridized with a probe specific for SV40-T antigen sequences (see Materials and Methods).

tial stability or efficiency of translation. Indeed, we have previously demonstrated that murine type I transcripts contain AUG codons within their 5' UTRs that reduce the translational efficiency of these mRNAs (36). The human and mouse type II transcripts also have a single out-of-frame AUG codon in their 5' UTRs, but the effect of this sequence on translation has not been tested.

Dual promoters have been described in several other eukaryotic contexts. In some cases, the differential regulation of the two promoters appears to be due to stage- or organ-specific factors (35, 37). In others, transcription from the distal promoter interferes with transcription from the proximal promoter, resulting in mutually exclusive promoter usage (38, 39). Since the thymocyte subsets we studied all contain transcripts from both the proximal and distal *lck* promoters, mutually exclusive usage due to transcriptional interferences is unlikely. Instead, the lack of sequence similarity between the proximal and distal promoters suggests control via different *trans*-acting factors having different target sequences.

The mouse and human distal *lck* promoters share a high degree of primary sequence similarity. The validity of inferring functional importance for regions of the promoter whose sequences have been conserved in evolution depends on the assumption that the *trans*-acting regulatory factors and their binding specificities are also conserved. Our ability to target SV40 T antigen–induced tumor formation to T cells found in the peripheral lymphoid organs using the human *lck* distal promoter in transgenic mice shows that the human *cis* elements are sufficient to drive appropriate tissue-specific expression in a mouse milieu, lending credence to this approach.

The significance of the occurrence of canonical binding sequences for E4TF1, CREB, and the PU.1 protein is questionable as none of them has been implicated in T cell-specific transcriptional regulation, and the sites for the latter two are not conserved between species. Since both the proximal and distal *lck* promoters lack other convincing sequence elements previously identified as the binding sites for *trans*-acting regulatory proteins, study of the regulation of the *lck* promoters is likely to reveal novel regulators of transcription. Indeed, Fig. 6 indicates that the conserved sequence element serves as a target for binding of nuclear proteins, and the behavior of these proteins is different in thymocytes, splenocytes, and hepatocytes. Detailed dissection of transcriptional control mechanisms regulating *lck* expression will almost certainly provide insights into the control of thymocyte development via gene activation, and may help define the function of p56[lck] in T cells.

tant to recognize that type I and type II transcripts encode identical proteins. Hence, the advantageous differential control must act at or prior to translation, i.e., on the transcription rate, nascent transcript processing, mRNA stability, or translational efficiency. Differential transcriptional control could simply reflect the need for more *lck* mRNA in developing thymocytes, thus requiring a stronger promoter, as in the mouse α-amylase dual promoter system (35). Alternatively, response to certain stimuli, e.g., T cell activation, might differ between the two *lck* promoters. The quite dissimilar yet evolutionarily conserved sequences of the 5' UTRs of type I and type II *lck* transcripts suggest the potential for differen-

---

References

1. Marth, J.D., R. Peet, E.G. Krebs, and R.M. Perlmutter. 1985. A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpressed in the murine T cell lymphoma LSTRA. Cell. 43:393.

2. Perlmutter, R.M., J.D. Marth, D.B. Lewis, R. Peet, S.F. Ziegler, and C.B. Wilson. 1988. Structure and expression of lck transcripts in human lymphoid cells. J. Cell. Biochem. 38:117.

3. Koga, Y., N. Kimura, J. Minowada, and T.W. Mak. 1988. Expression of the human T cell-specific tyrosine kinase YT16 (lck) message in leukemic T cell lines. Cancer Res. 48:856.

4. Turner, J.M., M.H. Brodsky, B.A. Irving, S.D. Levin, R.M. Perlmutter, and D.R. Littman. 1990. Interaction of the unique N-terminal region of tyrosine kinase $p56^{lck}$ with cytoplasmic domains of CD4 and CD8 is mediated by cysteine motifs. Cell. 60:755.

5. Veillette, A., J.B. Bolen, and M.A. Bookman. 1989. Alterations in tyrosine protein phosphorylation induced by antibody-mediated cross-linking of the CD4 receptor of T lymphocytes. Mol. Cell. Biol. 9:4441.

6. Takadera, T., S. Leung, A. Gernone, Y. Koga, Y. Takihara, N.G. Miyamoto, and T.W. Mak. 1989. Structure of the two promoters of the human lck gene: differential accumulation of two classes of lck transcripts in T cells. Mol. Cell. Biol. 9:2173.

7. Perlmutter, R.M. 1989. T cell signalling. Science (Wash. DC). 245:344.

8. Garvin, A.M., S. Pawar, J.D. Marth, and R.M. Perlmutter. 1988. Structure of the murine lck gene and its rearrangement in a murine lymphoma cell line. Mol. Cell. Biol. 8:3058.

9. Voronova, A.F., H.T. Adler, and B.M. Sefton. 1987. Two lck transcripts containing different 5' untranslated regions are present in T cells. Mol. Cell. Biol. 7:4407.

10. Sartor, O., F.S. Gregory, N.S. Templeton, S. Pawar, R.M. Perlmutter, and N. Rosen. 1989. Selective expression of alternative lck mRNAs in human malignant cell lines. Mol. Cell. Biol. 9:2983.

11. Perlmutter, R.M., J.D. Marth, S.F. Ziegler, A.M. Garvin, S. Pawar, M.P. Cooke, and K.M. Abraham. 1988. Specialized protein tyrosine kinase proto-oncogenes in hematopoietic cells. Biochim. Biophys. Acta. 948:245.

12. Perlmutter, R.M., B. Berson, J.A. Griffin, and L. Hood. 1985. Diversity in the germline antibody repertoire. J. Exp. Med. 162:1998.

13. Adler, H.T., P.J. Reynold, C.M. Kelly, and B.M. Sefton. 1988. Transcriptional activation of lck by retrovirus promoter insertion between two lymphoid-specific promoters. J. Virol. 62:4113.

14. Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74:5463.

15. Lewis, D.B., K.S. Prickett, A. Larsen, K. Grabstein, M. Weaver, and C.B. Wilson. 1988. Restricted production of interleukin 4 by activated human T cells. Proc. Natl. Acad. Sci. USA. 85:9743.

16. Wilson, C.B., J. Westall, L. Johnston, D.B. Lewis, S.K. Dower, and A.R. Alpert. 1986. Decreased production of interferon γ by human neonatal cells: intrinsic and regulatory deficiencies. J. Clin. Invest. 77:860.

17. Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry. 18:5294.

18. Lewis, D.B., A. Larsen, and C.B. Wilson. 1986. Reduced interferon-γ mRNA levels in human neonates. J. Exp. Med. 163:1018.

19. Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.

20. Littman, D.R., and S.N. Gettner. 1987. Unusual intron in the immunoglobulin domain of the newly isolated murine CD4 (L3T4) gene. Nature (Lond.). 325:453.

21. Buskin, J.N., and S.D. Hauschka. 1989. Identification of a monocyte nuclear factor that binds to the muscle-specific enhancer of the mouse muscle creatine kinase gene. Mol. Cell. Biol. 9:2627.

22. Garvin, A.M., K.M. Abraham, K.A. Forbush, A.G. Farr, B.L. Davidson, and R.M. Perlmutter. 1990. Disruption of thymocyte development and lymphomagenesis induced by SV40 T-antigen. International Immunology. 2:173.

23. Marth, J.D., D.B. Lewis, C.B. Wilson, M.E. Gearn, E.G. Krebs, and R.M. Perlmutter. 1987. Regulation of $pp56^{lck}$ during T cell activation: functional implications for the src-like protein tyrosine kinases. EMBO (Eur. Mol. Biol. Organ.) J. 6:2727.

24. Marth, J.D., D.B. Lewis, M.P. Cooke, E.D. Mellins, M.E. Gearn, L.E. Samelson, C.B. Wilson, A.D. Miller, and R.M. Perlmutter. 1989. Lymphocyte activation provokes modification of a lymphocyte-specific protein tyrosine kinase ($p56^{lck}$). J. Immunol. 142:2430.

25. Reynolds, P.J., J.L. Lesley, J. Trotter, R. Schulte, R. Hyman, and B.M. Sefton. 1990. Changes in the relative abundance of type I and type II lck mRNA transcripts suggest differential promoter usage during T cell development. Mol. Cell. Biol. 10:4266.

26. Scollay, R., A. Wilson, A. D'Amico, K. Kelly, M. Egerton, M. Pearse, L. Wu, and K. Shortman. 1988. Developmental status and reconstitution potential of subpopulations of murine thymocytes. Immunol. Rev. 104:81.

27. von Boehmer, H. 1988. Developmental biology of T lymphocytes. Annu. Rev. Immunol. 6:309.

28. Baldwin, A.S., Jr., and P. Sharp. 1987. Binding of a nuclear factor to a regulatory sequence in the promoter of the mouse $H\text{-}2k^b$ class I major histocompatibility gene. Mol. Cell. Biol. 7:305.

29. Baniahmad, A., C. Steiner, A.C. Köhne, and R. Renkawitz. 1990. Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. Cell. 61:505.

30. Jones, N.C., P.W.J. Rigby, and E.B. Ziff. 1988. Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes & Dev. 2:267.

31. Watanabe, H., T. Wada, and H. Handa. 1990. Transcription factor E4TF1 contains two subunits with different functions. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:841.

32. Goebl, M.G. 1990. The PU.1 transcription factor is the product of the putative oncogene *Spi-1* (letter). *Cell.* 61:1165.

33. Klemsz, M.J., S.R. McKercher, A. Celada, C.V. Beveren, and R.A. Maki. 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. *Cell.* 61:113.

34. Sturm, R., T. Baumruker, J.B. Robert Franza, and W. Herr. 1987. A 100-kD HeLa cell octamer binding protein (OBP100) interacts differently with two separate ocatamer-related sequences within the *SV40* enhancer *Genes & Dev.* 1:1147.

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

36. Marth, J.D., R.W. Overell, K.E. Meier, E.G. Krebs, and R.M. Perlmutter. 1988. Translational activation of the *lck* proto-oncogene. *Nature (Lond.).* 332:171.

37. Colbert, M.C., and E. Ciejek-Baez. 1988. Alternative promoter usage by aldolase A during in vitro myogenesis. *Dev. Biol.* 130:392.

38. Woisetschlaeger, M., C.N. Yandava, L.A. Furmanski, J.L. Strominger, and S.H. Speck. 1990. Promoter switching in Epstein-Barr virus during the initial stages of infection of B lymphocytes. *Proc. Natl. Acad. Sci. USA.* 87:1725.

39. Corbin, V., and T. Maniatis. 1989. Role of transcriptional interference in the Drosophila melanogaster *Adh* promoter switch. *Nature (Lond.).* 337:279.